When less is more - Endogenous tagging with TurboID increases the sensitivity of proximity labelling-based experiments

Alexander Stockhammer[1], Laila S. Benz[1], Christian Freund[2], Benno Kuropka[2] and Francesca Bottanelli[1,*]

[1]Membrane trafficking laboratory, Institute for Chemistry and Biochemistry, Freie Universität Berlin, Thielallee 63, 14195 Berlin

[2]Laboratory of Protein Biochemistry, Institute for Chemistry and Biochemistry, Freie Universität Berlin, Thielallee 63, 14195 Berlin

[*]Author to whom any correspondence should be addressed: Francesca.bottanelli@fu-berlin.de

Abstract

In recent years, proximity labelling has established itself as an unbiased and powerful approach to map the interactome of specific proteins. Generally, protein fusions with labelling enzymes are transiently overexpressed to perform these experiments. Using a pipeline for the rapid generation CRISPR-Cas9 knock-ins (KIs) based on antibiotic selection, we were able to compare the performance of commonly used labelling enzymes when endogenously expressed. We found TurboID and its shorter variant miniTurboID to be superior above other labelling enzymes at physiological expression levels. Endogenous tagging of the μ subunit of the AP-1 complex increased the sensitivity for detection of interactors in a proximity labelling experiment and resulted in a more comprehensive mass spectrometry data set. We were able to identify several known interactors of the complex and cargo proteins that simple overexpression of a labelling enzyme fusion protein could not reveal. Our approach greatly simplifies the execution of proximity labelling experiments for proteins in their native cellular environment and allows going from CRISPR transfection to mass spectrometry data in just over a month.

Introduction

The biological role of a protein is shaped by its subcellular localisation and its interaction with other biomolecules. Therefore, mapping the interactors of a given protein can be crucial for understanding its biological function. In the past several years, proximity labelling with biotin in living cells has emerged as a complementary approach to classic affinity purification/mass spectrometry (AP/MS)-based methods for mapping protein-protein interactions in living cells and organisms[1,2]. The proximity labelling is carried out by enzymes genetically fused to the protein of interest (POI) that catalyse the formation of a highly reactive biotin intermediate labelling proteins within a small radius (1-10 nm)[3,4] in a promiscuous manner. A key advantage of proximity labelling-based interactome mapping compared to traditional approaches is that also very weak and transient interactions can be captured. The biotinylation itself provides a unique chemical moiety that can be used for subsequent enrichment and identification.

The enzymes used for proximity labelling are either engineered peroxidases (APEX[5], APEX2[6]) or engineered biotin ligases (BioID[7], BioID2[8], TurboID[9] and miniTurboID[9]). APEX and APEX2 use H_2O_2 as co-substrate to rapidly generate a highly reactive phenoxyl radical from biotin-phenol that reacts specifically with electron-rich side chains (primarily tyrosine)[9]. An attractive feature of APEX peroxidases are the fast labelling kinetics (labelling time: <1 min) that enables probing with a high temporal resolution. However, on the downside, they require H_2O_2 that causes oxidative stress in living cells and thus cannot be used for proximity labelling in living organisms. In contrast, biotin ligases simply need non-toxic, highly soluble biotin as a substrate

which in an ATP-dependent reaction is converted into a reactive biotinoyl-5'-AMP intermediate that covalently tags proximal lysine residues[7]. Although the labelling time could be reduced from >18h for BioID[7] to less than an hour with TurboID[9], labelling with biotin ligases is significantly slower than with APEX peroxidases.

Proximity labelling experiments are routinely performed by transiently overexpressing a POI fused to a labelling enzyme. Overexpression can cause mislocalisation, protein aggregation and other artifacts[10-12], and a general shift to tagging at the endogenous protein level could improve the quality of the proximity labelling experiments. Physiological expression of the fused labelling enzyme could not only prevent artefacts but at the same time might reveal less abundant interactors that would be hidden in the unspecific labelling background due to overexpression. So far, endogenously expressed biotin ligases have been successfully used in mice and *C. elegans*[13-16], but for cultured cell lines that can be easily transfected, transient overexpression of the labelling enzyme is generally preferred as it presents a faster way to perform the experiments.

Here, we combine a rapid KI strategy based on antibiotic selection of positively edited cells[17,18] with proximity-based proteomics to detect interactors at physiological expression levels. This approach enabled us to endogenously tag a protein with the four most commonly used labelling enzymes (APEX2, BioID2, miniTurboID and TurboID) and compare their performance when expressed endogenously. We tagged the μ1A subunit of the adaptor protein complex AP-1 (AP1μA) at the C-terminus and clathrin light chain (CLC) at the N-terminus. AP-1, together with CLC, mediates specific transport between the *trans*-Golgi network (TGN) and endosomes[19,20] and transiently associates to membranes, leaving excess, non-membrane bound subunits in the cytosol. Our data reveals that endogenous tagging allows proximity labelling with higher specificity than achieved with overexpression of the labelling enzyme. Known interactors of AP-1 as well as specific cargo were higher enriched or even exclusively found in the experiments performed with a KI cell line. We identified the TurboID and the miniTurboID to be best suited for knock-in proximity labelling and propose a pipeline for rapid endogenous tagging to improve the quality of proximity-based mass spectrometry experiments.

Results

Endogenous tagging of AP1μA with labelling enzymes

To evaluate the performance of commonly used labelling enzymes, we used the CRISPR-Cas9 system to genetically fuse APEX2, BioID2, miniTurboID and TurboID to the C-terminus of our chosen target AP1μA. Insertion of a geneticin (G418) resistance cassette downstream of a polyadenylation signal into the targeted gene locus (AP1M1), allowed us to rapidly select for cells that were successfully modified[18]. The CRISPR strategy is illustrated in Fig. 1A. Cells were transfected with plasmids encoding for a gRNA targeting the genomic locus and homology repair plasmids containing sequences of the various labelling enzymes. Three days post transfection G418 was added, and cells were allowed to grow back to confluency to perform downstream analysis. The four KIs were verified with a western blot (Fig. 1B) and the correct localisation of the AP1μA fusion proteins was confirmed via immunofluorescence microscopy (Fig. 1C). The cargo adaptor AP-1 orchestrates transport between the TGN and endosomes[19-21] and is reported to predominantly localise to the TGN[22]. We found all AP1μA fusions localised in a juxtanuclear area typical for Golgi-localised proteins.

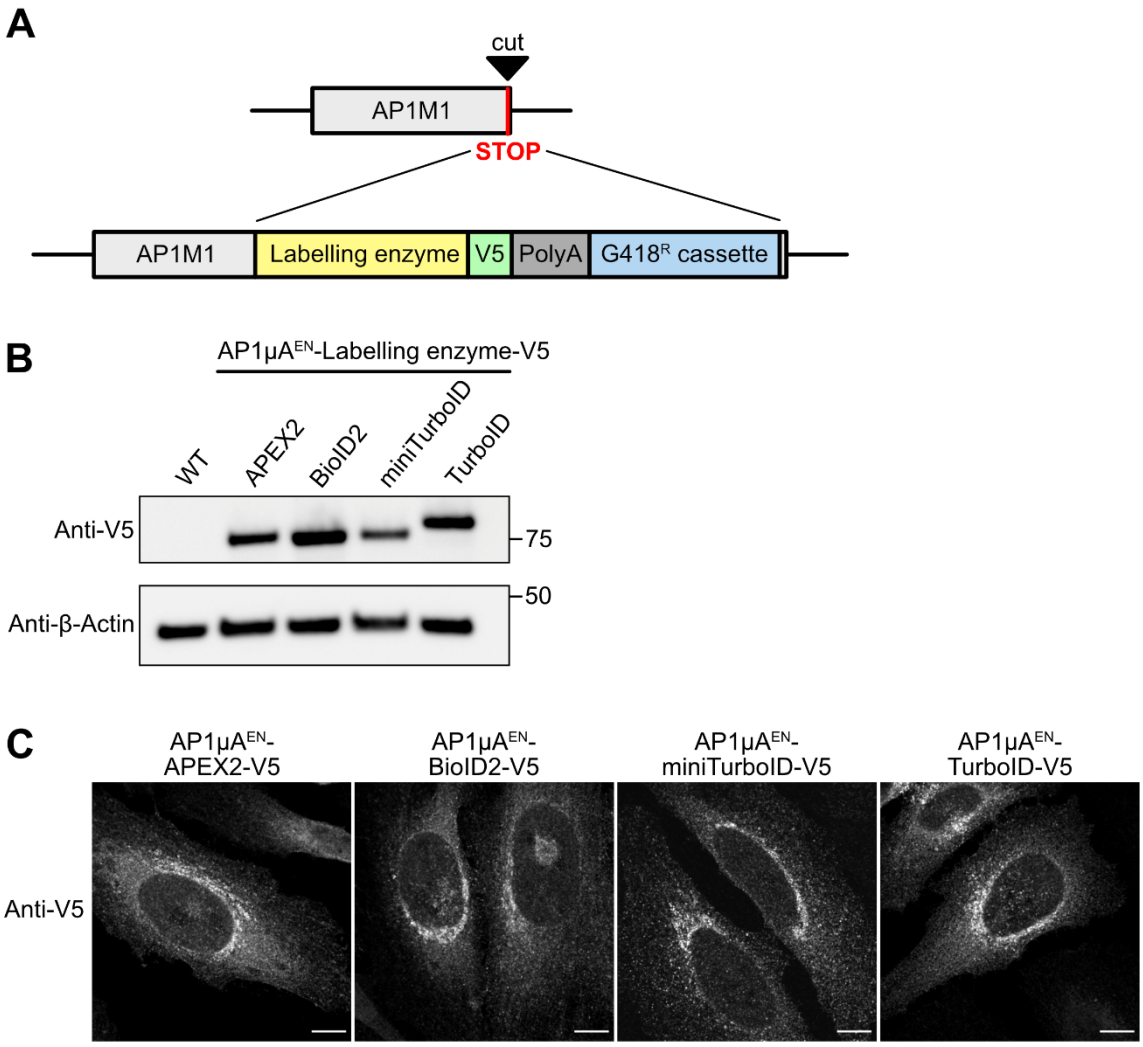


Fig. 1: **Rapid KI-strategy allows for endogenous tagging of AP1μA with different labelling enzyme for proximity biotinylation.**
A: Scheme of KI strategy. AP1μA was C-terminally tagged with the labelling enzyme, a V5 tag and a resistance cassette that allows for rapid selection of positive cells. **B:** Blots of whole cell lysates from generated cell lines to verify the KI of the labelling enzymes by anti-V5 blotting. **C:** Cell lines

expressing the labelling enzymes endogenously were fixed and stained for the V5 tag to detect the labelling enzyme expression and localisation. ^{EN}= endogenous. Scale bars are 10 µm.

TurboID variants outperform other labelling enzymes at endogenous expression levels

We then wanted to compare the performance of the various labelling enzymes when expressed at endogenous levels. For biotin ligases labelling times between less than 1h (miniTurboID and TurboID)[9,23] and at least 16 h – 24 h (BioID2)[8,24] are reported. To ensure the amount of biotinylation is sufficient for visualisation and comparable between the different biotin ligases, we treated all cell lines with 50 µM biotin for 24 h to initiate the labelling. For the APEX2 peroxidase we used the established effective labelling time of 1 min[6,25] to avoid prolonged exposure to toxic H_2O_2. To visualise biotinylated proteins, we used fluorescently labelled streptavidin. The different biotin ligases BioID2, miniTurboID and TurboID are all fully functional when fused to endogenous AP1µA (Fig. 2A). Importantly, biotinylated proteins are strongly concentrated to the Golgi region populated by the AP1µA fusions. However, we could not find any specific biotinylation for the APEX2 peroxidase when expressed at physiological levels (Fig. 2A). To exclude general handling errors with the APEX2 sample we transiently overexpressed APEX2-Vimentin as well as AP1µA-APEX2 fusions and found for both constructs a specific biotinylation pattern (Supplementary Fig. 1A,B). Taken together, these results indicate that the APEX2 peroxidase reaction requires very high copy numbers of the APEX2 fusion protein for quantitative proximity labelling and might not be compatible with KI approaches and low abundance endogenous targets.

To quantitatively compare the performance of the various biotin ligases, we compared the biotinylation pattern after supplementing the culture media for either 2 h or 24 h with 50 µM biotin (Fig. 2B). We quantified the biotinylation in the Golgi area and found that shorter TurboID and miniTurboID labelling (2 h) are more efficiently than longer BioID2 labelling (24 h) at physiological expression levels (Fig. 2C). Moreover, we assessed biotinylation rate of the different labelling enzymes on a western blot (Fig. 2D) and quantified the amount of biotinylated proteins against the amount of expressed labelling enzyme (Fig. 2E). We again found a high biotinylation rate for miniTurboID and TurboID in contrast to lower levels of labelled proteins through the BioID2 or APEX2. Collectively, these results suggest that miniTurboID and TurboID are best suited for tagging endogenous proteins to perform proximity labelling experiments. The low biotinylation rate of BioID2 makes it unfavourable compared to the TurboID variants, as a high amount of biotinylation is crucial to enrich enough protein for the subsequent identification and quantification by mass spectrometry (MS). We decided to use the AP1µA^{EN}-TurboID-V5 (^{EN}= endogenous) cell line for our further experiments as in our single cell analysis we found the biotinylation rate slightly increased compared to the miniTurboID (Fig. 2C).

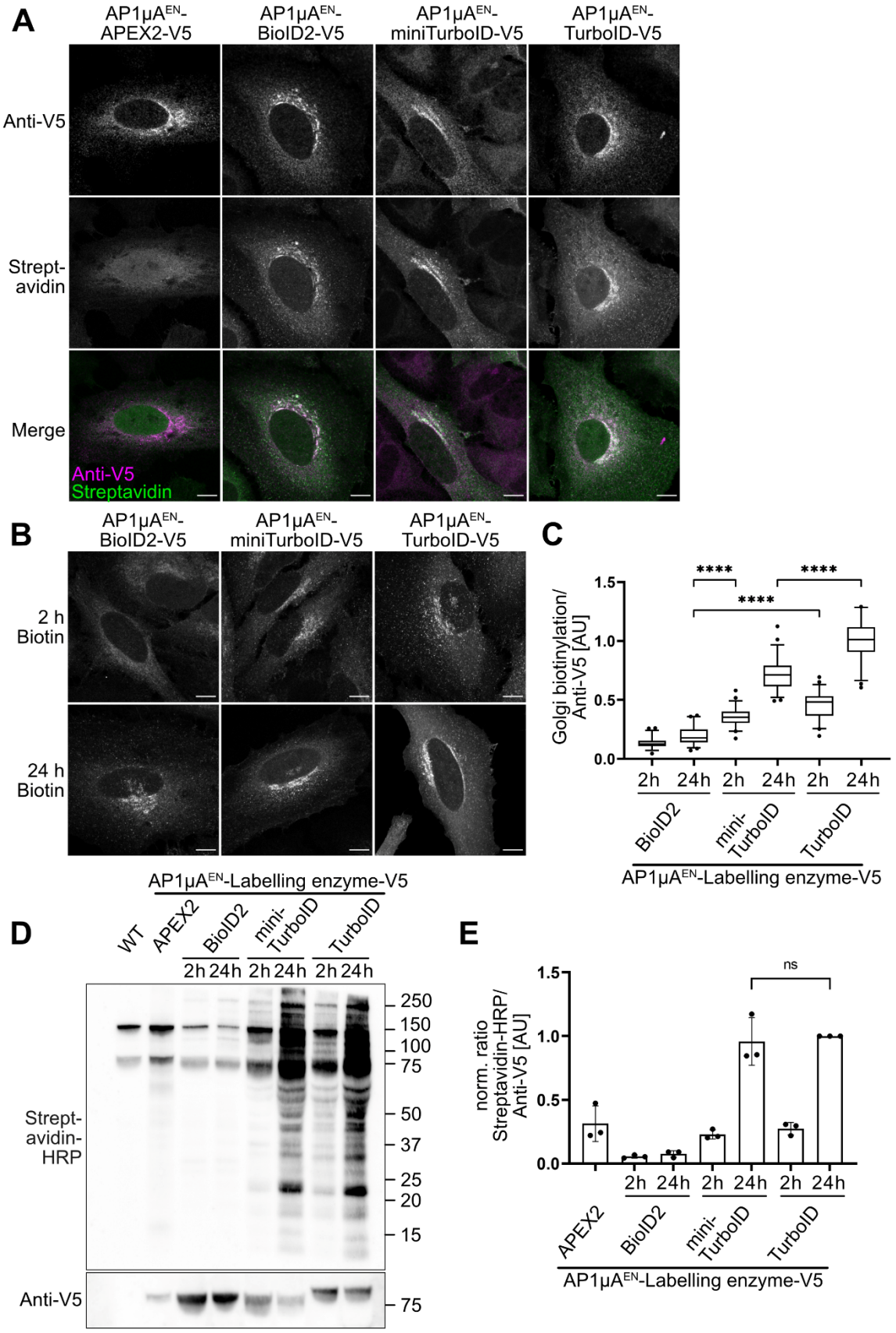


Fig. 2: **TurboID and MiniTurboID are most suitable for endogenous tagging with labelling enzymes.**

A: Cells endogenously expressing the different labelling enzymes were fixed and stained with anti-V5 antibody to detect the labelling enzyme and streptavidin-AF488 to detect biotinylated proteins. Cells were treated with 50 μM biotin for 24 h and AP1μA^{EN}-APEX2-V5 expressing cells were

incubated 30 min with 500 µM biotin-phenol and labelling was induced for 1 min with H_2O_2. **B:** Detection of biotinylated proteins with streptavidin-AF488 in fixed cells expressing different biotin ligases. 50 µM Biotin was added for either 2 h or 24 h before fixation. **C**: Ratio of biotinylated Golgi-localised proteins detected by streptavidin-AF488 to biotin ligase expression detected by anti-V5 antibody at the Golgi. Cells expressing the different biotin ligases endogenously fused to AP1µA were treated with 50 µM biotin for 2 h or 24 h and then fixed and prepared for microscopy as in **A**. For each condition 30-50 cells were analysed. All p-values from unpaired t-tests are <0.0001. **D:** Comparison of the labelling efficiency of different labelling enzymes endogenously fused to AP1µA on a western blot. Cells were treated with 50 µM biotin for 2 h or 24 h, WT (wild type) cells were treated for 24 h with 50 µM biotin and APEX2 were incubated 30 min with 500 µM biotin-phenol and labelling was induced for 1 min with H_2O_2. Whole cell lysates were blotted with streptavidin-HRP to detect biotinylated proteins, and anti-V5 antibody to compare ligase expression. This experiment was performed three times with similar results. **E:** The ratio of biotinylated proteins (Streptavidin-HRP) to the expressed labelling enzyme (Anti-V5) was calculated for each line on the blot in **D** and then normalised to the TurboID value for each independent replicate. P-value is 0.71. Scale bars are 10 µm.

Endogenous N-terminal tagging of CLC

To test whether we can apply our rapid KI strategy of labelling enzymes also for N-terminally tagged proteins, we fused APEX2 and TurboID endogenously to the N-terminus of clathrin light chain A (CLCa). For N-terminal fusions, the resistance cassette is inserted between LoxP sites upstream of the start codon of the tags[17]. The resistance cassette is then excised via transfection of Cre recombinase. The strategy used for N-terminal tagging is illustrated in Fig. 3A. Again, comparable to the C-terminally tagged AP1µA, we observed locally confined biotinylation for the biotin ligase (V5-TurboID-CLCa^{EN}) but no specific biotinylation for the APEX2 fusion protein (Fig. 3B).

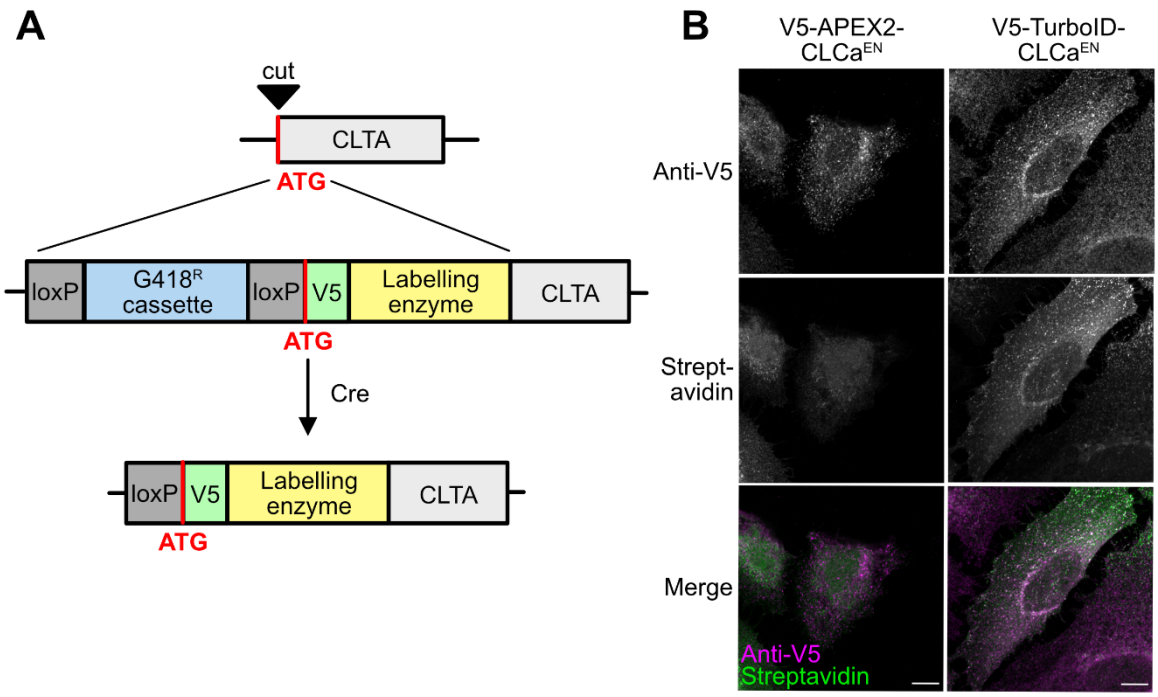


Fig. 3: Endogenous N-terminal tagging of CLCa with labelling enzymes.
A: Scheme of KI strategy. CLCa was N-terminally tagged with the labelling enzyme and a V5 tag. The integrated resistance cassette can be excised with the loxP-Cre-system after transfection with the Cre recombinase. **B:** HeLa V5-APEX2-CLCa^{EN} and V5-TurboID-CLCa^{EN} cells were fixed and stained with anti-V5 antibody to detect the labelling enzyme and streptavidin-AF488 to detect biotinylated proteins. V5-TurboID-CLCa^{EN} cells were treated with 50 µM biotin for 24 h and V5-

APEX2-CLCa^{EN} cells were incubated 30 min with 500 µM biotin-phenol and labelling was induced for 1 min with H_20_2. Scale bars are 10 µm.

Knock-in biotin ligases allow for locally confined proximity labelling

Transient overexpression of a protein can lead to artefacts such as mislocalisation or aggregation [10-12], increasing the chances of detecting non-specific, artificial interactors such as abundant cytosolic proteins. Physiological expression of the labelling enzyme should allow for highly specific, locally confined biotinylation of natural interactors. By applying two-colour stimulated emission depletion (STED) super-resolution microscopy, we were able to visualise the biotinylation pattern after 2 h of biotin addition in AP1µA^{EN}-TurboID-V5 knock-in cells (Fig. 4A) and to compare it with AP1µA-TurboID-V5 overexpression (Fig. 4B). The high resolution achieved on the STED microscope enabled visualisation of distinct nanodomains occupied by AP-1. Notably, the biotinylated proteins are primarily localised to the very same nanodomains as the endogenous AP1µA^{EN}-TurboID-V5 fusion (indicated by the white arrows in Fig. 4A), indicating a high local specificity of the proximity labelling of the TurboID. Overall, localisation of the overexpressed AP1µA-TurboID-V5 was more diffuse (Fig. 4B). Interestingly, biotinylated proteins localised to extended areas with no overlap with the fusion protein AP1µA-TurboID-V5 (white arrows in Fig. 4B). These unspecific, highly biotinylated zones are likely a result of excessive background biotinylation caused by overexpressed AP1µA-TurboID-V5. In addition, areas were found where AP1µA-TurboID is present but no obvious biotinylation occurred (yellow arrows in Fig. 4B). Many cells overexpressing AP1µA-TurboID-V5 exhibit a very high background biotinylation in the cytoplasm and in the nucleus (Fig. 4C) while background biotinylation was unnoticeable in the KI cells. Quantification of the background biotinylation revealed significantly higher cytosolic biotinylation in the transiently overexpressed cells compared to the KI cells, even after short biotin incubation (Fig. 4D). The drastic reduction of unspecific biotinylation in our CRISPR-KI cell line promises to achieve a higher sensitivity for specific low abundance interactors in proximity labelling experiments.

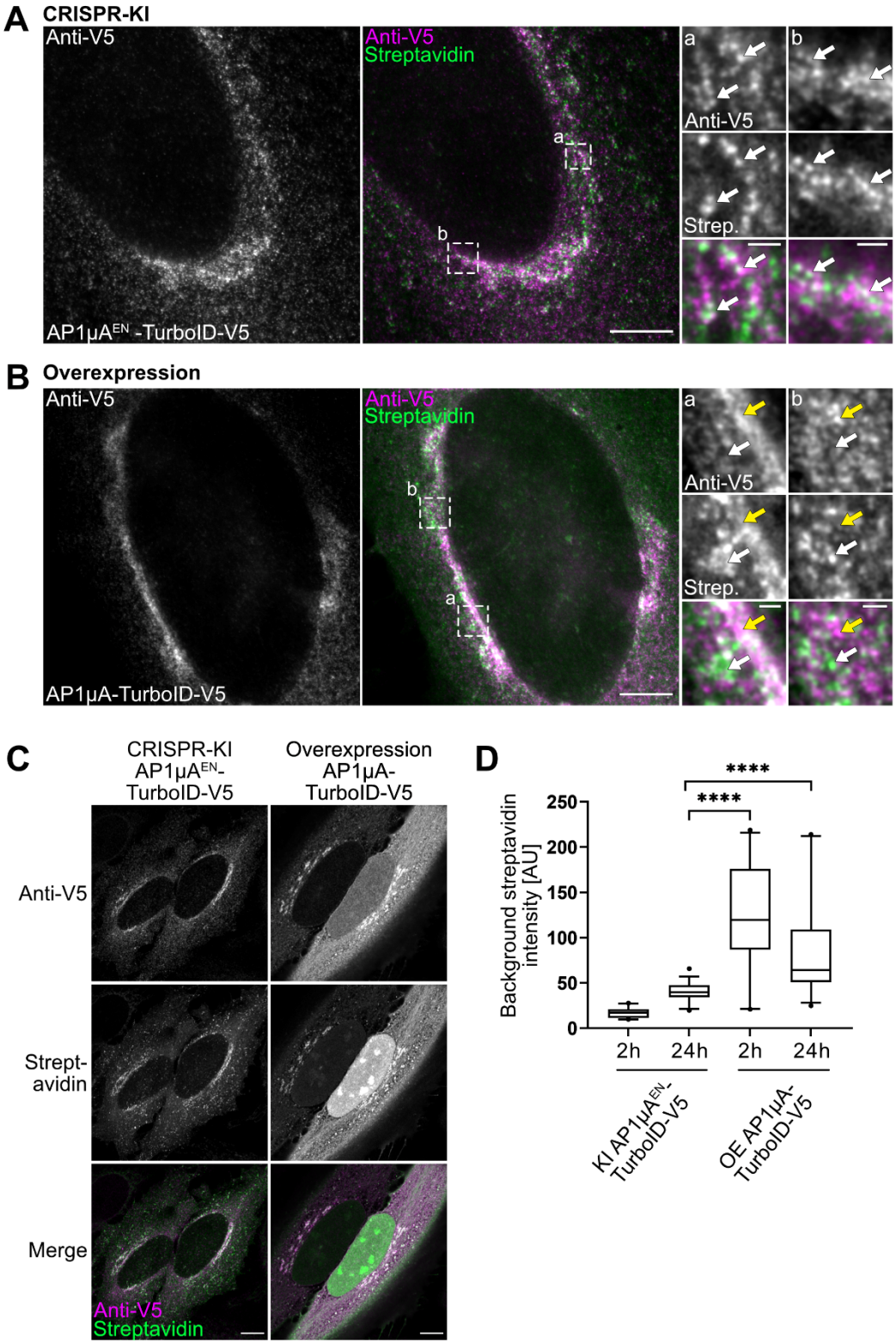


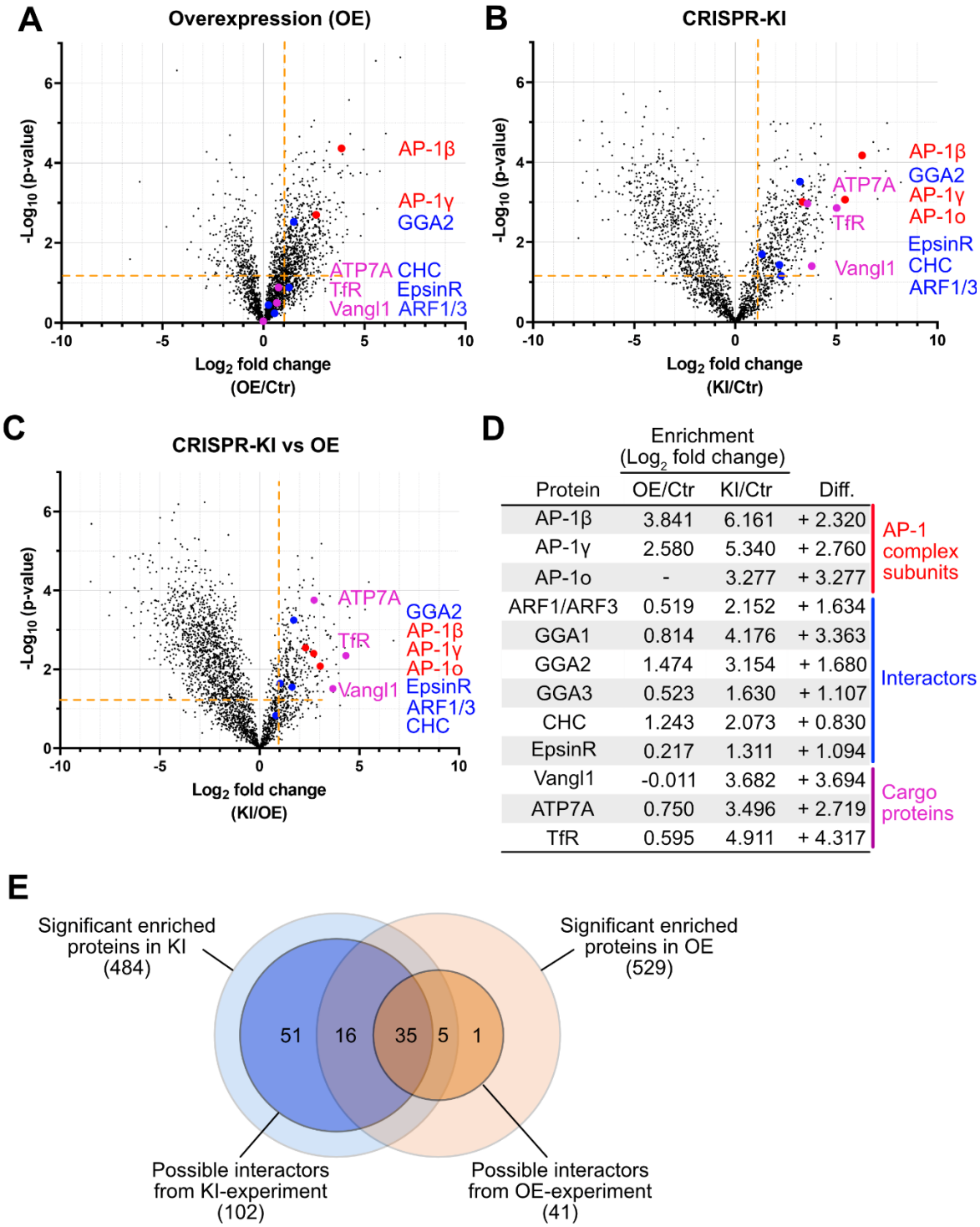
Fig. 4: Endogenous tagging allows a more spatially confined proximity labelling than overexpression of the labelling enzyme.
A: STED micrographs of a fixed AP1μA^{EN}-TurboID-V5 cell stained with anti-V5 antibody and streptavidin-STARORANGE to detect biotinylated proteins. Cells were treated with 50 μM biotin for 2 h before fixation. Crops show distinct overlap of biotinylated proteins and AP1μA^{EN}-TurboID-V5 (marked by white arrows). **B:** STED micrographs of a fixed cell transiently overexpressing AP1μA-TurboID-V5 that was treated as described in **A**. Crops show that biotinylated proteins and AP1μA-TurboID-V5 accumulate in distinct zones (white arrows indicate areas of biotinylation without AP1μA-TurboID-V5, yellow arrows indicate areas of accumulated biotin ligase without biotinylated proteins). **C:** Comparison of cells either endogenously expressing AP1μA^{EN}-TurboID-V5 or transiently

overexpressing AP1μA-TurboID-V5. Cells were treated for 2 h with 50 μM biotin, fixed and stained with anti-V5 antibody to detect the labelling enzyme and streptavidin-AF488 to detect biotinylated proteins. **D:** Background biotinylation in cells either endogenously expressing AP1μA^{EN}-TurboID-V5 (KI) or transiently overexpressing AP1μA-TurboID-V5 (OE). Background biotinylation was measured in cells that were treated with biotin for 2 h or 24 h by measuring streptavidin-AF488 signal intensity from a cytosolic area. At least 20 cells were analysed per condition. P-values are all <0.0001. Scale bars are 10 μm for **C**, 5 μm for STED images in **A,B** and 500 nm for crops in **A**,**B**.

Physiological expression of TurboID fusions permits highly specific interactome identification

To probe whether endogenous tagging with TurboID allows biotin labelling of AP1μA-specific interactors with increased sensitivity compared to transient overexpression, we analysed streptavidin-purified proteins by MS using label-free quantification. Biotin was added to the culture medium for 24 h to secure detectable protein labelling in the KI cells. In total, we identified and quantified 2481 proteins (Supplementary Table 1). Proteins with at least a 2-fold increase in relative intensity compared to control ($\log_2$ fold change >1) and a q-value ≤0.05 were considered significantly enriched. We found 529 proteins significantly enriched in the overexpression sample and 484 proteins significantly enriched in the KI sample. Overall, a larger number of proteins was enriched when AP1μA-TurboID was overexpressed compared to the KI condition. This may be the result of mislocalisation of AP1μA-TurboID which shows extensive cytoplasmic localisation, possibly leading to the biotinylation of a larger set of proteins that naturally would not interact with AP-1. We found the two large subunits of the AP-1 complex (AP-1β1 and AP-1γ)[19] significantly enriched in overexpression condition (Fig. 4A) and even higher enriched in the KI cells (Fig. 4B). Notably, the other small subunit of the AP-1 complex (AP-1σ)[19] was only identified when AP1μA^{EN}-TurboID-V5 was endogenously expressed. We next looked at known interactors of the AP-1 complex to test whether endogenous tagging improves enrichment of specific interactors. The AP-1 complex is recruited to the Golgi membranes by the small GTPase ARF1 (ADP ribosylation factor 1)[20] and is known to directly bind clathrin, EpsinR and GGAs (Golgi-localized, γ-ear-containing, Arf-binding proteins)[26-28]. All interactors were identified in the overexpression condition but only GGA2 was found significantly enriched. Physiological expression of AP1μA^{EN}-TurboID-V5 on the other hand allowed significant enrichment of clathrin heavy chain, ARF1, EpsinR and all GGA proteins (GGA1-3). The difference between physiological expression levels and overexpression was even more striking when we looked at the AP-1 specific cargo proteins Vangl1, ATP7A and the transferrin receptor (TfR). ATP7A and TfR are both reported to be trafficked through the AP-1 complex[22,29] and Vangl1 shares a conserved export motif with Vangl2 that is sorted by AP-1[30]. These AP-1 cargo proteins were all not or only slightly enriched in the overexpression experiment but significantly enriched in KI cells. Direct comparison between the two data sets illustrated that endogenous tagging of AP1μA with TurboID permits significantly higher enrichment of all known interactors in proximity labelling MS experiments compared to overexpression of labelling enzyme (Fig 4C, D). To finally evaluate the overall quality of the MS data from the CRISPR-KI AP1μA^{EN}-TurboID experiment, we identified possible interactors of AP1μA from the MS data and compared the two data sets (KI vs overexpression). We defined potential interactors as proteins that were at least 4-fold significantly enriched ($\log_2$ fold change >2, q-value ≤0.05) and are either known to localise to the Golgi/TGN, are known to be involved in cellular trafficking, are transmembrane proteins that might be trafficked by AP1μA or might be involved in the regulation of membrane homeostasis (e.g. regulatory kinases). For the KI cell line, we found a total of 102 potential interactors (Supplementary Table 2) of which 51 were not significantly enriched ($\log_2$ fold change >1, q-value ≤0.05) in the overexpression condition. Evaluation of the MS data from the overexpressed AP1μA-TurboID resulted in a list of only 41 possible interactors from which only 1 was not significantly enriched in the KI-condition (Supplementary Table 3). The results of our analysis are illustrated in Fig. 5E. Taken together, our findings highlight that tagging the

endogenous protein with the labelling enzyme increases the sensitivity of the proximity biotinylation and the MS measurement so that very transient but specific interactors, such as specific cargo proteins, can be identified.



Protein	Enrichment (Log2 fold change) OE/Ctr	KI/Ctr	Diff.	
AP-1β	3.841	6.161	+ 2.320	AP-1 complex subunits
AP-1γ	2.580	5.340	+ 2.760	
AP-1o	-	3.277	+ 3.277	
ARF1/ARF3	0.519	2.152	+ 1.634	Interactors
GGA1	0.814	4.176	+ 3.363	
GGA2	1.474	3.154	+ 1.680	
GGA3	0.523	1.630	+ 1.107	
CHC	1.243	2.073	+ 0.830	
EpsinR	0.217	1.311	+ 1.094	
Vangl1	-0.011	3.682	+ 3.694	Cargo proteins
ATP7A	0.750	3.496	+ 2.719	
TfR	0.595	4.911	+ 4.317	

Fig. 5: Enrichment of specific interactors with endogenously expressed AP1μA^{EN}-TurboID. **A:** Volcano plot showing the changes in relative protein intensity between the overexpression (OE) experiment and control (cytosolic overexpressed TurboID). Significant hits are shown in the top right corner (q-value ≤ 0.05 and log_2 fold change > 1) separated by the orange lines. Subunits of the AP-1 complex (red), known interactors (blue) and known cargos (magenta) are marked. **B:** Volcano plot showing the changes in relative protein intensity between the KI experiment (AP1μA^{EN}-TurboID-V5) and control. Same parameters as in **A**. **C**: Volcano plot showing the changes in relative protein intensity between KI experiment and

the OE experiment. Same parameters as in **A. D:** Table of the analysed subunits, interactors and cargo proteins. Differences (Diff.) in log_2 fold enrichment are indicated. Entire protein list is shown in Supplementary Table 1. **E**: Venn diagram showing the number of potential interactors (defined by protein localisation and function (see methods) and log_2 fold change >2, q-value ≤0.05) that we were able to identify in the KI and in the OE condition compared with the overall number of enriched proteins for each condition. Lists of potential interactors are shown in Supplementary Tables 2 and 3.

Discussion

By tagging endogenous AP1µA with the commonly used labelling enzymes APEX2, BioID2, miniTurboID and TurboID we were able to identify the biotin ligases miniTurboID and TurboID as best suited to perform KI proximity labelling experiments (Fig. 2). Previous work has already demonstrated the TurboID variants outperform BioID or BioID2[9,23,24,31], however BioID2 is still a commonly used enzyme. When wanting to combine gene editing with proximity biotinylation, we recommend switching to the faster kinetics of TurboID. To our surprise, we could not achieve any biotinylation with the APEX2 peroxidase expressed as a low abundance endogenous fusion (Fig. 2,3). The short labelling time in combination with the low physiological expression does not allow detectable proximity labelling. We were unable to combine APEX2-tagging with the presented KI proximity labelling approach as APEX2 requires higher expression levels to achieve quantitative labelling with biotin. Additionally, the incubation time with the biotin might be adapted according to the abundance of the POI. AP-1 is a relatively abundant protein[32], but for lower expressed proteins the time of biotin treatment might be extended while for more abundant protein it could be shortened. Recently, new variants of the BioID2 were introduced, microID and ultraID[33]. With a molecular weight below 20 kDa they are significantly smaller than the here presented labelling enzymes. Especially the ultraID is reported to have labelling kinetics similar to TurboID with less background activity. It would be interesting to see how these variants perform when physiologically expressed. Reduction of background biotinylation could be particularly beneficial in cell lines with continuous endogenous expression of the labelling enzymes.

Using an AP1µA^{EN}-TurboID-V5 cell line, we were able to compare the enrichment of known interactors and cargo proteins of the AP-1 complex in quantitative MS measurements against data sets generated with cells overexpression AP1µA-TurboID-V5. Strikingly, we found known interactors significantly higher enriched when AP1µA^{EN}-TurboID-V5 was expressed at endogenous levels (Fig. 5). Especially relevant for our research is the strong enrichment of AP-1 cargo proteins that can be only observed when TurboID is endogenously fused to AP1µA. The increased sensitivity for real interactors is likely a result of less unspecific labelling due to mislocalisation- or aggregation artefacts induced by overexpression of the labelling enzyme fused to the POI. A larger background of peptides from unspecifically labelled proteins increases the sample complexity and therefore lowers the overall sensitivity of the MS sample as all peptides compete for ionisation and detection. Endogenous protein tagging with a biotin ligase that allows for highly confined biotinylation (Fig. 4) therefore increases the chances for detection of specific interactors, especially if they are of low abundance. Importantly, endogenous expression of the AP1µA^{EN}-TurboID fusion resulted in a more comprehensive list of potential interaction partners compared to overexpression of AP1µA-TurboID (Supplementary Tables 2, 3 + Fig. 5E).

In summary, endogenous tagging with biotin ligases enables highly specific proximity labelling and increases the sensitivity for real interactors that might be transient or of low abundance. The pipeline for the generation of CRISPR KIs, based on the integration of a resistance cassette, allows rapid generation of endogenously tagged cells. The ease and speed of the KI

generation make it an attractive alternative to transient overexpression of the labelling enzyme as the MS experiments can be done in about 4-5 weeks after the CRISPR transfection.

Acknowledgments

We acknowledge the whole lab for the fruitful discussion during the planning and writing phase of this manuscript. This project was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grants SFB958 (Project A25), SFB/TRR186 (Project A05 (CF) and Project A20 (FB)) and a major research instrumentation grant for the acquisition of the STED microscope. For mass spectrometry, we acknowledge the assistance of the Core Facility BioSupraMol supported by the Deutsche Forschungsgemeinschaft (DFG).

Methods

Antibodies and Streptavidin conjugates

All dyes and antibodies used in this study are provided in Supplementary Table 4 in the supplementary information.

Mammalian cell culture

All experiments were carried out in HeLa cells ATCC CCL-2 (ECACC General Collection) grown in a humidified incubator at 37°C with 5% CO_2 in DMEM (Gibco-ThermoFisher) supplemented with 10% fetal bovine serum (Corning) and penicillin/streptomycin (Lonza Bioscience) to prevent contamination. For transient transfection, HeLa cells at 70-80% confluency were transfected with FuGENE HD Transfection Reagent (Promega) according to the supplier's protocol.

Generation of CRISPR-Cas9 knock-in cell lines

See supplementary information.

Plasmid design of overexpression plasmids

See supplementary information.

Immunofluorescence

For all immunofluorescence samples, 40 000 cells were seeded on fibronectin-coated coverslips. Biotinylation was induced by replacing the growth media with medium containing 50 µM biotin and samples were incubated for 24 h or 2 h as indicated in the figures. Biotinylation for AP1µA^{EN}-APEX2-V5 cells was induced as described by Hung et al.[25]. All cells were washed twice with PBS and then fixed in 4% PFA for 10 min at RT. Subsequently, they were rinsed three times with PBS and incubated for 3 min in permeabilization buffer (0.3% NP40 (Roth), 0.05% Triton-X 100 (Sigma Aldrich) and 0.1% BSA (IgG free) (Roth) in PBS). Cells were blocked for 1 h in blocking buffer (0.05% NP40, 0.05% Triton-X 100 and 5% goat serum (Jackson ImmunoResearch) in PBS) at RT and then incubated with an anti-V5 rabbit antibody in blocking buffer overnight rocking at 4°C. On the next day, the samples were washed three times 5 min in washing buffer (0.05% NP40, 0.05% Triton-X 100 and 0.2% BSA (IgG free) in PBS) before incubation with an anti-rabbit ATTO647N-coupled secondary goat antibody in blocking buffer for 1 h rocking at RT. For visualisation of biotinylated proteins either Streptavidin-Alexa488 (for confocal microscopy) Streptavidin-STARORANGE (for STED microscopy) was added to the secondary antibody mix. The cells were then washed three times 5 minutes with wash buffer and then dipped in dH_2O before mounting with ProLong Gold Antifade Reagent (Thermo Fisher Scientific). Mounted samples were let to harden overnight at RT and then stored at 4°C until imaging.

Imaging and image processing

Confocal and STED imaging was carried out on a commercial expert line Abberior STED microscope equipped with 488 nm, 561 nm and 645 nm excitation lasers. For two-colour STED experiments both dyes were depleted with a 775 nm depletion laser. The detection windows for the dyes were set to 498 to 551 nm, 571 to 630 nm and 650 to 756 nm. Excitation power was kept constant between samples in the same experiment to be able to quantify differences in expression levels and biotinylation. The pixel size was set to 60 nm for confocal and 20 nm for STED.

All images shown were smoothed using a Gaussion filter with 1-pixel SD using ImageJ [34].

Image analysis and statistical analysis

All image analysis was carried out with ImageJ. To determine the ratio of biotinylated proteins and V5-tagged AP1μA-labelling enzyme fusions in Fig. 2C, a small region in the Golgi area was selected in the raw image and the average grey values were measured for both channels. The ratio between the mean intensity fluorescence of the biotinylated proteins in the Golgi area versus the mean fluorescent intensities from the V5 channel was then calculated. For each condition at least 40 cells were analysed from three independent experiments.

To analyse background biotinylation in Fig. 4D, a small region in a cytoplasm was selected and the mean intensity fluorescence was measured. At least 20 cells from each condition were analysed.

Statistical analysis (t-tests) was carried out with Prism. P-values are indicated in figure legends.

Knock-in verification via western blot

For each knock-in cell line, 180 000 cells were seeded on a 6-well plate. 24 h later cells were washed twice with PBS and harvested in 350 µl of Laemmli sample buffer. The samples were boiled for 10 min at 98°C before loading 30 µl of each sample on two separate 4-12% SDS-polyacrylamide gels. After electrophoresis, proteins were transferred to a nitrocellulose membrane (Amersham) via wet blotting. Both membranes were washed once with phosphate buffered saline with 0.05% Tween (PBST) and then blocked for 1 h with 5% (wt/vol) milk powder and 1% BSA in PBST rocking at RT. After that, membranes were washed once 5 min with PBST and two times 5 min with PBS before incubation with either an anti-V5 antibody or an anti-β-actin antibody overnight rocking at 4°C. On the next day, the membranes were washed three times 5 min with PBST and incubated with a secondary antibody coupled to HRP in 5% (wt/vol) milk powder and 1% BSA in PBST for 30 min rocking at RT. The membrane was washed twice with PBST and twice with PBS for 5 min each. To develop the membrane the ECL western blot substrate was added for 2 min and then the membrane was imaged.

Western blot analysis of biotinylated proteins

For detection of biotinylated proteins on a western blot, 800 000 HeLa WT and all KI cell lines were seeded on a 10 cm cell culture dish. Starting on the next day, the media was replaced with biotin-containing medium (50 µM biotin) and samples were incubated for 24 h or 2 h at 37°C. The biotin addition was timed in a way that all samples could be harvested at the same time. Biotinylation for AP1μA^{EN}-APEX2-V5 cells was induced as described in ref. (add quotation). For harvesting, the cells were washed twice with ice-cold PBS and then extracted in 400 µL of ice-cold lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.5% NP-40, 0.5 mM DTT and protease inhibitors). Cells were then centrifuged for 10 min at 4°C at 14,000g to clarify the cell lysates. 20 µl of the whole-cell lysates were mixed with Laemmli buffer and boiled at 95°C for 10 min before loading on a 4-12% SDS-polyacrylamide gel. Two separate gels were used, one for detection of biotinylated proteins and the other for detection

of the fusion protein. After electrophoresis, proteins were transferred to a nitrocellulose membrane (Amersham) via wet blotting. To visualise biotinylated proteins on the membrane, after blocking (5% (wt/vol) milk powder in PBST), the blot was incubated with 0.3 µg/mL streptavidin-HRP in 3% BSA in PBST for 30 min rocking at RT. The western blot to detect the fusion protein with the V5 tag was carried about as described above.

For quantification of the biotinylated proteins in relation to the available labelling enzyme the intensity of the line on the biotinylation blot was measured in ImageJ and set in relation to the intensity of the corresponding V5 line. To measure the intensity of a line of biotinylated proteins in ImageJ, a box was drawn in the middle of the measured signal and the average intensity value of this area was used for analysis. The respective position and size of the box was kept constant for each line of a blot. The V5 tag intensity was measured accordingly except that the box was always placed around the area of the strongest signal. For each experiment the values were normalised to the AP1µA^{EN}-TurboID-V5 sample that was incubated for 24 h with biotin. The experiment was repeated three times.

Preparation of MS samples, LC-MS and data analysis

See supplementary information.

Identification of potential interactors of AP1µA

For identification of potential interactors only proteins that were at least 4-fold significantly enriched in the MS experiment were considered. These proteins were then screened with Uniprot to identify proteins that are are either known to localise to either the Golgi or TGN, are thought to be involved in cellular trafficking, are transmembrane proteins that might be trafficked by AP1µA or could control trafficking events (e.g. kinases, phosphatases). Together with all uncharacterised proteins these proteins were then considered potential interactors and summarised in Supplementary Tables 2,3.

References

1 Trinkle-Mulcahy, L. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res* **8**, F1000 Faculty Rev-1135, doi:10.12688/f1000research.16903.1 (2019).

2 Qin, W., Cho, K. F., Cavanagh, P. E. & Ting, A. Y. Deciphering molecular interactions by proximity labeling. *Nature Methods* **18**, 133-143, doi:10.1038/s41592-020-01010-5 (2021).

3 Martell, J. D. *et al.* Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nature Biotechnology* **30**, 1143-1148, doi:10.1038/nbt.2375 (2012).

4 Kim, D. I. *et al.* Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences* **111**, E2453, doi:10.1073/pnas.1406459111 (2014).

5 Rhee, H.-W. *et al.* Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* **339**, 1328, doi:10.1126/science.1230593 (2013).

6 Lam, S. S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nature Methods* **12**, 51-54, doi:10.1038/nmeth.3179 (2015).

7 Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology* **196**, 801-810, doi:10.1083/jcb.201112098 (2012).

8 Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Molecular Biology of the Cell* **27**, 1188-1196, doi:10.1091/mbc.E15-12-0844 (2016).

9 Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology* **36**, 880-887, doi:10.1038/nbt.4201 (2018).

10 Rizzo, M. A., Davidson, M. W. & Piston, D. W. Fluorescent Protein Tracking and Detection: Applications Using Fluorescent Proteins in Living Cells. *Cold Spring Harbor Protocols* **2009**, doi:10.1101/pdb.top64 (2009).

11 Gibson, T. J., Seiler, M. & Veitia, R. A. The transience of transient overexpression. *Nature Methods* **10**, 715-721, doi:10.1038/nmeth.2534 (2013).

12 Moriya, H. Quantitative nature of overexpression experiments. *Molecular Biology of the Cell* **26**, 3932-3939, doi:10.1091/mbc.E15-07-0512 (2015).

13 Rudolph, F. *et al.* Deconstructing sarcomeric structure–function relations in titin-BioID knock-in mice. *Nature Communications* **11**, 3133, doi:10.1038/s41467-020-16929-8 (2020).
14 Feng, W. *et al.* Identifying the Cardiac Dyad Proteome In Vivo by a BioID2 Knock-In Strategy. *Circulation* **141**, 940-942, doi:10.1161/CIRCULATIONAHA.119.043434 (2020).
15 Artan, M. *et al.* Interactome analysis of C. elegans synapses by TurboID-based proximity labeling. *Journal of Biological Chemistry*, 101094, doi:10.1016/j.jbc.2021.101094 (2021).
16 Murata, K. *et al.* Efficient induction of proximity-dependent labelling by biotin feeding in BMAL1-BioID knock-in mice. *The Journal of Biochemistry*, doi:10.1093/jb/mvab059 (2021).
17 Schmidt, J. C., Zaug, A. J. & Cech, T. R. Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. *Cell* **166**, 1188-1197.e1189, doi:10.1016/j.cell.2016.07.033 (2016).
18 Bottanelli, F. *et al.* A novel physiological role for ARF1 in the formation of bidirectional tubules from the Golgi. *Molecular Biology of the Cell* **28**, 1676-1687, doi:10.1091/mbc.e16-12-0863 (2017).
19 Hirst, J. & Robinson, M. S. Clathrin and adaptors. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1404**, 173-193, doi:10.1016/S0167-4889(98)00056-1 (1998).
20 Ren, X., Farías, Ginny G., Canagarajah, Bertram J., Bonifacino, Juan S. & Hurley, James H. Structural Basis for Recruitment and Activation of the AP-1 Clathrin Adaptor Complex by Arf1. *Cell* **152**, 755-767, doi:10.1016/j.cell.2012.12.042 (2013).
21 Robinson, M. S. Adaptable adaptors for coated vesicles. *Trends in Cell Biology* **14**, 167-174, doi:10.1016/j.tcb.2004.02.002 (2004).
22 Gravotta, D. *et al.* The clathrin adaptor AP-1A mediates basolateral polarity. *Dev Cell* **22**, 811-823, doi:10.1016/j.devcel.2012.02.004 (2012).
23 Cho, K. F. *et al.* Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nature Protocols* **15**, 3971-3999, doi:10.1038/s41596-020-0399-0 (2020).
24 Zhang, Y. *et al.* TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. *Nature Communications* **10**, 3252, doi:10.1038/s41467-019-11202-z (2019).
25 Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nature Protocols* **11**, 456-475, doi:10.1038/nprot.2016.018 (2016).
26 Doray, B. & Kornfeld, S. Gamma subunit of the AP-1 adaptor complex binds clathrin: implications for cooperative binding in coated vesicle assembly. *Molecular biology of the cell* **12**, 1925-1935, doi:10.1091/mbc.12.7.1925 (2001).
27 Mills, I. G. *et al.* EpsinR : an AP1/clathrin interacting protein involved in vesicle trafficking. *Journal of Cell Biology* **160**, 213-222, doi:10.1083/jcb.200208023 (2003).
28 Doray, B., Ghosh, P., Griffith, J., Geuze Hans, J. & Kornfeld, S. Cooperation of GGAs and AP-1 in Packaging MPRs at the Trans-Golgi Network. *Science* **297**, 1700-1703, doi:10.1126/science.1075327 (2002).
29 Holloway, Z. G. *et al.* Trafficking of the Menkes copper transporter ATP7A is regulated by clathrin-, AP-2–, AP-1–, and Rab22-dependent steps. *Molecular Biology of the Cell* **24**, 1735-1748, doi:10.1091/mbc.e12-08-0625 (2013).
30 Guo, Y., Zanetti, G. & Schekman, R. A novel GTP-binding protein–adaptor protein complex responsible for export of Vangl2 from the trans Golgi network. *eLife* **2**, e00160, doi:10.7554/eLife.00160 (2013).
31 May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* **9**, doi:10.3390/cells9051070 (2020).
32 Nagaraj, N. *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* **7**, 548-548, doi:10.1038/msb.2011.81 (2011).
33 Zhao, X. *et al.* ultraID: a compact and efficient enzyme for proximity-dependent biotinylation in living cells. *bioRxiv*, 2021.2006.2016.448656, doi:10.1101/2021.06.16.448656 (2021).
34 Abramoff, M., Magalhães, P. & Ram, S. J. Image Processing with ImageJ. *Biophotonics International* **11**, 36-42 (2004).