ultraID: a compact and efficient enzyme for proximity-dependent biotinylation in living cells

Xiyan Zhao*, Sebastian Bitsch*, Lea Kubitz*, Kerstin Schmitt, Lukas Deweid, Amélie Roehrig, Elisa Cappio Barazzone, Oliver Valerius, Harald Kolmar, Julien Béthune**

* equal contribution  
** corresponding author

Abstract

Proximity-dependent biotinylation (PDB) combined with mass spectrometry analysis has established itself as a key technology to study protein-protein interactions in living cells. A widespread approach, BioID, uses an abortive variant of the E. coli BirA biotin protein ligase, a quite bulky enzyme with slow labeling kinetics. To improve PDB versatility and speed, various enzymes have been developed by different approaches. Here we present a novel small-size engineered enzyme: ultraID. We show its practical use to probe the interactome of Argonaute-2 after a 10 min labeling pulse and express at physiological levels. Moreover, using ultraID, we provide a membrane-associated interactome of coatomer, the coat protein complex of COPI vesicles. To date, ultraID is the smallest and most efficient biotin ligase available for PDB and offers the possibility of investigating interactomes at a high temporal resolution.

Key words: protein-protein interactions, PPI, BioID, proximity-dependent biotinylation, directed evolution, biotin ligase, proteomics, interactome

Introduction

Proximity-dependent biotinylation (PDB) in living cells is a powerful technique to identify protein-protein interactions in living cells. In the original approach BioID, the R118G abortive variant of the E. coli protein biotin ligase (PBL) BirA, is fused to a protein of interest. In cells, BioID uses ATP and biotin as substrates to produce biotinyl-AMP (bioAMP). By contrast to the wild type (WT) enzyme, BioID does not retain bioAMP in its active site but releases it, leading to the non-enzymatic biotinylation of lysine residues on surrounding proteins within an estimated radius of about 10 nm (ref. 3). Such labeled proteins can then be affinity purified on a streptavidin matrix and identified by mass spectrometry.

BioID is a valuable and widely used tool but also suffers from drawbacks. First, it is rather bulky (approx. 36 kDa), which sometimes hampers the proper localization of the corresponding fusion proteins. Second, its kinetic of labeling is slow, necessitating hours of incubation with exogenous biotin. Finally, as E. coli’s BirA is a class II biotin ligase, it has an N-terminal DNA-binding domain that shows structural homology to linker histone H5 and may lead to artefactual non-specific binding to the host DNA and/or chromatin-interacting proteins.

To circumvent these issues, new enzymes were developed (reviewed in 7). BioID2 is derived from Aquifex aeolicus and is the smallest enzyme for PDB described to date (26.4 kDa). As a class I PBL, it also lacks a DNA-binding domain. BASU is a variant of the B. subtilis BirA that lacks its DNA-binding domain and was claimed to
show 1000 fold faster biotinylation kinetics than BioID\(^8\), though this has been disputed\(^9\). A directed evolution approach led to TurboID and miniTurbo, two super active variants of BioID\(^9\). TurboID has the size of BioID and is the most active enzyme described to date supporting labeling kinetics down to 10 min\(^9\). This strong activity, however, comes at the cost of a high background labeling before addition of extra biotin to the medium\(^9,10\). miniTurbo is smaller than TurboID and does no show high background activity. However, miniTurbo has about half the activity of TurboID\(^9\) and was reported as unstable\(^10\). Furthermore, both TurboID and miniTurbo can lead to toxicity when constitutively expressed in mammalian cell culture or model organisms\(^9,10\).

Finally, AirID is an enzyme faster than BioID and less toxic than TurboID but still requiring hours of labeling time\(^11\). Altogether, each of these enzymes come with advantageous properties and some drawbacks. Clearly, an enzyme that combines most advantages would be a valuable tool for PDB.

Here we present two novel enzymes: microID, a truncation variant of BioID2, and its directed evolution-deduced variant ultraID. With a molecular weight below 20 kDa, microID and ultraID are by far the smallest PDB enzymes available to date. Moreover, ultraID exhibits enzyme kinetics similar to TurboID but with less background activity. UltraID supports efficient labeling in mammalian cell culture, \textit{E. coli} and S. \textit{cerevisiae}. In a direct comparison in which the protein Argonaute-2 (Ago2) was either fused to ultraID or BioID and expressed at physiological levels, we show that a 10 min labeling time with ultraID can substitute for a 24h labeling with BioID. Finally, using ultraID, we probed the membrane-associated interactome of the COPI coat protein complex coatomer in living cells. Altogether, ultraID is the smallest enzyme for PDB and supports efficient short labeling with reduced background activity.

**Results**

**microID, a small size biotin ligase derived from BioID2**

We previously developed a split-BioID assay in which BioID is split into two inactive fragments that reassemble an active enzyme when fused to two interacting proteins\(^{12}\). We explored whether a similar assay could be set up with BioID2. Based on the structures of both proteins we split BioID2 between amino acids K\(^{171}\)/S\(^{172}\) (Fig. 1A), corresponding to the split-BioID site (E\(^{256}\)/G\(^{257}\)). The resulting fragments NBioID2 (BioID2 [2-171]) and CBioID2 (BioID2 [172-233]) were then respectively fused to the proteins FKBP (12-kDa FK506 binding protein) and FRB (FKBP-rapamycin binding domain) that interact in the presence of rapamycin\(^{13}\). To test split-BioID2, the fusion proteins were expressed in HeLa cells and PDB analyzed in the presence or absence of rapamycin (Fig. 1B). Biotinylation was observed irrespective of rapamycin treatment (Fig. 1C, lanes 1 and 2). In addition, CBioID2-FRB was hardly detectable and only in the presence of rapamycin (Fig. 1C, lanes 1 and 2), suggesting that it is unstable. This suggested that NBioID2, which contains the enzyme catalytic site, retains biotinylation activity. This is indeed the case as when NBioID2-FKBP was expressed in the absence of CBioID2-FRB, similar biotinylation activity was observed to when both NBioID2 and CBioID2 fusions were expressed (Fig. 1C, compare lane 3 to 1&2). NBioID2, with a size of 19.7 kDa, is the smallest enzyme for PDB to date and is termed from now on microID (or µID).
**Figure 1:** A truncation fragment of BioID2 with promiscuous biotinylation activity. (A) BioID and BioID2 crystal structures (Protein Data Bank 1BIB and 2EAY) showing the splitting site of split-BioID (E^{256}/G^{257}) and its corresponding position in BioID2 (K^{171}/S^{172}). (B) Principle of the rapamycin-induced dimerization set-up to test split-BioID fragments: FKBP is fused to the N-terminal fragment and FRB to the C-terminal fragment. PDB is tested in the presence or absence of rapamycin. (C) Blots of lysates of HeLa cells transiently transfected with the indicated constructs and cultivated for 24h in biotin-containing medium in the presence or absence of rapamycin. Biotinylation was analyzed using IRDye680-labeled streptavidin and expression levels of the fusion proteins with antibodies against FLAG and Myc tag as indicated. The asterisk (*) indicates a non-specific cross reactivity.

**microID shows efficient biotinylation at short labeling times**

We then analyzed how microID performs in comparison to other PDB enzymes. First, we compared BioID, BASU and microID. After a 24 h labeling time (Supplementary Fig. 1, left), all three enzymes led to promiscuous biotinylation. By contrast only BASU and microID showed clear biotinylation after much shorter labeling times of 1h (Supplementary Fig. 1, middle) and 10 min (Supplementary Fig. 1, right). Hence, in addition to its small size, microID is an enhanced activity enzyme in comparison to BioID.

In a second benchmarking round, we compared BioID2, BASU, TurboID and microID (Figure 2). We did not include miniTurbo as it was reported to be less active than TurboID\(^9\) and highly unstable\(^10\). To assess the background activity of each enzyme, we compared the biotinylation obtained in the absence or presence of additional biotin for each timepoint. All four enzymes showed clear biotinylation over background levels at 24 h (Figure 2, left) and 1 h (Figure 2, middle) labeling time. At 10 min labeling, the clearest biotinylation over background was observed for BASU, TurboID and microID (Figure 2, right). Note that in these experiments the expression levels of TurboID were lower than those of the other enzymes which likely led to lower resulting biotinylation (Figure 2, anti-Myc signal).
Figure 2: microID shows efficient biotinylation at short labeling times. Blots of lysates of HeLa cells transiently transfected with the indicated constructs and incubate with 50 µM biotin for 24 h (left), 1 h (middle) or 10 min (right). Biotinylation was analyzed using IRDye680-labeled streptavidin and expression levels of the fusion proteins with antibodies against the Myc tag as indicated.

MicroID, like its parental enzyme BioID2\textsuperscript{4}, requires lower concentrations of added biotin than the standard 50 µM concentration used with BioID. Indeed, we observed clear biotinylation activity down to 0.5 µM biotin after a 2h labeling time (Supplementary Fig. 2). This property may prove useful for PDB experiments performed in model organisms in which providing an excess of biotin can be problematic.

Together, these experiments reveal microID as a novel small size enzyme that supports efficient PDB at short labeling times and reduced biotin concentrations.

Directed evolution of microID

We next tried to improve the activity of microID through protein engineering. A single substitution, R118G, renders BirA of *E. coli* abortive. Situated in a conserved biotin/bio-AMP binding site, the corresponding mutation (R40G) had been introduced into BirA of *A. aeolicus* to obtain BioID2\textsuperscript{4}. Substituting R118 with a serine instead of a glycine results in twofold increased activity in BioID\textsuperscript{9}. We thus tested if the corresponding mutation, R40S, also leads to enhanced activity in microID. However, a side-by-side comparison revealed that R40S rather leads to weaker biotinylation when compared to R40G (Supplementary Fig. 3).

The hyper active enzyme TurboID was obtained from a yeast display directed evolution approach using BioID R118S as a starting point. We reasoned that we could engineer a higher activity enzyme using a similar strategy applied to microID. To this end, we generated, through error-prone PCR\textsuperscript{14}, a surface display library of approximately 8×10\textsuperscript{7} random microID mutants with an estimated average mutagenesis rate of 2.2 amino acid exchanges per variant. The random variants were
expressed as 6xHis-tagged fusion proteins to Aga2p, a yeast protein that is exposed to the cell surface through its disulfide bonds linkage to Aga1p (Figure 3A).

Figure 3: Directed evolution of microID. (A) Scheme of the biotinylation assay. microID random variants were surface-presented via Aga1p:Aga2p. The yeast surface was biotinylated in presence of ATP and biotin. The His-tag served as a presentation marker. Biotin was detected with a streptavidin-APC conjugate. (B) Density plots of surface presentation (Y-axis) vs. cell surface biotinylation (X-axis). The error-prone PCR based-yeast surface display library was screened by stepwise decreased biotinylation time. Upper row: Comparison of microID with the initial library (R1) after a 17 h biotinylation assay. Bottom row: same with the Round 3 (R3) library after a 10 min biotinylation assay. (C) Single clone analysis from the Round 3 library. (D) Positions of the mutated residues (structure based on PDB: 3EFR processed with the program ChimeraX). (E) Activity measurement of microID variants with an ELISA-based biotinylation assay. Serum albumin biotinylation mediated by the indicated enzymes was measured through the absorbance at 405 nm and was normalized to the activity of microID (set to 100%). Error bars are standard deviation. N=5; n=3.

A cell surface biotinylation assay was performed by incubating the ligase-expressing cells with 50 μM biotin and 2.5 mM ATP for 17h as a starting point. After cell staining using fluorescently labeled streptavidin (ligase activity) and anti-pentaHis antibodies (surface presentation), three consecutive FACS rounds with decreasing labeling times were conducted, sorting those cells that exhibited a strong fluorescence signal for biotin ligase activity and clear surface presentation (Figure 3A and B). As a result, 0.77, 2.24 and 0.53 % of the cells were enriched in round 1, 2 and 3, respectively. MicroID, which served as a control, showed negligible activity when the reaction time was shorter than 1h (Supplementary Fig. 4). By contrast, after three rounds of enrichment, sorted cells showed significant activity after a 10 min labeling time (Figure 3B). Eight individual clones of the final selection round were selected and
analyzed by flow cytometry. The six clones that showed the strongest activity (Supplementary Fig. 5) were sequenced, revealing three individual mutants (Figure 3C). Due to a single nucleotide deletion leading to a stop-codon loss, one microID variant was extended by 22 amino acids at its C-terminus (NSSRSDNNSVDVTSTLFPLYF). Interestingly, the other two identified triple mutants shared a L41P mutation. This site is located next to the R40G mutation within the biotin-binding site. All other substitutions are distributed at the surface of the enzyme, away from the catalytic site (Figure 3D).

To investigate whether the observed enhanced cell biotinylation was the result of increased labeling of proximal proteins rather than enhanced self-biotinylation of the microID variants, enzymes were removed from the cell surface after the biotin ligase reaction by reducing the disulfide bonds between Aga1p and Aga2p with DTT. Surface exposition signals were considerably reduced after DTT treatment (Supplementary Fig. 6) while cell surface biotinylation persisted for both triple mutants but not for the C-terminally extended variant (Supplementary Fig. 6). Most likely, the C-terminal extension, which contains a lysine residue serves as an acceptor labeling substrate, eventually resulting in an enhanced overall biotinylation. Therefore, this variant was excluded from further analysis. The triple mutants were named ultraID-4 (with the mutations S24C, L41P, K169R) and -5 (L41P, K115M, K156E).

The biotinylation activity of ultraID-4 and -5 was then assessed in HeLa cells after 10 min labeling time and compared to microID, the template for directed evolution, and TurboID, the most active enzyme described so far (Figure 4A). Both variants yielded biotinylation signals stronger than that of microID and seemingly similar to that of TurboID. However, as enzyme expression differed between the ultraID clones and TurboID, precise conclusions about their biotinylation activities were not drawn at this point.

![Figure 4: ultraID variants show improved activity compared to microID.](https://doi.org/10.1101/2021.06.16.448656)
The mutation L41P is responsible for the enhanced activity of the ultraID variants

The directed evolution strategy that led to TurboID resulted in 14 mutations on the original BioID sequence. None of them are close to the catalytic pocket and they do not cluster to a special part of the enzyme. Hence there is no obvious explanation on how these mutations lead to a higher activity apart that they somehow indirectly affect the catalytical site. The outcome of our directed evolution effort is quite different: ultraID-4 and -5 are triple mutants of microID that share an L41P mutation. Since L41 is adjacent to the R40G mutation within the biotin-binding site of microID, we reasoned that the enhanced activity of the ultraID variants can probably be mapped down to the single additional L41P mutation in microID. To address this, ultraID-4 and ultraID-5 as well as microID, microID R40G/L41P and a microID L41P-only variant were produced as recombinant proteins in E. coli (Supplementary Fig. 7A). Enzymatic activity was then determined in an ELISA-based proximity-dependent biotinylation assay (see methods). In accordance to the single clone cell surface display analysis, ultraID-4 and 5 showed an increased activity of 171% and 187% respectively compared to microID (Figure 3E). Interestingly, the microID L41P-only mutant, with wild-type R40 position, showed a higher biotinylation activity than microID (Figure 3E). Hence L41P, like R40G, is sufficient to render the BioID2 [2-170] fragment abortive. The effects of both mutations are additive as the microID variant with combined R40G/L41P mutations showed the highest activity of all tested variants (207%). Similar observations were made in HeLa cells in which both microID (R40G) and microID L41P-only (with intact R40 position) showed biotinylation activity after a 10 min incubation with biotin, and combining both mutations resulted in even stronger biotinylation signals (Figure 4B). Moreover, similar to microID (Supplementary Fig. 2), this latter variant also shows clear activity at reduced concentrations of exogenously added biotin (Supplementary Fig. 8).

To further characterize the enzymes, we determined their melting temperatures (Tm) with a thermal shift assay (see methods). Compared to microID (Tm = 64°C), ultraID-5 showed decreased thermal stability (Tm = 52°C) whereas ultraID-4 showed two independent melting points (55°C and 64.5°C), which could be a consequence of the S24C mutation that might promote the formation of an intermolecular disulfide bond. The microID L41P-only variant had a slightly decreased stability with a melting point of 60.5°C whereas the double mutations R40G/L41P in microID-R40G/L41P did not impact its thermal stability (Tm = 63°C) compared to microID (Supplementary Fig. 7B).

Altogether, the combination of R40G and L41P is the minimal set of mutations required for enhanced enzymatic activity, and the corresponding variant has a thermostability profile comparable to microID. The variant microID-R40G/L41P was thus selected for further characterization and is from here on referred to as ultraID.

ultraID outperforms current enzymes for PDB

Having defined ultraID with its minimal set of mutations, we compared its activity to that of other PDB enzymes. BioID, BioID2, BASU, TurboID, microID and ultraID were expressed in HeLa cells, and the extent of protein biotinylation after 24 h, 1 h and 10 min labeling time was determined by western blot and compared to respective samples without biotin supplementation (Figure 5A). Each experiment was
replicated for quantitative estimation of the biotinylation activity. Three additional experiments for TurboID and ultraID with 10 min biotinylation were included in the quantitative analysis. The signals for streptavidin were integrated over each lane, excluding the band corresponding to the self-biotinylation of the ligases, and were normalized to the expression levels of each ligase (Myc signal).

**Figure 5: ultraID outperforms current PDB enzymes.** (A) Blots of lysates of HeLa cells transiently transfected with the indicated constructs and incubated (+) or not (-) with 50 µM biotin for 24h (left), 1h (middle) or 10 min (right). Biotinylation was analyzed using IRDye680-labeled streptavidin and expression levels of the fusion proteins with antibodies against the Myc tag as indicated. (B) Quantification of (A). Circles and squares of the same color belong to the same replicate experiments.

All the ligases showed a clear activity over background biotinylation after a 24 h and 1 h labeling time with ultraID and TurboID having the strongest relative activity. With a 10 min labeling time, hardly any biotinylation over background levels was observed for BioID and BioID2 whereas BASU and microID still showed a clear activity. All the enzymes were however outperformed by TurboID and ultraID that yielded 5-9 times higher biotinylation than microID and BASU under these conditions. The average relative signal for TurboID in the biotin-treated sample after a 10 min pulse labeling was on average ca. 25% higher than that for ultraID, though this was not consistent in all experiments (Figure 5B). In accordance with previous reports⁹,¹⁰, the high activity of TurboID came however at the cost of a high background signal from the sample without biotin addition that was consistently observed in all experiments and on average almost 3 times higher than for ultraID (Figure 5A, B).
Together, our data show that ultraID yields a labeling efficiency similar to TurboID at labeling times down to 10 min, with the advantage of a lower background biotinylation activity, making it a more efficient enzyme for PDB.

**PDB with microID and ultraID in yeast and bacteria**

An important feature of novel enzymes for PDB is their suitability for various model organisms. As the directed evolution of microID was based on yeast display, it was expected that the enzymes are also active in *Saccharomyces cerevisiae*. As we previously applied BioID in yeast to probe the microenvironment of the ribosomal protein Asc1/RACK1\(^{15}\), experiments in *S. cerevisiae* were conducted in which BioID, microID and ultraID were fused to Asc1. Comparisons of the three enzymes showed that while they all work in yeast, microID and ultraID showed stronger activity than BioID (Figure 6A). We note, however, that in *S. cerevisiae* expression of the microID and ultraID fusion proteins both lead to some background activity from the low biotin concentrations (8.2 nM) in the YNB medium without additional biotin (- condition on Figure 6A).

---

**Figure 6: PDB in yeast and bacteria with microID and ultraID.** (A) Blots of lysates of *S. cerevisiae* stains expressing the indicated Asc1 fusion proteins and incubated or not (-) with 10 µM biotin for the indicated times. The fusion proteins were expressed from plasmids in a Δasc1 strain, and the ASC1 wild-type strain carrying the empty vector served as a control. Biotinylation was analyzed using HRP-labeled streptavidin and expression levels of the fusion proteins with antibodies against Asc1 as indicated. (B) Blots of lysates of *E. coli* cells transformed with expression plasmids for ultraID, microID or the control unrelated protein CNOT9. The cells had been incubated (+) or not (-) with biotin for 16 h. The asterisk (*) indicate the signal for the endogenous bacterial biotinylated protein BCCP.
We also tested the enzymes in *E. coli*. Also in this organism, a clear biotin-dependent labeling activity was observed upon expression of microID and ultraID but not of the non-related protein CNOT9 (Figure 6B).

In conclusion, both microID and ultraID are well suited for PDB in bacteria, yeast and mammalian cell culture.

**UltraID supports PDB-MS within a 10 min timeframe in authentic experimental conditions**

To investigate the performance of ultraID in authentic conditions, in which the enzyme is fused to a target protein expressed at physiological level, we then established doxycycline-inducible cell lines that express ultraID fused to the N-terminus of Argonaute-2 (Ago2) from a single gene locus as we described previously. To filter background biotinylation from the proteomic data, we constructed an additional control cell line that expressed ultraID fused to the unrelated protein Rab11. As a comparison, stable cell lines for BioID-Ago2 and BioID-Rab11 were already available in our laboratory. Expression levels were tuned so that the Ago2 fusion proteins were expressed close to endogenous levels (Supplementary Fig. 9) and Ago2 and Rab11 fusions yielded similar biotinylation levels. The fusion proteins showed a localization pattern typical for Ago2 and Rab11 in indirect immunofluorescence experiments (Supplementary Fig. 10). PDB was then performed upon addition of exogenous biotin for 10 min (ultraID fusions) or 24h (BioID fusions) to four biological replicate samples. The resulting biotinylated proteins were processed for MS analysis after their capture with trypsin-resistant streptavidin beads and on-beads trypsin digestion (Figure 7A). Peptide identification was performed with the MaxQuant software and the relative enrichment of each protein across samples was quantified by label-free quantification (LFQ) using the proDA (probabilistic drop out analysis) package to infer for each protein its mean LFQ intensity and associated variance, taking into account the missing values. To obtain the proxome of Ago2, we tested the differential abundance of proteins between the Ago2 and Rab11 datasets by comparing the inferred LFQ mean intensities for the two conditions (Figure 7B). When examining the data, we noticed that the LFQ-based enrichment analysis occasionally missed some valid candidates that were detected in all Ago2 replicate samples but in none of the Rab11 control samples. We thus also added to the list of true-positive hits those proteins that did not pass the thresholds of the LFQ analysis but fulfilled the following criteria: detection in all four replicates of Ago2 samples with at least two peptides in three replicate samples and no peptide detected in any Rab11 sample. The complete list of hits for the BioID and ultraID datasets is provided in Supplementary Data 1.

A side-by-side comparison of BioID and TurboID in which both enzymes were fused to the same specific bait proteins revealed that a 10 min labeling TurboID-derived dataset is larger than an 18 h labeling BioID-derived dataset but that the additional hits are mainly non-relevant proteins. This is probably due to a permanent background labeling activity of TurboID leading to a much wider labeling radius. Our data suggest that ultraID is not affected by this shortcoming. Indeed, the ultraID-derived Ago2 interactome obtained after 10 min labeling was somewhat smaller (50 vs 68 hits) than that obtained from BioID after 24 h labeling (Figure 7B) but ultraID performed equally well to BioID in identifying known Ago2-associated proteins such as...
as TNRC6 proteins\textsuperscript{21}, Dicer\textsuperscript{22}, CNOT1\textsuperscript{23, 24}, GIGYF proteins\textsuperscript{12} or XRN1\textsuperscript{25} (Figure 7B, C) and the volcano plots resulting from both datasets were similar (Figure 7C).

Hence, ultraID allows obtaining relevant proteomic datasets in PDB experiments with a labeling time of 10 min at physiological expression levels.

Figure 7: 10 min labeling with ultraID under experimental PDB conditions. (A) Experimental setup for the side-by-side comparison of 24h labeling with BioID vs. 10min labeling with ultraID. (B) Venn diagram displaying the proteomic dataset sizes and overlap of the significant hits from (A), the gene names of the dataset overlap are specified. (C) Volcano plots of the proteins identified by LC-MS/MS in the BioID and ultraID experiments. The Rab11 datasets served as negative control. Significant hits (p-value < 0.01 and log\textsubscript{2} fold change > 2) are indicated in red. Some hits, already known as Ago2 interactors or previously identified in other BioID datasets are labeled.

Defining the membrane-associated interactome of coatomer with ultraID

To further validate the practicality of ultraID we analyzed a membrane-associated interactome of coatomer, the COPI coat protein complex. During the formation of COPI transport vesicles, the small GTPase Arf1 (or one of its paralogs) recruits coatomer to the Golgi membrane when bound to GTP. There, coatomer polymerizes and together with Arf1 promotes membrane deformation, cargo protein capture and membrane scission. As a cargo-loaded transport vesicle sheds from its donor membrane, GTP-hydrolysis by Arf1, stimulated by its ArfGAPs, leads to the release of coatomer to the cytosol\textsuperscript{26} (Supplementary Fig. 11). Because the active form of coatomer is the population that transiently associates with endomembranes, it is not amenable to proteomic studies using affinity purification approaches. Defining a
relevant interactome of coatomer has thus been challenging. Currently available datasets came from cellular fractionation assays or reconstitution experiments in which COPI vesicles were purified, out of their native context under non-physiological conditions\textsuperscript{27, 28}. To probe the interactome of coatomer in living cells, we used ultraID as a C-terminal fusion to the \( \gamma \)-COP subunit of coatomer. Of the seven coatomer subunits, we picked \( \gamma \)-COP because it can be tagged at its C-terminus without affecting its functionality\textsuperscript{29} and because this subunit exist as two paralogs, \( \gamma \text{I}- \)COP and \( \gamma \text{II}- \)COP, with overlapping but not fully identical functions\textsuperscript{29, 30}.

To ensure a physiological relevance of the datasets, \( \gamma \text{I}- \)COP-ultraID and \( \gamma \text{II}- \)COP-ultraID were expressed as inducible rescue constructs respectively in \( \gamma \text{I}- \)COP and \( \gamma \text{II}- \)COP knock-out P19 mouse pluripotent cells\textsuperscript{29}, with their expression levels tuned to those of the endogenous proteins (Figure 8A). To filter background biotinylation from the proteomic data, we constructed an additional control cell line that expressed ultraID fused to Ago2 (Figure 8A). The drug brefeldin A (BFA) inhibits the activation of Arf1 at the Golgi membrane and thus prevents association of coatomer with membranes\textsuperscript{31, 32} (Supplementary Fig. 11), which leads to the collapse of the Golgi into the ER\textsuperscript{33}. In indirect immunofluorescence experiments \( \gamma \text{I}- \)COP-ultraID and \( \gamma \text{II}- \)COP-ultraID showed the expected BFA-sensitive Golgi localization assessed by co-localization with the Golgi marker protein GM130 (Supplementary Fig. 12 upper and middle panels). By contrast, in both mock- and BFA-treated cells, Ago2-ultraID showed a punctuated staining pattern (Supplementary Fig. 12 lower panel) reflecting the partial localization of Ago2 to P-bodies\textsuperscript{34}. Moreover, whereas in the mock condition biotinylated proteins by \( \gamma \text{I}- \)COP-ultraID and \( \gamma \text{II}- \)COP-ultraID after PDB were concentrated in the Golgi area, they showed a diffuse cytosolic staining under BFA-treatment (Supplementary Fig. 12).

PDB was then performed on the three cell lines under mock- or BFA-treatment conditions (for the \( \gamma \text{I}- \)COP-ultraID and \( \gamma \text{II}- \)COP-ultraID cell lines) with three biological replicates for each condition (total of 15 samples, see Fig. 8B). To assess the specificity of the proteomic datasets obtained after a longer labeling time with ultraID, cells were incubated with biotin for 4h. The proteomic data were analyzed as described above. We first considered the \( \gamma \text{I}- \)COP and \( \gamma \text{II}- \)COP proxiones under mock-treated conditions in which coatomer is present in both its membrane-associated and cytosolic forms (Supplementary Fig. 13 and Supplementary Data 2) and identified 56 and 53 proximal proteins to \( \gamma \text{I}- \) and \( \gamma \text{II}- \)COP respectively. Of these, 59\% of \( \gamma \text{I}- \)COP proximal proteins and 42\% of \( \gamma \text{II}- \)COP’s have a known Golgi localization. Among them, we found as expected the other COP subunits of the coatomer complex, the paralogs of the small GTPase Arf1 and their three ArfGAPs.
Figure 8: Probing a membrane-associated interactome with ultraID. (A) Blots of lysates of P19 cells stably expressing ultraID-Ago2, of P19 cells knock-out for endogenous γ1-COP and rescued by γ1-COP-ultraID, and of P19 cells knock-out for γ2-COP and rescued by γ2-COP-ultraID lines. The fusion proteins were detected with antibodies against γ1- and γ2-COP and biotinylation with IRDye680-labeled streptavidin. (B) Experimental setup for the determination of the membrane-associated interactome of γ1-COP and γ2-COP. (C) Volcano plots of the proteins identified by LC-MS/MS with the mock vs. BFA samples. Significant hits (p-value < 0.01 and log₂ fold change > 2 in this analysis, and p-value < 0.01 and log₂ fold change > 2 in the γ-COP vs Ago2 comparison under mock conditions), are indicated with gene names in red. (D) Proteins identified as membrane interactors of γ1-COP and γ2-COP (see text).
To gain a better insight into the potential interactome of membrane-associated coatomer we considered the proteins enriched in the mock condition compared to the BFA-treated samples in which coatomer is released from Golgi membranes (Fig. 8C). Proteins that are enriched in the mock samples are indeed likely to represent membrane-associated proteins proximal to coatomer. With this analysis, we identified 21 and 17 proximal proteins to membrane-associated γ1- and γ2-COP respectively (Fig. 8D and Supplementary Data 2). Remarkably, all but one are known Golgi proteins demonstrating the specificity of the assay. The only non-Golgi protein is emerin, a membrane-anchored protein of the inner nuclear membrane (INM). Another INM membrane protein, Sun2, uses an arginine-based COPI-binding signal that mediates retrieval from the Golgi and participates in nuclear localization. Such an arginine motif is also found on other INM proteins including emerin\(^ {35} \). As a partial Golgi localization has been described for emerin\(^ {36} \), there may be a substantial amount of that protein at the Golgi in P19 cells that, like Sun2, relies on COPI vesicles to reach the INM.

**ultraID identifies expected membrane-associated interactors of coatomer**

Unsurprisingly, Arf proteins and their ArfGAPs, which are respectively responsible for coatomer recruitment to and release from membranes\(^ {26} \), were identified as membrane-associated interactors of γ1- and γ2-COP. Golgins are long membrane-anchored coiled-coil proteins that tether specific transport vesicles to distinct regions of the Golgi apparatus\(^ {37} \). In cooperation with additional proteins, golgins then stimulate the fusion of the tethered vesicles with their target membrane by allowing the formation of a tight SNARE complex\(^ {38} \). Accordingly, the Golgi-localized SNAREs syntaxin-5 and GOS-28, both involved in COPI-dependent intra-Golgi trafficking\(^ {39} \), as well as SLY-homolog, an SM-family protein that is required for Golgi SNARE pairing\(^ {40} \), were identified in the membrane-associated datasets. Golgins known to interact with COPI vesicles: giantin\(^ {41} \) and TMF\(^ {42} \) (found in both datasets), as well as golgin-B4\(^ {43} \) (found in the γ1-COP datasets) were also identified. In accordance with the role of COPI vesicles in retrograde intra-Golgi trafficking these golgins are located at the Golgi rims\(^ {44} \). Interestingly, we also identified trans-Golgi network (TGN) localized golgins (GCC88, GCC185, and golgin-245) as membrane-associated interactors of coatomer. These are normally involved in tethering vesicles coming from endosomes\(^ {45} \). Combined with the recent identification of a COPI-mediated transport route from endosomes to TGN in yeast\(^ {46} \), this calls for further investigating a role of the COPI pathway in endosomal trafficking in mammals that was suggested by several reports more than 20 years ago\(^ {47, 48} \).

Scyl1 in association with GORAB recruits coatomer to specific Golgi domains\(^ {49} \) and Scyl3 is known to interact with coatomer\(^ {50} \). In our datasets Scyl1 is associated with both γ1-COP and γ2-COP at the membrane whereas Scyl3 is rather specifically associated with γ2-COP. This is in apparent contradiction with previous reports that indicated a specific interaction of Scyl1 with γ2-COP\(^ {51} \), and interaction of Scyl3 with both paralogs\(^ {50} \). The discrepancy might come from the different cell lines used in the different studies. Additional proteins specifically proximal to membrane-associated γ1-COP were WWOX a protein that localizes to the Golgi but its function there is currently unknown\(^ {52, 53} \), ATP7A, a copper transporter that resides at the TGN\(^ {54} \), TMEM165, a Golgi-localized transporter for Mn\(^ {2+} \) which is an important co-factor of glycosylation enzymes, and ERGIC-53. ERGIC-53 is a cargo receptor that constitutively cycles, through COPII and COPI vesicles, between the ER and the ER-Golgi intermediate
compartment (ERGIC), and can be found at low concentrations at the cis-cisternae of the Golgi but not beyond\(^{45}\). Transport vesicles that act at the cis-side of the Golgi are tethered by golgin-84\(^{45}\). Interestingly, we identified both ERGIC-53 and golgin-84 as specific membrane-associated proximal proteins to \(\gamma_{1}\)-COP. This is in agreement with the observation that \(\gamma_{1}\)-COP-containing coatamer is preferentially localized to the cis-side of the Golgi whereas \(\gamma_{2}\)-COP-containing coatamer localizes preferentially at the trans-side\(^{30}\).

In summary, ultraID labeling allowed us to define a membrane-associated interactome of coatamer that faithfully reflects its expected proximal landscape.

**Discussion**

Here we describe ultraID, a novel engineered enzyme for efficient PDB in living cells. UltraID combines several advantages of previously described enzymes: it lacks a DNA-binding domain like BioID2, BASU and miniTurbo, like BioID2 it is active at lower concentrations of added biotin, and it allows robust biotinylation within short labeling times to levels similar to TurboID. To date, ultraID is by far the smallest enzyme for PDB, 44% smaller than TurboID/BioID and 25% smaller than BioID2 (19.7 vs. 35 and 26.4 kDa respectively). This is also advantageous as it is generally considered that the larger the size of a tag, the greater the chance of interfering with the function, trafficking and interactions of the fusion protein\(^{56}\). Finally, ultraID produces markedly less labeling background than TurboID when no exogenous biotin is added. This is important as the background activity of TurboID can lead to a considerable loss of specificity\(^{10}\).

Using ultraID we analyzed the remodeling of the interactome of COPI coat proteins after incubation with BFA. This allowed us to define a membrane-associated interactome. As transport vesicles are transient carriers that rely on dynamic protein/membrane interactions for their biogenesis and functions, defining their proteome is a challenging task. Previously, a landmark core proteome of COPI vesicles was defined based on the LC-MS analysis of isolated COPI vesicles generated in vitro from permeabilized cells and recombinant coat proteins\(^{28}\). A striking difference between this dataset and ours is the presence of glycosylation enzymes and cargo receptor proteins in the former, and the presence of many golgins and the membrane recruitment machinery in the latter. This illustrates the advantages and drawbacks of each approach. The in vitro isolation removes vesicles from their native cellular environment. As a consequence, it is less likely to capture the interactions that occur at the Golgi membrane. The PDB approach allowed us to probe the direct neighborhood of \(\gamma_{1}\)- and \(\gamma_{2}\)-COP in living cells, but it also limited our ability to analyze the content of COPI vesicles. Indeed, soluble cargo proteins inside the vesicles are inaccessible to biotAMP released in the cytosol. Similarly, transmembrane COPI cargo proteins such as Golgi-resident glycosylation enzymes typically expose short tails to the cytosol whereas the largest parts of the proteins are inserted in the lumen where they exert their functions\(^{57}\). This limits the amount of potentially available acceptor lysine residues and thus the likelihood for biotinylation to occur. All in all, this stresses that complementary approaches are needed to comprehensively define the proteome and interactome of transport vesicles.

Potential specific functions of the \(\gamma_{1}\)- and \(\gamma_{2}\)-COP paralogs have remained elusive. In vitro data suggest functional redundancy\(^{28, 58}\) but cellular assays point to
some differences\textsuperscript{29, 30}. Notably, we showed that \( \gamma_1 \)-COP plays a unique role in promoting neurite outgrowth during the neuronal differentiation of P19 cells\textsuperscript{29}. The molecular mechanisms underlying this observation are currently unknown. Having established ultraID-tagged P19 cell lines, it will now be possible to analyze the interactome of \( \gamma_1 \)- and \( \gamma_2 \)-COP during neuronal differentiation and possibly to identify differences that may help understand their specific functions in this context.

To summarize, we have developed ultraID a compact and hyperactive enzyme for PDB that favorably compares to previously available BPLs. Ultimately, ultraID is the product of two engineering steps: the removal of the C-terminal domain of the class I BPL BioID2, and the addition of one point mutation (L41P) to the canonical R40G substitution within the biotin-binding domain. The position of the L41P substitution suggests that it strengthen the destabilization of the biotin-binding site induced by the R40G substitution, thereby stimulating the release of bioAMP. This is supported by the observation that L41P alone is sufficient to render the truncated BioID2 [2-171] abortive and that the effects of R40G and L41P are additive. The engineering of ultraID may thus provide a simple framework to improve the activity of PDB enzymes based on class I BPLs. It also suggests that specifically targeting the biotin-binding site of BPLs for random mutagenesis may be a promising approach to generate hyperactive enzymes.

\section*{Methods}

\textbf{Plasmids and antibody}
The plasmids, yeast strains and antibodies used in this study are provided in the Supplementary Tables 1–3

\textbf{Mammalian cell culture}
All cell culture experiments were conducted with low-passage HeLa 11ht cells\textsuperscript{59} obtained from Dr. Kai Schönig (Zentralinstitut für Seelische Gesundheit Mannheim, Germany) or P19 cells obtained from Sigma-Aldrich. HeLa cells were maintained as described in\textsuperscript{12} and P19 cell as described in\textsuperscript{29}. The cells were regularly tested for mycoplasma contamination.

\textbf{Construction of stable cell lines}
Fip-recombinase-mediated construction of stable HeLa 11ht cell lines was performed as described in ref.\textsuperscript{59}. P19 stable cell lines were constructed using the Piggy-BAC transposon system as described in ref.\textsuperscript{29}.

\textbf{Immunofluorescence microscopy}
9'000 HeLa cells or 20'000 P19 cells per well were seeded in 8-well \( \mu \)-slides (Ibidi). On the following day, the expression of the ligase fusion proteins was induced with doxycycline. For BioID HeLa cell lines, 50 \( \mu \)M biotin was added at the same time as doxycycline and cells were incubated at 37\( ^\circ \)C for 24 h. For ultraID HeLa cell lines, the biotinylation time was 10min one day after induction. For P19 cells, BFA-treated samples were first incubated for 1h with BFA and then for an additional 4h with BFA and biotin. The mock samples were only incubated with biotin for 4h. After
biotinylation, cells were carefully rinsed with ice cold PBS, subsequently fixed and permeabilised with 150 μL ice-cold methanol and incubated at -20 °C for 10 min. All further steps were performed at RT. Methanol was removed and the cells were washed three times with PBS. The samples were blocked with PBS + 5 % bovine serum albumin (BSA, Carl Roth) (blocking solution) for 15 min. Incubation with the relevant primary antibodies was performed for 30 min at RT or overnight at 4°C. The cells were washed three time with blocking solution and incubated with the secondary antibodies and streptavidin-AlexaFluor647 for 30 min in darkness. Then, the cells were washed three times with PBS. After a 10 min staining with DAPI (0.1μg/mL in PBS) and brief washing with water, the cells were mounted with Mowiol mounting medium. After incubation at RT overnight, the slides were stored at 4°C until imaging. Microscopy images were acquired with a Nikon Eclipse Ti2 spinning disk confocal microscope (Nikon Imaging Center, Heidelberg) using the acquisition software Volocity 6.3 (Perkin Elmer).

**Screening for biotinylation**

*In HeLa cells.* Screening for biotinylation was performed as described in ref. 60, 61.

*In S. cerevisiae.* Cells were transformed with the pME4478 (ASC1-βirA*), pME5086 (ASC1-μID) and pME5087 (ASC1-ultraID) plasmids and cultivated in yeast nitrogen base medium (YNB: 1.5 g/L YNB without amino acids, 5 g/L ammonium sulfate, 2% glucose) with 20 mg/L L-tryptophan at 30°C. Yeast strains used in this work are listed in Supplementary Table 2. After incubation with biotin (50 μM), cells were lysed through alkaline treatment and subsequent boiling in SDS-PAGE loading buffer62. An equal number of cells of each sample (volume of culture harvested in mL = 5 divided by OD600 of the culture) was used to generate cell lysates. The cells were harvested by centrifugation, washed with water and resuspended in 400 μL 0.1 M NaOH. After incubation for 5 min at room temperature, samples were centrifuged, supernatants removed, and 100 μL of 1:4 diluted loading dye (0.25 M Tris-HCl pH 6.8, 30% glycerol, 15% β-mercaptoethanol, 7% SDS, 0.3% bromophenol blue) added. Samples were incubated for 3 min at 95°C, centrifuged at 16,200 x g for 5 min, and 5 μl of the supernatants were subjected to SDS-PAGE followed by protein transfer onto nitrocellulose membrane. Proteins on the membrane were stained with Ponceau Red (0.2% PonceauS, 3% trichloroacetic acid), and blocked with either 5% milk powder or 1% BSA dissolved in phosphate buffered saline (PBS). The membranes were then incubated with polyclonal rabbit anti-Asc1 (in 5% milk powder, 0.1% Tween 20, PBS) followed by incubation with a peroxidase-coupled secondary antibody or with Pierce™ High Sensitivity Streptavidin-HRP (Thermo Fisher, #21130, diluted 1:2000 in 1% BSA, 0.1% Tween 20, PBS). Chemiluminescent signals were detected with the FUSION-SL-4 imaging system (Peqlab).

*In E. coli.* For each sample, one colony of T7 Express competent cells (NEB), transformed with a pET-15b plasmid for the expression of microID, ultraID or CNOT9, was used to inoculate 5 mL LB medium containing 100 μg.mL⁻¹ ampicillin (LB-amp). After overnight incubation at 37°C, the samples were diluted to an OD600 of 0.6 in 6 mL LB-amp medium. Protein expression was then induced with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 1 h and then the cultures were split into two tubes with 3 mL culture volume. For each strain, biotin was added to one of those two tubes at a concentration of 50 μM while the other tube without biotin served as a negative control. The cultures were then incubated at 37°C for another
Thereafter, $OD_{600}$ was determined and volumes with cell numbers equivalent to 200 $\mu$L of $OD_{600} = 0.6$ were centrifuged at 4'000×g for 10 min. The cell pellets were resuspended in 50 $\mu$L 5× sodium dodecyl sulfate (SDS) loading buffer (5 % β-mercaptoethanol, 0.02 % (w/v) bromophenol blue, 10 % SDS, 30 % glycerol, 250 mM Tris-HCl pH 6.8), incubated at 95°C for 5 min and subsequently diluted to 1× with 200 $\mu$L water. 10 $\mu$L of these dilutions were used for western blot analysis.

**Western blot analysis**
For each sample, 10 $\mu$g of protein extract in 1× SDS loading buffer was loaded onto a discontinuous SDS polyacrylamide gel. After electrophoresis, proteins were transferred to a nitrocellulose membrane (Cytiva) by wet blotting at 4°C for 1h. Subsequent steps and fluorescence-based western blot quantitative analysis was performed as described in ref. 12 with the LI-COR Image Studio software.

**Preparative proximity-dependent biotinylation**
PDB in HeLa cell lines was performed with four biological replicates for each sample (ultraID-Ago2 and ultraID-Rab11, and BioID-Ago2 and BioID-Rab11). Cells were seeded in 15 cm dishes (2 per cell line) at 2.0 $\times$ 10^6 cells per dish. The following day fusion protein expression was induced with 25 ng/mL doxycycline-containing medium for 24 h. For the BioID cell lines, biotin was also added at that point at a concentration of 50 $\mu$M. For the ultraID cell lines, following the 24h induction, the medium was replaced with 50 $\mu$M biotin and 25 ng/mL doxycycline containing-medium and the cells were incubated at 37 °C for 10 min. PDB reactions were stopped by putting the plates on ice, removing the medium and washing three times with ice-cold PBS. All further steps were performed at 4°C. The cells were scraped in 1.5 mL PBS and collected in 15 mL tubes by centrifugation at 1′200×g for 5 min at 4°C. Cell pellets were frozen in liquid nitrogen and stored at −80 °C until further processing.

The procedure was similar for the P19 stable cell lines except that PDB was performed with three biological replicates for each sample (ultraID-Ago2, γ1-COP-ultraID, and γ2-COP-ultraID, with or without BFA addition) and fusion protein expression was induced with 5 ng/ml doxycycline for 24h. After induction, the medium of the BFA-treated samples was exchanged for 5μg/mL brefeldin A-containing medium and the cells incubated at 37°C for 1h. Thereafter, the medium was exchanged for 5μg/mL brefeldin A and 50 $\mu$M biotin-containing medium, and the cells incubated for another 4h. The mock-treated cells were incubated with 50 $\mu$M biotin-containing medium for 4h. PDB was then stopped and sample storage was performed as for the HeLa cells.

**Streptavidin affinity pulldowns and sample preparation for MS analysis**
Cell pellets were thawed on ice and resuspended in 1 mL ice-cold RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1 % SDS, 0.5 % sodium deoxycholate, 1 % Triton X-100, 1x Protease Inhibitor Cocktail, 1 mM DTT). All subsequent steps were performed at 4°C. 1 $\mu$L benzonase (Sigma) was added and the cell suspensions were transferred to RNase-free microcentrifuge tubes. Lysates were sonicated with a Bioruptor (Diagenode) at high intensity with 4 cycles of 30 s ON and 30 s OFF. The sonicated lysates were centrifuged at 16'000×g for 10 min. The cleared lysates were then transferred to a fresh microcentrifuge tube and 30μL were stored as input material. Protein concentration was determined by a Bradford assay and all samples were adjusted to 3 mg protein in 1 mL RIPA buffer. High performance streptavidin
sepharose beads (Cytiva, #17-5113-01, Lot 10280314) were chemically modified to yield resistance against tryptic digestion as described in ref. 17. 40 μL modified streptavidin-sepharose beads equilibrated in RIPA buffer were used per pulldown. The protein concentration-adjusted lysate samples were incubated with the equilibrated beads at 4 °C for at least 1h. The beads were then harvested by centrifugation and washed with four consecutive buffers. For each wash, 1mL of buffer was incubated with the beads at RT for 5min, the beads were then harvested by centrifugation. The washing procedure included two washes with wash buffers 1, 2 and 3 each and three washes with wash buffer 4 (wash buffer 1: 2 % SDS; wash buffer 2: 0.1 % sodium deoxycholate, 1 % Triton X-100, 500 mM NaCl, 1 mM EDTA, 50 mM HEPES pH 7.4; wash buffer 3: 10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5 % NP-40, 0.1 % sodium deoxycholate; wash buffer 4: 50 mM ammonium bicarbonate). The buffer from the final wash was carefully removed and 1 μg of mass spectrometry grade trypsin (Serva, Heidelberg) in 60 μL wash buffer 4 was added to each sample. On-beads digestion was performed at 37°C overnight. On the following day, another 0.5 μg of trypsin were added to each sample and the digestion was resumed at 37°C for another 2h. The beads were centrifuged and the supernatant was carefully transferred to a fresh RNase-free microcentrifuge tube. The beads were washed with 30 μL of high-performance liquid chromatography (HPLC) grade water at RT for 5 min. After centrifugation, the supernatant was mixed with the previously collected supernatant and the elution procedure was repeated once. The eluates were centrifuged at 16’000× g for 5 min and 100 μL of the supernatant was transferred to a fresh microcentrifuge tube. The samples were acidified with 4 μL of 50 % formic acid and dried in a SpeedVac centrifugal evaporator (ThermoFisher) at 65°C for 2 h. The dried peptides were stored at −80°C until analysis by mass spectrometry.

Mass spectrometry data acquisition
Dried peptide samples were reconstituted in 20 μL LC-MS sample buffer (2% acetonitrile, 0.1% formic acid). 2 μL of each sample were subjected to reverse phase liquid chromatography for peptide separation using an RSLCnano Ultimate 3000 system (Thermo Fisher Scientific). Therefore, peptides were loaded on an Acclaim PepMap 100 pre-column (100 μm x 2 cm, C18, 5 μm, 100 Å; Thermo Fisher Scientific) with 0.07% trifluoroacetic acid at a flow rate of 20 μL/min for 3 min. Analytical separation of peptides was done on an Acclaim PepMap RSLC column (75 μm x 50 cm, C18, 2 μm, 100 Å; Thermo Fisher Scientific) at a flow rate of 300 nL/min. The solvent composition was gradually changed within 94 min from 96 % solvent A (0.1 % formic acid) and 4 % solvent B (80 % acetonitrile, 0.1 % formic acid) to 10 % solvent B within 2 minutes, to 30 % solvent B within the next 58 min, to 45% solvent B within the following 22 min, and to 90 % solvent B within the last 12 min of the gradient. All solvents and acids had Optima grade for LC-MS (Thermo Fisher Scientific). Eluting peptides were on-line ionized by nano-electrospray (nESI) using the Nanospray Flex Ion Source (Thermo Fisher Scientific) at 1.5 kV (liquid junction) and transferred into a Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Full scans in a mass range of 300 to 1650 m/z were recorded at a resolution of 30,000 followed by data-dependent top 10 HCD fragmentation at a resolution of 15,000 (dynamic exclusion enabled). LC-MS method programming and data acquisition was performed with the XCalibur 4.0 software (Thermo Fisher Scientific).
**Mass spectrometry data analysis**

The raw files from the MS analysis were analyzed with the MaxQuant software\(^{18}\) for peptide and protein identification and assignment of LFQ intensities. The standard pre-assigned MaxQuant parameters were used with no activation of the “match between run” option (only peptides detected by MS/MS were considered). For enrichment analysis, the Maxquant data were analyzed in R using the packages proDA\(^{20}\), EnhancedVolcano\(^{63}\) and VennDiagram\(^{64}\). Few obvious hits were missed by the enrichment analysis by proDA and were added to the list of significantly enriched proteins. The criteria for enrichment for these hits were for the HeLa-11ht cell lines: detected in all four Ago2 samples with 2 unique peptides in at least three replicates, and no peptide detected in any Rab11 control sample. For the P19 cells: detected in all three γ1/2-COP samples with 2 unique peptides in all three replicates or 3 unique peptides in at least two replicates, and no peptide detected in any Ago2 control sample.

**Library generation and yeast cultivation**

For randomization of microID, error-prone PCR was performed using the GeneMorph II Random Mutagenesis Kit (Agilent). Hereby, three individual PCRs with different mutation rates ranging from low to high were carried out. PCR products were purified using the Promega Wizard SV Gel and PCR Clean-Up System and amplified by conventional PCR using primers that incorporated 30 bp sequences homologous to the pCT display vector\(^{65}\). A yeast surface-display library was generated as described in\(^{14, 66}\). Gap repair cloning was performed into a BamHI, Xhol and Nhel hydrolyzed pCT vector. Library size was estimated by plating a serial dilution of the regenerated yeasts on SD-Trp agar. The resulting library was cultivated in 1 L SD-Trp medium at 30°C and 180 rpm. The next day, surface presentation was induced by inoculating the library in 50 mL SG-Trp medium to an optical density of 1 and incubation at 30°C and 180 rpm overnight.

**Yeast-surface biotinylation assay and staining**

1×10^7 - 1×10^8 yeasts were washed twice and resuspended in 1 mL PBS. Biotin and ATP were added to a final concentration of 50 μM and 2.5 mM, respectively. Cells were incubated at 30°C and 900 rpm to prevent settling. Incubation time was decreased for each screening round. To prepare yeast cells for FACS screening, biotinylated cells were washed twice with PBS, resuspended in 100 μL PBS containing 0.1 μg anti-penta-His antibody (Qiagen) and 1 μg streptavidin allophycocyanin (Invitrogen) and incubated on ice for 20 min. After two washing steps, cells were stained as described before using 0.2 μg anti-mouse FITC antibody (Sigma-Aldrich). Finally, cells were washed again twice and were suspended in 1.5 mL of PBS.

**FACS screening**

FACS screening was performed after cell staining. Sorting gates were set in comparison to a microID negative control (no biotin and ATP). Cells exhibiting both, a fluorescence signal for biotin ligase activity (670 nm) and surface presentation (530 nm), were sorted using a BD Influx system and cultivated on SD-agar plates. After two days, cells were scraped off the plates, inoculated in SD liquid medium and cultivated overnight. After induction of surface presentation, biotinylation and staining as described above, a next round of screening was performed.
Recombinant expression
microID, ultraID-4 and ultraID-5 genes were amplified by PCR and subcloned into a pET22b expression vector, which was used for transformation of E. coli BL21 (DE3) cells. R40G / L41P and L41P-only mutants were generated by site-directed mutagenesis using the pET22b_microID plasmid as a template. A 50mL dYT overnight culture was used to inoculate 1L TB-medium to an OD600 = 0.1 and was incubated at 37°C and 220 rpm. Enzyme expression was induced by addition of 0.5 mM IPTG (final concentration) after reaching an OD600 ~0.6. Expression was carried out overnight at 18°C and 220 rpm. After cell lysis by sonification, the enzymes were purified by IMAC, dialyzed against 50 mM TRIS-HCl pH 7.5, 150 mM NaCl and 5 mM MgCl2 overnight and stored at -80°C.

ELISA-based biotinylation assay
To quantify enzymatic activity, a 96-well Nunc MaxiSorp plate (ThermoFisher scientific) was coated with 200 µL of a 40 mg/mL human serum albumin fraction V (CarlRoth, Germany) for 1 h. Wells were washed three times with PBST (6.4 mM Na2PO4, 2 mM KH2PO4, 10 mM KCl, 140 mM NaCl, 0,1% (w/v) Tween20, pH 7.4). For biotinylation reactions, 50 µM Biotin, 2.5 mM ATP, and 0.06 mg/ml biotin ligase (final concentration, each) in 50 mM TRIS-HCl pH 7.5, 150 mM NaCl and 5 mM MgCl2 were added and incubated for 10 min at 37°C. Reaction was stopped by adding 0.25 mM EDTA. Subsequently, wells were washed three times with PBST. Alkaline phosphatase-conjugated streptavidin (Sigma-Aldrich) was diluted 1:500 and added in a volume of 50 µl for 1 h at RT. Staining reaction was carried out after three washing steps and equilibration of the wells with AP-buffer. 1 mg/ml para-nitrophenyl phosphate was added and absorbance at 405 nm was measured with a TECAN Infinite® 200 PRO plate reader after 20 min of staining.

Data availability
The MS proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository69 with the data set identifier XXXXXXXX. The data that support the findings of this study are available from the corresponding author on request.

References
1. Gingras, A.C., Abe, K.T. & Raught, B. Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. Curr Opin Chem Biol 48, 44-54 (2019).
2. Roux, K.J., Kim, D.I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol 196, 801-810 (2012).
3. Kim, D.I. et al. Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proc Natl Acad Sci U S A 111, E2453-2461 (2014).
4. Kim, D.I. et al. An improved smaller biotin ligase for BioID proximity labeling. Mol Biol Cell 27, 1188-1196 (2016).
5. Xu, Y. & Beckett, D. Evidence for interdomain interaction in the Escherichia coli repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. *Biochemistry* **35**, 1783-1792 (1996).

6. Groft, C.M., Uljon, S.N., Wang, R. & Werner, M.H. Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. *Proc Natl Acad Sci U S A* **95**, 9117-9122 (1998).

7. Samavarchi-Tehrani, P., Samson, R. & Gingras, A.C. Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches. *Mol Cell Proteomics* **19**, 757-773 (2020).

8. Ramanathan, M. et al. RNA-protein interaction detection in living cells. *Nat Methods* **15**, 207-212 (2018).

9. Branon, T.C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880-887 (2018).

10. May, D.G., Scott, K.L., Campos, A.R. & Roux, K.J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* **9** (2020).

11. Kido, K. et al. AirID, a novel proximity biotinylation enzyme, for analysis of protein-protein interactions. *eLife* **9** (2020).

12. Schopp, I.M. et al. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat Commun* **8**, 15690 (2017).

13. Banaszynski, L.A., Liu, C.W. & Wandless, T.J. Characterization of the FKBP.rapamycin.FRBox ternary complex. *Journal of the American Chemical Society* **127**, 4715-4721 (2005).

14. Bogen, J.P., Grzeschik, J., Krah, S., Zielonka, S. & Kolmar, H. Rapid Generation of Chicken Immune Libraries for Yeast Surface Display. *Methods Mol Biol* **2070**, 289-302 (2020).

15. Opitz, N. et al. Capturing the Asc1p/Receptor for Activated C Kinase 1 (RACK1) Microenvironment at the Head Region of the 40S Ribosome with Quantitative BioID in Yeast. *Mol Cell Proteomics* **16**, 2199-2218 (2017).

16. Béthune, J., Artus-Revel, C.G. & Filipowicz, W. Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep* **13**, 716-723 (2012).

17. Rafiee, M.R. et al. Protease-resistant streptavidin for interaction proteomics. *Molecular systems biology* **16**, e9370 (2020).

18. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).

19. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526 (2014).

20. Ahlmann-Eltze, C. & Anders, S. proDA: Probabilistic Dropout Analysis for Identifying Differentially Abundant Proteins in Label-Free Mass Spectrometry. *bioRxiv*, 661496 (2020).

21. Eulalio, A., Huntzinger, E. & Izaurralde, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat Struct Mol Biol* **15**, 346-353 (2008).
22. Maniataki, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* **19**, 2979-2990 (2005).
23. Braun, J.E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44**, 120-133 (2011).
24. Chekulaeva, M. et al. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat Struct Mol Biol* **18**, 1218-1226 (2011).
25. Orban, T.I. & Izaurralde, E. Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* **11**, 459-469 (2005).
26. Béthune, J. & Wieland, F.T. Assembly of COPI and COPII Vesicular Coat Proteins on Membranes. *Annu Rev Biophys* **47**, 63-83 (2018).
27. Gilchrist, A. et al. Quantitative proteomics analysis of the secretory pathway. *Cell* **127**, 1265-1281 (2006).
28. Adolf, F. et al. Proteomic Profiling of Mammalian COPII and COPI Vesicles. *Cell reports* **26**, 250-265 e255 (2019).
29. Jain Goyal, M. et al. A paralog-specific role of COPI vesicles in the neuronal differentiation of mouse pluripotent cells. *Life Sci Alliance* **3** (2020).
30. Moelleken, J. et al. Differential localization of coatomer complex isoforms within the Golgi apparatus. *Proc Natl Acad Sci U S A* **104**, 4425-4430 (2007).
31. Donaldson, J.G., Finazzi, D. & Klausner, R.D. Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. *Nature* **360**, 350-352 (1992).
32. Helms, J.B. & Rothman, J.E. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature* **360**, 352-354 (1992).
33. Sciaky, N. et al. Golgi tubule traffic and the effects of brefeldin A visualized in living cells. *J Cell Biol* **139**, 1137-1155 (1997).
34. Liu, J., Valencia-Sanchez, M.A., Hannon, G.J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7**, 719-723 (2005).
35. Turgay, Y. et al. A classical NLS and the SUN domain contribute to the targeting of SUN2 to the inner nuclear membrane. *EMBO J* **29**, 2262-2275 (2010).
36. Buchwalter, A., Schulte, R., Tsai, H., Capitanio, J. & Hetzer, M. Selective clearance of the inner nuclear membrane protein emerin by vesicular transport during ER stress. *eLife* **8** (2019).
37. Lowe, M. The Physiological Functions of the Golgin Vesicle Tethering Proteins. *Front Cell Dev Biol* **7**, 94 (2019).
38. Witkos, T.M. & Lowe, M. Recognition and tethering of transport vesicles at the Golgi apparatus. *Curr Opin Cell Biol* **47**, 16-23 (2017).
39. Linders, P.T., Horst, C.V., Beest, M.T. & van den Bogaart, G. Stx5-Mediated ER-Golgi Transport in Mammals and Yeast. *Cells* **8** (2019).
40. Laufman, O., Kedan, A., Hong, W. & Lev, S. Direct interaction between the COG complex and the SM protein, Sly1, is required for Golgi SNARE pairing. *EMBO J* **28**, 2006-2017 (2009).
41. Sonnichsen, B. et al. A role for giantin in docking COPI vesicles to Golgi membranes. *J Cell Biol* **140**, 1013-1021 (1998).
42. Miller, V.J. et al. Molecular insights into vesicle tethering at the Golgi by the conserved oligomeric Golgi (COG) complex and the golgin TATA element modulatory factor (TMF). *J Biol Chem* **288**, 4229-4240 (2013).

43. Sohda, M. et al. Interaction of Golgin-84 with the COG complex mediates the intra-Golgi retrograde transport. *Traffic* **11**, 1552-1566 (2010).

44. Gillingham, A.K. At the ends of their tethers! How coiled-coil proteins capture vesicles at the Golgi. *Biochem Soc Trans* **46**, 43-50 (2018).

45. Wong, M. & Munro, S. Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. *Science* **346**, 1256898 (2014).

46. Xu, P. et al. COPI mediates recycling of an exocytic SNARE by recognition of a ubiquitin sorting signal. *eLife* **6** (2017).

47. Aniento, F., Gu, F., Parton, R.G. & Gruenberg, J. An endosomal beta COP is involved in the pH-dependent formation of transport vesicles destined for late endosomes. *J Cell Biol* **133**, 29-41 (1996).

48. Whitney, J.A., Gomez, M., Sheff, D., Kreis, T.E. & Mellman, I. Cytoplasmic coat proteins involved in endosome function. *Cell* **83**, 703-713 (1995).

49. Witkos, T.M. et al. GORAB scaffolds COPI at the trans-Golgi for efficient enzyme recycling and correct protein glycosylation. *Nat Commun* **10**, 127 (2019).

50. Kuliyev, E. et al. Overlapping Role of SCYL1 and SCYL3 in Maintaining Motor Neuron Viability. *J Neurosci* **38**, 2615-2630 (2018).

51. Hamlin, J.N. et al. Scyl1 scaffolds class II Arfs to specific subcomplexes of coatomer through the gamma-COP appendage domain. *J Cell Sci* **127**, 1454-1463 (2014).

52. Ludes-Meyers, J.H., Bednarek, A.K., Popescu, N.C., Bedford, M. & Aldaz, C.M. WWOX, the common chromosomal fragile site, FRA16D, cancer gene. *Cytogenet and genome research* **100**, 101-110 (2003).

53. Mahmud, M.A.A. et al. Cellular Expression and Subcellular Localization of Wwox Protein During Testicular Development and Spermatogenesis in Rats. *J Histochem Cytochem*, 22155421991629 (2021).

54. Hartwig, C. et al. Trafficking mechanisms of P-type ATPase copper transporters. *Curr Opin Cell Biol* **59**, 24-33 (2019).

55. Hauri, H.P., Kappeler, F., Andersson, H. & Appenzeller, C. ERGIC-53 and traffic in the secretory pathway. *J Cell Sci* **113** ( Pt 4), 587-596 (2000).

56. Vandemoortele, G., Eyckerman, S. & Gevaert, K. Pick a Tag and Explore the Functions of Your Pet Protein. *Trends Biotechnol* **37**, 1078-1090 (2019).

57. Stanley, P. Golgi glycosylation. *Cold Spring Harbor perspectives in biology* **3** (2011).

58. Sahlmuller, M.C. et al. Recombinant heptameric coatomer complexes: novel tools to study isoform-specific functions. *Traffic* **12**, 682-692 (2011).

59. Weidenfeld, I. et al. Inducible expression of coding and inhibitory RNAs from retargetable genomic loci. *Nucleic Acids Res* **37**, e50 (2009).

60. Schopp, I.M. & Béthune, J. Split-BioID - Proteomic Analysis of Context-specific Protein Complexes in Their Native Cellular Environment. *J Vis Exp* (2018).

61. Ramirez, C.A., Egetemaier, S. & Béthune, J. Context-Specific and Proximity-Dependent Labeling for the Proteomic Analysis of Spatiotemporally Defined Protein Complexes with Split-BioID. *Methods Mol Biol* **2247**, 303-318 (2021).

62. Kushnirov, V.V. Rapid and reliable protein extraction from yeast. *Yeast* **16**, 857-860 (2000).
63. Blighe, K., Rana, S. & Lewis, M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. (online, 2018).

64. Chen, H. & Boutros, P.C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 12, 35 (2011).

65. Boder, E.T. & Wittrup, K.D. Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15, 553-557 (1997).

66. Benatuil, L., Perez, J.M., Belk, J. & Hsieh, C.M. An improved yeast transformation method for the generation of very large human antibody libraries. Protein Eng Des Sel 23, 155-159 (2010).

Acknowledgments
We thank Valeria Mery and Jérôme Oswald for helping establishing PDB in bacteria with ultraID, Hans-Dieter Schmitt (MPI for Biophysical Chemistry, Göttingen, Germany) for providing codon usage-optimised BioID2 for S. cerevisiae as a template for generating ASC1-microID and ASC1-ultraID plasmids. LC-MS analysis was done in the Service Unit LCMS Protein Analytics of the Göttingen Center for Molecular Biosciences (GZMB) at the Georg-August-University Göttingen (Grant DFG-GZ: INST 186/1230-1 FUGG to Stefanie Pöggeler).

Author contributions
X.Z. performed all experiments with P19 cells. S.B. and L.D. performed the directed evolution of microID. L.K. performed all experiments with HeLa cells and bacteria. A.R. performed all the preliminary experiments leading to the discovery and initial characterization of microID. E.C.B. performed preliminary experiments for the characterization of microID and ultraID. K.S. and O.V. contributed the yeast experiments and the LC-MS analysis. H.K. supervised the directed evolution approach. J.B. designed and supervised the study, and analyzed the data. L.K. and J.B. wrote the first draft of the manuscript. All authors edited the paper.