An automated proximity proteomics pipeline for subcellular proteome and protein interaction mapping

Xiaofang Zhong[1,2,3,†], Qiongyu Li[1,2,3,†], Benjamin J. Polacco[1,2,3], Trupti Patil[1,2,3], Jeffrey F. DiBerto[4], Rasika Vartak[1,2,3], Jiewei Xu[1,2,3], Aaron Marley[5], Helene Foussard[1,2,3], Bryan L. Roth[4], Manon Eckhardt[1,2,3], Mark Von Zastrow[1,3,5], Nevan J. Krogan[1,2,3], Ruth Hüttenhain[1,2,3]

[1]Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco, CA 94158, USA.
[2]J. David Gladstone Institutes, San Francisco, CA 94158, USA.
[3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA.
[4]Department of Pharmacology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.
[5]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA 94158, USA.

[†]Authors contributed equally

Abstract

Proximity labeling (PL) coupled with mass spectrometry has emerged as a powerful technique to map proximal protein interactions in living cells. Large-scale sample processing for proximity proteomics necessitates a high-throughput workflow to reduce hands-on time and increase quantitative reproducibility. To address this issue, we developed a scalable and automated PL pipeline, including generation and characterization of monoclonal cell lines, automated enrichment of biotinylated proteins in a 96-well format, and optimization of the quantitative mass spectrometry (MS) acquisition method. Combined with data-independent acquisition (DIA) MS, our pipeline outperforms manual enrichment and data-dependent acquisition (DDA) MS regarding reproducibility of protein identification and quantification. We apply the pipeline to map subcellular proteomes for endosomes, late endosomes/lysosomes, the Golgi apparatus, and the plasma membrane. Moreover, using serotonin receptor ($5HT_{2A}$) as a model, we investigated agonist-induced dynamics in protein-protein interactions. Importantly, the approach presented here is universally applicable for PL proteomics using all biotinylation-based PL enzymes, increasing both throughput and reproducibility of standard protocols.

Introduction

The subcellular localization of individual proteins and their interactions with other proteins determines protein function. Neither protein localization nor protein-protein interactions (PPIs) are static: proteins traffic between subcellular neighborhoods and dynamically change interactions to carry out their functions and respond to extracellular signals. Systematic characterization of protein localization and changing protein networks is thus essential to delineate molecular mechanisms of fundamental cellular processes.

Proximity labeling (PL) has emerged as a powerful and versatile proteomics approach to globally map the spatial organization of proteins (Rhee *et al*, 2013; Go *et al*, 2021) and their molecular interactions, including protein-protein (Go *et al*, 2021; Lobingier *et al*, 2017; Paek *et al*, 2017), protein-RNA (Ramanathan *et al*, 2018; Qin *et al*, 2021b), and protein-DNA interactions (Gao *et al*, 2018; Myers *et al*, 2018). PL relies on enzymes that are genetically fused to either a target protein or a protein localization domain of interest and subsequently catalyze the proximity-dependent biotinylation of biomolecules in live cells. Various PL enzymes have been developed representing two main enzyme classes: (1) peroxidases, including horseradish peroxidase (Li *et al*, 2014; Martell *et al*, 2016) and engineered ascorbate peroxidase 2 (APEX2) (Lam *et al*, 2015) and (2) biotin ligases including BioID (Roux *et al*, 2012), UltraID (Kubitz *et al*, 2022), TurboID and miniTurbo (Branon *et al*, 2018) and others (Samavarchi-Tehrani *et al*, 2020; Qin *et al*, 2021a).

The common principle underlying all of these PL approaches involves the expression of the PL enzyme fusion construct in cells. To avoid artifacts from overexpression and variability in the expression level of the fusion construct, most studies aim for moderate and inducible expression of the fusion construct in isogenic cell clones (Samavarchi-Tehrani *et al*, 2018), which requires characterization of multiple cellular clones for each PL construct. Once a cellular clone is selected, the biotinylation reaction is performed for different conditions to label proteins in the proximity of the protein of interest. These biotinylated proteins are then enriched for each sample by binding to streptavidin beads, followed by a series of washing steps to remove non-biotinylated proteins and proteolytic digestion to prepare samples for subsequent quantitative mass spectrometric (MS) analysis. This multistep sample preparation is time-consuming and has restricted scalability, limiting the sample throughput in current PL proteomics applications. Additionally, sample preparation across multiple batches can increase variability, thus compromising reproducibility in quantifying the biotinylated proteins across batches. While PL represents a powerful technology to investigate dynamics in subcellular localizations as well as protein interactions across many conditions, throughput has been a limiting factor for large-scale PL projects due to the multistep nature of the standard protocol (Go *et al*, 2021; Antonicka *et al*, 2020; Göös *et al*, 2022; Piette *et al*, 2021).

To address these issues, here we describe an automated, scalable, and reproducible pipeline for PL-based proteomics (**Figure 1**) including: generation and characterization of monoclonal

cell lines expressing the PL-fusion construct; automated enrichment of biotinylated proteins in 96-well plate format; and an optimized quantitative proteomics method for reproducible protein quantification. To obtain consistent expression levels of APEX2-tagged constructs we streamlined the process of monoclonal cell line generation and characterization using a multistep process based on flow cytometry to determine the construct expression level, immunofluorescence imaging to verify cellular localization, and western blot analysis to assess biotinylation efficiency of the APEX2 enzyme. To enable PL studies with high-throughput, we built an automated 96-well format platform for enriching biotinylated proteins using magnetic streptavidin beads on the KingFisher Flex platform. Commonly, data-dependent acquisition (DDA) has been used for PL proteomics, but the preference for high-abundance peptides in DDA can result in stochastic and irreproducible measurements (Aebersold & Mann, 2003; Michalski *et al*, 2011; Hu *et al*, 2016). Due to the systematic and unbiased fragmentation of all precursor ions within a predefined mass range, data-independent acquisition (DIA) achieves higher quantitative consistency at comparable proteome coverage. (Gillet *et al*, 2012; Collins *et al*, 2017; Ludwig *et al*, 2018). Thus, we established a DIA-MS method for biotinylated proteins based on several optimal parameters, including precursor scan range, normalized collision energy, maximum injection time, and the precursor isolation window overlap. The automated enrichment strategy in combination with the optimized DIA-MS method did not only increase throughput, but also provided a higher reproducibility of protein identification and quantification compared to manual sample processing.

We demonstrate the ability of our automated PL strategy in combination with DIA-MS to accelerate the quantitative analysis of subcellular proteomes as well as protein interaction network dynamics across a time course. We used APEX2-based PL to determine compartment-specific proteins for several subcellular compartments including plasma membrane, endosome, Golgi apparatus and lysosome. The compartment-specific proteins determined with our approach matched their expected locations based on the Human Cell Map (Go *et al*, 2021). We further applied our automated PL strategy to map the dynamics of proximal proteome changes of a target protein of interest. On the example of the serotonin 2A ($5HT_{2A}$) receptor, a G protein-coupled receptor (GPCR), activated with its endogenous ligand serotonin (5-HT), we demonstrate that our automated PL approach combined with DIA-MS provides high reproducibility to quantify ligand-induced proximal proteome changes over a time course. Among the ligand-dependent interaction network changes, we identified well known regulators of GPCR function, such as G protein-coupled receptor kinase 2 (GRK2), beta-Arrestin 2, as well as subunits of protein kinase C (PKC) and protein kinase D (PKD).

Overall, we demonstrate the applicability of our automated PL pipeline for a wide range of PL applications. While we focus on APEX2-based proximity labeling in our study, the automated strategy can be widely used to essentially accelerate any PL strategy which relies on biotinylation.

Results

Step-by-step guide for generating monoclonal cell lines for proximity labeling

To set up a PL proteomics experiment, the PL enzyme of choice is genetically fused to a protein localization domain (PLD) or a protein of interest (POI) and expressed in cells. Here, we provide a step-by-step guide for designing PL constructs as well as generating and characterizing monoclonal cell lines for PL proteomics. To allow for stable, inducible expression of the PL constructs in a broad range of cell types, we introduced the APEX2 constructs into a lentiviral backbone under a Doxycycline-inducible promoter containing Puromycin resistance for antibiotic selection (**Figure 2A**). The constructs were appended with a FLAG tag and/or GFP for detection by FACS, imaging, and western blot. For spatial proteomics, a protein localization sequence was selected to direct the PL construct to the subcellular compartment of interest. For protein interaction mapping, the APEX2 enzyme was fused to the protein of interest (POI) and separated by a flexible linker to minimize interference of the PL enzyme with the protein's function (**Figure 2A**). After generating polyclonal cells by lentiviral transduction and puromycin selection, the expression of PL constructs was confirmed through FACS and immunofluorescence microscopy. Finally, to achieve consistent and homogeneous expression of the APEX2-tagged constructs and to reduce shifts in the cellular population between biological replicates, we generated monoclonal cell lines by single cell sorting. To choose optimal clones after sorting, we designed a strategy for characterization of these clones by initially screening multiple clones using flow cytometry to evaluate expression uniformity of the fusion construct. Then the most promising ones were assessed by immunofluorescence and confocal imaging for subcellular localization of the fusion construct, followed by western blot analysis to evaluate the capacity to biotinylate (**Figure 2A**).

We exemplify the monoclonal cell line generation and characterization in HEK293T cells for a plasma membrane localized APEX2 construct, for which the enzyme was targeted to the plasma membrane (PM) using a Lyn11 motif appended with GFP for visualization (**Figure S1**). Following lentiviral transduction and puromycin selection, the polyclonal cell line expressed the PL construct heterogeneously, with some cells expressing GFP very faintly and others showing GFP expression in cytoplasm and nucleus in addition to the PM (**Figure 2B**). Following single cell sorting, 24 monoclonal PM-APEX2 cell lines were selected for flow cytometry analysis to test for expression consistency and intensity of the APEX2-tagged PM construct based on GFP fluorescence. Based on the flow cytometry analysis, we prioritized monoclonal cell lines containing over 85% GFP-positive cells for further characterization by immunofluorescence and confocal microscopy (**Figure 2C**). Confocal imaging showed localization of PM-APEX2 clones 7, 8, and 11 concentrated at the PM, while clone 1 is mainly diffused in the cytoplasm (**Figure 2D**), an observation also recapitulated in the wider fluorescence distribution seen by flow cytometry for this clone. Finally, western blot analysis was performed to evaluate the biotinylation efficiency of the monoclonal PM-APEX2 cell lines. Biotinylation was initiated by addition of biotin-phenol and H_2O_2, with the sole biotin-phenol treatment as the negative control. All PM-APEX2 clones demonstrated different levels of biotinylation labeling, with the C7 and the polyclonal cell line showing highest intensity, C8 and C11 medium intensity, and C1 lowest intensity. Meanwhile negative controls of all the clones showed low levels of biotinylation

(**Figure 2E**). We reasoned that the high biotinylation efficiency might result in higher unspecific labeling, and selected clones 8 and 11 for PL studies.

Using this step-be-step pipeline, we could generate and characterize monoclonal cell lines within 3 weeks. While we chose APEX2-based PL for our study, the guide is applicable to any other PL enzyme.

An automated, scalable biotin enrichment method coupled to quantitative MS

The enrichment of biotinylated proteins is a key component of every proximity labeling workflow. Manual enrichment including several washing steps is not only time-consuming but can also increase sample to sample variability, thereby limiting quantitative reproducibility. To facilitate high-throughput analysis in large-scale PL studies while maintaining high reproducibility and quantitative accuracy, we established an optimal PL proteomics strategy combining automated enrichment of biotinylated proteins in 96-well plate format with MS analysis using data-independent acquisition (DIA) (**Figure 3A**).

A two-day automated enrichment protocol using magnetic streptavidin beads for biotinylated proteins was adapted on the KingFisher Flex system (**Figure 3B**), for which the biotinylated proteins are bound to beads overnight, followed by various washing steps to remove non-biotinylated beads on the second day. For establishing the protocol, we considered parameters including incubation time for protein binding to beads, bead and sample mixing speed during binding and washing, duration of the washing steps, and bead collection time. The protocol, which is delineated in detail in **Table S2**, requires less than 5 minutes to load the corresponding plates into the system (**Figure 3B**) and eliminates the need for manual pipetting. To maximize the enrichment efficiency while minimizing unspecific binding to beads, we evaluated the effects of varying amounts of protein lysate input and streptavidin beads. Therefore, we performed a proximity labeling experiment using the PM-APEX2 cell line described above. Cells were incubated with biotin phenol, followed by treatment with H_2O_2 to initiate the biotin labeling reaction. Based on the number of proteins being identified in the PM-APEX2 biotinylation experiment, 1 mg of protein lysate (**Figure S2A**) and 80 µl of magnetic streptavidin beads were determined as the optimal starting amount for enrichment (**Figure S2B**). While 80 µL of streptavidin beads resulted in slightly lower protein identifications compared to 100 µL, the number of identified proteins in the negative control (without H_2O_2 treatment) was reduced using 80 µL of beads, indicating less unspecific binding to beads (**Figure S2B**). Notably, both the optimal protein lysate starting amount and streptavidin bead amount, might be dependent on the expression level of the PL construct and the cell type used.

To maximize protein identifications and quantitative accuracy, we optimized several parameters for the MS analysis of proximity labeling samples using DIA, including precursor scan range, normalized collision energy, maximum injection time, and the overlap width of the precursor isolation window. Optimized parameters for the DIA method consisted of one full MS scan over a range of 350-1050 *m*/*z*, MS^2 scans with a normalized HCD (higher-energy collisional dissociation) collision energy of 30%, a maximum injection time of 32 ms, and isolation windows of 20 *m*/*z* over 350-1050 *m*/*z* range with 2 *m*/*z* overlap between windows (**Figure S2C-F**).

To evaluate protein identification and quantification capabilities of the optimized DIA method

coupled to the automated enrichment method (Auto_DIA), we compared it to a manual biotinylation enrichment coupled to DIA analysis (Manual_DIA) as well as a data-dependent analysis (DDA) method with the automated enrichment (Auto_DDA). While the numbers of precursors, peptides, and proteins identified by all the three methods were comparable (**Figure S3A**), the reproducibility of Auto_DIA outperformed manual enrichment (Manual_DIA) and DDA analysis (Auto_DDA) based on the following evaluations: (1) The numbers of precursor features identified across all the three replicates were 36,992 (95%) for Manual_DIA, 31,743 (78.7%) for Auto_DDA, and increased to 40,211 (97.1%) for Auto_DIA (**Figure 3C**, **Table S3**); (2) Auto_DIA resulted in a narrow distribution of coefficients of variation (CV) with median CV of 7.16%, while the distributions for both Manual_DIA and Auto_DDA were wider with median CV of 12.8% and 11%, respectively (**Figure 3D**); (3) The correlation coefficients comparing peptide intensities of the three replicates ranged from 0.91-0.948 for Manual_DIA, 0.917-0.95 for Auto_DDA, but improved to 0.941-0.967 for Auto_DIA (**Figure S3B**). Collectively, we established a pipeline consisting of an automated protocol for enrichment of biotinylated proteins coupled to an optimized DIA MS analysis with the capacity of high-throughput sample processing is 96-well plate format, significantly reduced manual processing time, and higher reproducibility among technical replicates in terms of protein identification and quantification (**Figure 3E**).

Mapping subcellular proteomes using the automated proximity proteomics strategy

PL is the premier method for spatial proteomics and mapping the proteome composition of subcellular structures. To evaluate the performance of our optimized proximity proteomics strategy on subcellular proteome mapping, we generated and characterized monoclonal HEK293T cell lines expressing APEX2 proximity labeling constructs targeted to the endosome, Golgi apparatus, lysosome, and plasma membrane using the step-by-step approach described above. The protein localization sequence was selected as follows: Lyn11 for plasma membrane, 2xFYVE for endosome, LAMTOR1 and LAMP1 for late endosome/lysosome, β-1,4-galactosyltransferase (GalT) for Golgi apparatus and nuclear export signal (NES) as control (**Figure S1 and Table S1**). The final selected clones showed homogenous expression of the APEX2 construct with over 92% GFP-positive cells, except for the NES-APEX2 construct (**Figure S4A**). Moreover, the APEX2 constructs localized to the corresponding subcellular compartments of interest, as assessed by co-localization analysis with markers including Vimentin for cytosol, E-cadherin for plasma membrane, Rab5 for early endosome, and Rab9 and Golgin-97 for late endosome/lysosome and Golgi, respectively (**Figure 4A**). Lastly, the final monoclonal cell lines were selected based on similar biotinylation efficiency with minimum background biotinylation, which varied among different clones, likely due to the different expression levels of APEX2-tagged localization domains (**Figure S4B, C**). Following biotinylation, proteins were subjected to automated enrichment, on-bead proteolytic digestion, and quantitative DIA-MS. All cell lines showed successful biotinylation and efficient enrichment as assessed by western blot analysis (**Figure S5**).

Principal component analysis (PCA) comparing the biotinylated proteins for each localization construct showed separation of each location, but clustering of their biological replicates (**Figure 4B**). To determine compartment-specific proteins, we performed a two-step filtering process. First, the protein abundances for each location construct were compared to the control (NES-

APEX2) using MSstats (version 4.4.0) (Choi *et al*, 2014) (**Figure 4C**). All proteins with a log_2 fold change >1 and adjusted p-value <0.05 were considered as candidate compartment-specific proteins (**Table S4**). Second, we required the significant proteins to be exclusively present in one comparison, except for the proximal proteomes of endosome (2xFYVE) and lysosome/late endosome (LAMP1, and LAMTOR1) as we expect a significant overlap in their proteomes. In the end, 337, 55, and 228 location-specific proteins were identified for the endosome and lysosome/late endosome (Endo-Lyso), Golgi apparatus, and plasma membrane, respectively (**Figure 4D**). Based on the median normalized log_2FC (fold change) across all samples, the location-specific proteins were divided into three groups (**Figure 4E**). Gene ontology enrichment analysis demonstrated that, as expected, proteins specific for 2xFYVE, LAMP1, and/or LAMTOR1 (Group 1) were enriched for early endosome, endosome, and lysosome. Group 2 consisting of proteins specific for GalT were enriched for Golgi apparatus subcompartment and Golgi membrane and proteins in group 3, which are specific for Lyn11 (PM), were enriched in actin filament-based process, anchoring junction, and plasma membrane region. Finally, we compared the location-specific proteins derived from our PL analysis to the proteins defined for the respective compartments in the Human Cell Map (Go *et al*, 2021), a proximity map of the HEK293 proteome generated using PL using BioID. As expected, proteins specific for 2xFYVE, LAMP1, and LAMTOR1 were enriched for endosome and lysosome proteome determined by the Human Cell Map, proteins specific for GalT were enriched in Golgi apparatus, and proteins specific for Lyn11 (PM) were enriched in plasma membrane and cell junction (**Figure 4F**). Notably, our dataset also recovered location specific proteins which were not identified in the Human Cell Atlas, including Golgin-45 (Short *et al*, 2001), sphingomyelin synthases 2 (Villani *et al*, 2008), and palmitoyltransferase ZDHHC14 (Adachi *et al*, 2016) for Golgi apparatus; cell surface hyaluronidase (Yamamoto *et al*, 2017), pleckstrin homology domain-containing family O member 1 (Olsten *et al*, 2004), and guanine nucleotide-binding protein G(i) subunit alpha-1 (Cho & Kehrl, 2007) for plasma membrane; and proteins involved in endosomal trafficking, like Ras-related protein Rab-4A (RAB4A), Charged multivesicular body protein 2a (CHMP2A), sorting nexin-11 (SNX11) and sorting nexin-16 (SNX16) and vacuolar protein sorting-associated protein 41 homolog (VPS41) for endosome/lysosome. Taken together, using our automated PL strategy we identified location-specific proteins in various cellular compartments, demonstrating its applicability to spatial proteomics.

Mapping ligand-dependent proximal interaction network changes for the $5HT_{2A}$ receptor

Given the quantitative reproducibility of our automated proximity proteomics strategy coupled to DIA-MS, we sought to evaluate its applicability to study protein interaction network dynamics. Therefore, we investigated the changes in proximal interaction network of the serotonin 2A receptor ($5HT_{2A}$) upon activation by its endogenous ligand, 5-hydroxytryptamine (5HT). Despite the therapeutic relevance of $5HT_{2A}$ (Kwan *et al*, 2022; McClure-Begley & Roth, 2022), little is known about the interaction networks that regulate $5HT_{2A}$ cellular responses. Mapping its proximal interaction networks and how these dynamically change upon receptor activation could reveal novel factors regulating the activity of the receptor which, upon further validation, might be used to finetune $5HT_{2A}$ signaling responses.

To generate the $5HT_{2A}$ PL construct, the APEX2 enzyme was placed within the C-terminal tail of the receptor to prevent disturbing the C-terminal PSD-95/Discs-large/ZO-1 (PDZ)-binding domain, which is essential for receptor-mediated signaling and trafficking (Xia *et al*, 2003) (**Figure S6A**). Following monoclonal HEK293T cell line generation, we validated that the $5HT_{2A}$-APEX2 construct localized to the plasma membrane (**Figure S6B**) and retained signaling activity determined based on Gq heterotrimer dissociation (**Figure S6C**). To perform PL, the receptor was activated with 5HT over a time course of 30 min, biotin phenol and H_2O_2 were added to initiate biotinylation, followed by quenching, cell lysis, and automated enrichment of biotinylated proteins (**Figure 5A**). Our optimized DIA-MS method quantified 6429 proteins across the whole dataset. To determine proteins with agonist-dependent labeling changes, we performed an analysis of variance (ANOVA) by fitting a polynomial curve for each protein across the time course (**Figure 5A**) thereby identifying 744 proteins with significant differences following 5HT treatment (adjusted p-value <0.05) (**Figure 5B, Table S5**). An overrepresentation analysis of the location-specific proteins for plasma membrane, endosome, late endosome/lysosome (**Figure 4D**) suggested a decrease in biotinylation of plasma membrane markers and increase of endosome and lysosome markers across the time-course, indicating ligand-induced trafficking of the receptor (**Figure 5B**). To distinguish between (1) location-specific proteins that reside in the local environment of the receptor but do not participate in its function and (2) protein networks that may regulate receptor signaling and trafficking in the PL dataset, we examined the fold change (FC) distribution of the location specific proteins and further filtered the significant proteins with FC > 1.5 (**Figure 5C**). Based on the temporal profiles the remaining 172 proteins formed two distinct clusters: A cluster representing proteins with rapid kinetics and transient change in biotin labeling at 1 min after agonist treatment and the second cluster with later, sustained difference in labeling across 30 min of 5HT treatment (**Figure 5D**).

The transient cluster was enriched for proteins functioning in GPCR signaling and Rho GTPases signaling (**Figure 5E**), including well-characterized interactors of $5HT_{2A}$ such as G protein-coupled receptor kinase 2 (GRK2), which phosphorylates $5HT_{2A}$, and β-arrestin 2 (ARRB2), which binds to the phosphorylated receptor to meditate receptor desensitization and internalization. We also observed increased transient labeling of downstream $5HT_{2A}$ signaling targets such as protein kinase C (PKCD, PKCA, and PKCQ), protein kinase D (PRKD2), and diacylglycerol kinase zeta (DGKZ), which have long been implicated in $5HT_{2A}$ signaling (Roth & Chuang, 1987; Allen *et al*, 2008). Moreover, the activation of $5HT_{2A}$ triggered the signaling by Rho GTPases, which play a key role in the regulation of actin dynamics through activation of WASPs (Wiskott-Aldrich syndrome proteins) and WAVEs (WASP family verprolin-homologous). Our dataset indicates four proteins of WASP/WAVE complex with 5HT-dependent increase in PL. While the transient cluster contains mainly proteins involved in receptor signaling, the sustained cluster shows, amongst others, enrichment of proteins involved in membrane trafficking (**Figure S6D**). Specifically, we identified regulators of clathrin-mediated endocytosis including BMP-2 inducible kinase (BMP2K), the mannose 6-phosphate receptor (M6PR), as well as the clathrin adaptor proteins Epsin-1 (EPN1) and Epsin-2 (EPN2) in our dataset. Prior studies have demonstrated the involvement of clathrin components for $5HT_{2A}$ mediated endocytosis (Bhatnagar *et al*, 2001). As exemplified by the $5HT_{2A}$ receptor study, these results

demonstrate that our automated PL strategy can capture rapid and sustained dynamics of proximal protein networks. We envision that the automated enrichment in 96-well plate format combined with DIA-based MS accelerates PL studies with many conditions in parallel and given the quantitative reproducibility allows for sensitive detection of PL dynamics.

Discussion

Here, we developed an automated, scalable, and reproducible approach for PL-based proteomics to accelerate large scale PL studies (**Figure 1**). We first provide a step-by-step guide for the generation and characterization of monoclonal cell lines for PL (**Figure 2**), which is critical for reproducible proximity labeling levels across conditions. We then implemented an automated 96-well format method on the KingFisher Flex instrument to enrich biotinylated proteins, which reduces variability resulting from manual sample processing (**Figure 3**). Instead of 3 hours for handling up to 16 samples using the manual enrichment, our automated approach allows processing 96 samples in parallel with only 30 min of manual work. For the automated protocol, we optimized several parameters, such as the amounts of protein lysate input and streptavidin beads. While these parameters might be dependent on the PL enzyme used, the expression level of the PL construct as well as the cell type, our study provides a guide for optimization of these parameters. Finally, in addition to improving reproducibility in sample preparation, we implemented a DIA-based MS method for the measurement of biotinylated proteins with higher reproducibility and consistency in protein identification and quantification (**Figure 3**). While the manual enrichment combined with DIA-based protein quantification already improved the reproducibility in protein identification from 78.7% to 95% overlap among three replicates, the addition of the automated enrichment not only increased the throughput, but also reduced variability in protein quantification from 12.8 to 7.1% median CV. Notably, in the future, the presented strategy could be further optimized by shortening the time for binding of the biotinylated proteins to streptavidin beads. Moreover, the feasibility for large scale applications could be facilitated by scaling down the protocol for lower protein input amount.

We demonstrated the applicability of our automated proximity proteomics approach to map subcellular proteomes by targeting the APEX2 PL enzyme to early and late endosomes, Golgi apparatus, lysosome, and plasma membrane (**Figure 4**). The inter-compartmental crosstalk and membrane contact sites make it difficult to dissect compartment-specific proteins based on prior knowledge (Go *et al*, 2021; Eisenberg-Bord *et al*, 2016). Thus, we used the PL proteome generated by cytosolic APEX2-construct (NES-APEX2) as control and required proteins that were significantly overrepresented in the subcellular compartments compared to the control to be exclusively present in one compartment. Using these filtering criteria, we identified 337, 55, and 228 location-specific proteins for endosome and lysosome/late endosome, Golgi, and plasma membrane, respectively. The majority of our compartment-specific proteins matched their location defined in the Human Cell Map (Go *et al*, 2021), a proximity map of the HEK293 proteome generated using PL. Differences between our dataset and the Human Cell Map could be explained by the usage of distinct proximity labeling enzymes (APEX2 versus BioID), different proteins or protein localization domains to target the enzyme to a subcellular location, as well as different data analysis criteria. We envision that our scalable proximity labeling strategy will allow investigating dynamics in subcellular protein localization across many conditions.

Finally, we applied the automated proximity proteomics strategy to investigate temporal protein interaction dynamics. Previously, we demonstrated that APEX2-based PL can capture protein

interaction networks of GPCRs and their agonist dependent changes with temporal and spatial precision (Lobingier *et al*, 2017; Paek *et al*, 2017). To ensure consistent and reproducible quantification over the time course, we used a two-step quantitative MS workflow consisting of DDA for identification followed by targeted proteomics for quantification of biotinylated proteins in our previous study (Lobingier *et al*, 2017). Here, we demonstrate that our automated PL approach combined with DIA-MS provides high reproducibility to quantify agonist-induced proximal proteome changes over a time course. We map the proximal protein interaction network changes of $5HT_{2A}$ upon activation with its endogenous ligand 5HT (**Figure 5**). The 744 proteins significantly changing in the $5HT_{2A}$ PL dataset showed down-regulation of plasma membrane proteins and up-regulation of endosomal proteins indicating internalization and trafficking of the receptor following activation. To account for location-specific proteins that reside in the local environment of $5HT_{2A}$ and enrich for functionally relevant proteins in the proximal interaction network of the receptor, we filtered the dataset for significant proteins with a defined fold change cutoff. This filtering criteria retained proteins with well-characterized functions in GPCR signaling and trafficking such as GRK2, ARRB2, PKC and PKD. Notably, given the transient nature of these interactions, it has been difficult to capture them using classic affinity purification coupled to MS. However, the simplified assumption that all location specific proteins follow a similar trend in labeling might not hold true for all proteins in the dataset. Furthermore, while HEK293T cells provide a great model system to demonstrate the power of the automated PL approach to detect PPI dynamics with high sensitivity, these cells do not express the $5HT_{2A}$ receptor endogenously. Therefore, we emphasize that PPI mapping in other cell types with endogenous $5HT_{2A}$ expression or functional characterization is necessary to corroborate any hypotheses derived from the proximal interaction networks.

Collectively, we established a proximity labeling proteomics approach allowing for mapping of subcellular proteomes as well as the spatial- and time-resolved proximal interaction networks with high throughput and reproducible quantification. The method presented here is attractive for any large-scale proximity labeling studies. While we demonstrated this automated proximity proteomics for APEX2-based PL and implemented it on a KingFisher Flex platform, we expect that it can be (1) widely extended for any other PL enzymes, which function by biotinylating proteins, and (2) readily transferred to other automated platforms which can handle magnetic beads.

Materials and Methods

Reagents and Tools

Reagent/resource	Reference or source	Identifier or catalogue number
Cell lines		
HEK293T/17 cells	ATCC	CRL-11268
Antibodies		
Mouse anti-FLAG (M1)	Sigma	F-3040
Mouse anti-Vimentin	Invitrogen	MA3-745
Rabbit anti-E Cadherin	CST	3195T
Mouse anti-RAB5A	Fisher	89333555
Mouse anti-RAB9	Invitrogen	MA3-067
Rabbit anti Golgin-97	Invitrogen	A21270
Chicken anti-GFP	VWR	RL600-901-215

AF488-labeled goat anti-Chicken	Invitrogen	A11039
AF555-labeled goat anti-Mouse	Invitrogen	A21422
AF647-labeled goat anti-Rabbit	Invitrogen	A21244
Rabbit HRP anti-GAPDH	Bio-legend	607904
HRP anti-Streptavidin	VWR	N100
Recombinant DNA		
NES-APEX2	This study	N/A
2xFYVE-APEX2	This study	N/A
GalT-APEX2	This study	N/A
LAMP1-APEX2	This study	N/A
LAMTOR1-APEX2	This study	N/A
Lyn11-APEX2	This study	N/A

$5HT_{2A}$-APEX2	This study	N/A
$G\alpha_q$-RLuc8	Addgene	140982
$G\beta_3$	Addgene	140988
GFP2-$G\gamma_9$	Addgene	140991
$5HT_{2A}$	(Kim *et al*, 2020)	N/A
Chemical, enzymes and other reagents		
Dulbecco's Modification of Eagle's Medium (DMEM)	Corning	10-013-CV
Fetal bovine serum (FBS)	Gibco	A31605-01
Dulbecco's Phosphate-Buffered Saline (DPBS)	Corning	21-031-CV
0.05% Trypsin-EDTA	Gibco	25300-054
Penicillin streptomycin solution	Corning	30-002-CI
PolyJet	SignaGen	SL100688
0.45 μm PVDF filter unit	Millipore	SE1M003M00

Cellstripper	Corning	25-056-CI
Attune performance tracking beads	Thermo Fisher	4449754
ECL western blotting substrate	Pierce	32106
Streptavidin magnetic beads	Pierce	88817
Biotin phenol (Biotin Tyramide)	Iris Biotech	LS-3500
Hydrogen peroxide	Sigma	H1009-100ML
Sodium azide	Sigma	S2002
Sodium ascorbate	Spectrum	S1349
Trolox (6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid)	Sigma	238813
16% Formaldehyde (w/v)	Pierce	28908
EverBrite mounting medium with DAPI	Biotium	23002
cOmplete protease inhibitor cocktail tablets mini, EDTA-free	Roche	11846170001
660nm protein assay kit	Pierce	22662

TCEP	Pierce	20490
Dithiothreitol (DTT)	Sigma	D0632
Iodoacetamide (IAA)	Sigma	I1149
Sequencing-grade modified trypsin	Promega	V5111
Lysyl endopeptidase	VWR	100369-822
BioPureSPE 96-well C18 plate	NEST	HNS S18V
Ascorbic Acid	Sigma	A5960
BSA, Free Fatty Acid	Akron	AK8909
Coelenterazine 400a	Nanolight	340-1
Dialyzed FBS	Omega Scientific	FB-03
Hank’s Balanced Salt Solution	Gibco	14065056
Sodium chloride	Sigma	S7653-250G
Tris HCl pH 7.5	Corning	46-030-CM

Tris HCl pH 8.0	Corning	46-031-CM
Sodium deoxycholate	Sigma	D6750-500G
Triton-x100	Sigma	T9284-100ML
Sodium dodecyl sulfate	Fisher BioReagents	BP166-500
Urea	Promega	V3171
Potassium chloride	Sigma	P9541-1KG
Sodium carbonate	Sigma	S2127-500G
Pierce 660 nm Protein Assay Reagent	Thermo Scientific	22660
Software		
Spectronaut (version 16.0)	Biognosys	https://biognosys.com/shop/spectronaut
MSstats (version 4.4.0)	Bioconductor	https://www.bioconductor.org/packages/release/bioc/html/MSstats.html
Attune NxT software	Thermo Fisher Scientific	

FlowJo	FlowJo, LLC	https://www.flowjo.com/
Fiji	imageJ	https://imagej.net/software/fiji/
NIS-Elements software (v. 5.30.01 build 1541)	Nikon	
Prism (v8.0)	GraphPad Software	
Other		
Orbitrap Exploris 480 MS with internal calibration option	Thermo Fisher Scientific	BRE725533
EASY-nLC 1200 system	Thermo Fisher Scientific	
Nikon Ti2-E microscope	Nikon	
Attune NxT Flow Cytometer	Thermo Fisher Scientific	
KingFisher Flex system	Thermo Fisher Scientific	

Methods and Protocols

Mammalian Cell Culture Conditions

HEK293T cells (CRL-1583, ATCC) were maintained in Dulbecco's Modified Eagle's Medium (DMEM, GIBCO or Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS) (GIBCO) and 1% penicillin/streptomycin solution at 37 ^{o}C in a 5% CO_2 humidified incubator.

cDNA constructs

Standard cloning techniques were used to generate all constructs. Vectors were digested using restriction enzymes (EcoRI and BamHI) and PCR fragments were amplified using Q5 DNA polymerase (NEB). NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, GalT-APEX2, and $5HT_{2A}$-APEX2 were cloned into pLVX-TetOne-Puro by in-Fusion HD cloning. Ligated plasmid products were transformed into stellar competent cells (*E. coli* HST08 strain, Takara Bio). See **Table S1** for full descriptions of constructs.

Lentiviral generation and transductions

To generate lentiviruses, 8 x 10^6 HEK293T cells were seeded in T175 flask. The next day, cells were transfected with 5 µg of each of the pLVX-TetOne-Puro lentiviral plasmids containing the gene of interest, 3.33 µg of Gag-Pol-Tat-Rev packaging vector, and 1.66 µg of VSV-G envelope vector with 30 µL PolyJet reagent (3 µL per µg DNA) in 250 µL serum-free DMEM (25 µL per µg DNA). After 24 hrs, the culture media was replaced with fresh DMEM supplemented with 10% FBS and 1% Penicillin-Streptomycin solution. Virus was harvested at 48 hrs post media change. Briefly, the media was collected and centrifuged at 2500 rpm for 5 minutes. The supernatant was filtered through a 0.45-µm PVDF filter and mixed with the PEG-salt solution (a final 8.5% PEG and 0.3 M NaCl). The virions were aggregated and precipitated at 4 ^{o}C for at least 2 hrs, followed by centrifugation at 3500 rpm for 20 minutes. Immediately after spinning, the virus pellet was resuspended with 500 µL sterile PBS. For generation of stable APEX2-tagged cell lines, HEK293T cells were infected with lentivirus at ~50% confluency and selected by 2 µg/mL puromycin in growth medium.

Flow cytometric analysis

To create monoclonal stable cell lines, polyclonal cells were single cell sorted in a 96-well plate by fluorescence-activated cell sorting (FACS) using BD FACS Aria or BD FACS Fusion. Polyclonal cells were washed once with DPBS, collected by using Cellstripper (Corning). In the case of non-GFP tagged expression construct ($5HT_{2A}$-APEX2), cells were incubated with Alexa 647-conjugated M1-anti-FLAG antibody (1:1000 dilution) at 4 ^{o}C for 40 min. After spinning down, cells were washed with ice-cold PBS and resuspended in 1mL of sample buffer (PBS/1mM EDTA/25 mM HEPES/1% FBS) for sorting. 24 clones were selected and expanded in a 24-well plate for characterization by flow cytometry using the Attune NxT Flow Cytometer. The cellular morphology was first gated based on an FSC-A vs. SSC-A (forward scatter - area vs. side scatter - area) plot. Next, gating was generated based on the FSC-A vs. FSC-H (area vs. height) plot to select singlets. These singlets were further analyzed to identify the cells expressing $5HT_{2A}$ or the localization domains based on Alexa 647 or GFP fluorescence signal, respectively. The data was processed in FlowJo.

Confocal fluorescence imaging

HEK293 cells expressing NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, and GalT-APEX2 were plated onto 12 mm collagen coated glass coverslips in 12-well plate and treated with 1 µg/mL doxycycline for 24 hours. For initial polyclonal and

monoclonal cell lines selection, after aspirating culture media, cells were fixed with 4% paraformaldehyde in PBS at room temperature (RT) for 15 min. Cells were then mounted onto glass slides with EverBrite mounting medium with DAPI (Biotium Cat #23002), and ready for imaging. For the final monoclonal cell line characterization, the localization of all spatial APEX2 constructs were determined by the co-localization analysis with subcellular location markers. Cells were first washed with PBS twice and fixed in 4% paraformaldehyde in PBS for 15 min at room temperature (RT). After washing with PBS, cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min at RT and blocked in 5% BSA (w/v) in PBS with 0.1% Tween-20 (PBS-T) for 1h at RT. Incubation with primary antibodies diluted in 1% BSA in PBS-T against GFP (1:500) and Vimentin (1:100) for NES, GFP (1:500) and E-cadherin (1:500) for Lyn11, Rab 5 (1:200) for 2xFYVE , Rab9 (1:200) for LAMTOR1 and LAMP1, Golgin-97 (1:200) for GalT, and Alexa 647-conjugated M1-anti-FLAG antibody (1:200) for $5HT_{2A}$ was performed at 4 ^{o}C overnight. After washing with PBS-T three times, cells were incubated with AF488-labeled goat anti-Chicken, AF555-labeled goat anti-Mouse, and AF647-labeled goat anti-Rabbit secondary antibodies diluted in 1% BSA in PBS-T (1:1000) for 1 hour at RT. Cells were then washed in PBS-T twice and one final wash in PBS, mounted onto glass slides using EverBrite mounting medium with DAPI (Biotium Cat #23002), and imaged by confocal fluorescence microscopy. Images were generated using a Nikon Ti2-E microscope equipped with a Crest X-Light-V2 spinning disk confocal (Crest Optics), emission filters, 438/24 (DAPI), 511/20 (GFP), 560/25 (RFP), and 685/40 (Cy5), and Celeste Light Engine excitation lasers 405/477/546/638 nm used respectively (Lumencor), Piezo stage (Mad City Labs), and a Prime 95B 25mm CMOS camera (Photometrics) using a Plan Apo VC 100x/1.4 Oil (Nikon). The data was captured with NIS-Elements software (v. 5.41.01 build 1709, Nikon) and processed in Fiji/ImageJ2 (Schindelin *et al*, 2012).

BRET assay

Bioluminescence resonance energy transfer (BRET) assays were performed to validate the functionality of the $5HT_{2A}$-APEX2 construct. Here, the TRUPATH platform for measuring G protein dissociation was used as previously described (Olsen *et al*, 2020; Kim *et al*, 2020; DiBerto *et al*, 2022). In brief, wildtype or APEX2 constructs of 5-HT_{2A} were co-transfected with Ga_q-RLuc8, Gb_3, and GFP2-Gg_9 in a 1:1:1:1 ratio in HEK293T cells maintained in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (pen-strep). After 8 hours, the media were exchanged for DMEM supplemented with 1% dialyzed FBS and 1% pen-strep to minimize receptor desensitization by serum serotonin, and at least 12 hours before the experiment, cells were plated in 96-well microplates in DMEM supplemented with 1% dialyzed FBS and 1% pen-strep. To conduct the experiments, plates were first vacuum aspirated and 60 mL of assay buffer (1x Hank's Balanced Salt Solution in phosphate buffered saline, 20 mM HEPES, pH 7.4) were added to the wells. Next, 10 mL of coelenterazine 400a diluted in the assay buffer (50 mM) were added to the wells, and allowed to incubate for 5 minutes. Next, 30 mL of 3x drug diluted in drug buffer (assay buffer supplemented with 3 mg/mL fatty acid-free bovine serum albumin and 0.3 mg/mL ascorbic acid) were added to the wells, and allowed to incubate for 5 minutes. Last, plates were read using a BMG Labtech PHERAstar FSX with $BRET^2$ plus optic module. All data were analyzed using GraphPad Prism (V8.0).

APEX proximity labeling

HEK293T cells stably expressing $5HT_{2A}$-APEX2 and protein localization domains (NES-APEX2, Lyn11-APEX2, 2xFYVE-APEX2, LAMTOR1-APEX2, LAMP1-APEX2, and GalT-APEX2) were seeded with 350K cells/well in 6-well plate for testing biotinylation by western blot, and $3.5x10^6$ cells/dish in 10cm dish for enrichment of biotinylated proteins followed by quantitative MS. Following 24 hrs doxycycline induction, cells were incubated with 500 μM biotin phenol at 37 °C for 30 min in complete medium (DMEM/10% FBS/1% pen/strep), but $5HT_{2A}$-APEX2 cells were seeded with 10% dialyzed FBS/DMEM/1% pen/strep and changed to 1% dialyzed FBS/DMEM/1% pen/strep during doxycycline induction. Meanwhile, 0.1 μM 5HT was added for incubation of 1, 5, 10, and 30 min in cells expressing $5HT_{2A}$-APEX2 to capture the protein interaction dynamics. There was no agonist incubation for all APEX2-tagged localization domains. 2 mM H_2O_2 diluted in complete medium was freshly prepared prior to use. The H_2O_2 containing media was mixed with biotin phenol containing media at 1:1 ratio (*v*/*v*) to initiate APEX labeling. The labeling reaction was quenched after 45 sec by removing the medium, washing cells three times with ice-cold quenching solution (PBS supplemented with 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide). Cells were collected in a quenching solution (1 mL for 6-well plate and 8 mL for 10cm dish) and pelleted by centrifugation at 3,000 g for 10 min at 4 °C. The supernatant was removed and the cell pellet was resuspended in lysis buffer (50 mM Tris, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide, 1 mM DTT, and cOmplete protease inhibitor) supplemented with 10 mM sodium ascorbate, 5 mM Trolox, and 10 mM sodium azide, 1mM DTT, and protease inhibitor (100 μL for small-scale analysis and 1 mL for large-scale analysis). With a freeze-thaw cycle, each sample was sonicated for 5 seconds twice, centrifuged at 13,000 g for 10 min at 4 °C. The supernatant was subjected to the automated enrichment of biotinylated proteins. The manual enrichment was performed in parallel with the automated enrichment protocol, but using a magnetic rack (DynaMag-2, Thermo Fisher). The protein concentration was determined using Protein assay 660 (Pierce). For each proteomic sample, 25 μL lysate before and after binding was saved for western blot analysis to evaluate biotinylation and enrichment efficiency of biotinylated proteins.

Western blot analysis

Cell lysates were mixed with equal volume of sample loading buffer (200 mM β-mercaptoethanol in NuPAGE LDS sample buffer), boiled at 95 °C for 20 min, separated on 4%-12% SDS-PAGE gels (BioRad), and transferred to PVDF membranes (Trans-blot turbo transfer system, BioRad). The blots were blocked in 5% milk in TBS-T (Tris-buffered saline with 0.1% Tween 20) at 4 °C overnight and washed three times with TBS-T for 10 min. Blots were incubated with horseradish peroxidase (HRP)-conjugated anti-Streptavidin antibody (1:5000 dilution using 2% BSA in TBS-T) or HRP-conjugated anti-GAPDH antibody (1:10000 dilution) at room temperature for 1 hour. After washing three times in TBS-T for 10 min, the blots were incubated with ECL western blotting substrate (Pierce) for 1 min. Chemiluminescent signals were captured on the Azure 400 (Azure biosystem).

Automated enrichment protocol for biotinylated proteins

For the automated biotinylation enrichment protocol, the Kingfisher Flex system (Thermo Fisher) is programmed to simultaneously process a maximum of 96 samples. This protocol below includes two parts, where part 1 (Plate 1-3) is for washing magnetic streptavidin beads and binding of the biotinylated proteins to beads, and part 2 (Plate 4-12) is for washing and collecting beads prior to Lys-C/trypsin digestion. The enrichment protocol is conducted in the cold room using deep-well plates. See **Table S2** for full description of the Kingfisher program.

Plate 1. Add 100 µL of 50% streptavidin magnetic bead slurry (Pierce) to each well of a 96-well plate and add 900 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, pH 7.4) to each well. The beads are mixed for 2 seconds and slowly washed for 30 seconds. At the end of this step, the beads are collected for 5 counts with collection time of 10 seconds/count. A total of 4 min for washing.

Plate 2. Add 1 mL RIPA buffer. The beads are transferred to this plate and slowly washed for 4 min.

Plate 3. Add 1 mg protein from cell lysate to each well and top-up to 1 mL with RIPA buffer. The beads are transferred to this plate for overnight binding (~12 hrs).

Plate 4-6. Add 1 mL RIPA buffer to each plate. The beads with bound biotinylated proteins are collected from Plate 3 and transferred to Plate 4-6 for washing three times with RIPA buffer. Each wash step takes 7 min.

Plate 7. Add 1 mL 1 M KCl solution. Wash the beads for 4 min.

Plate 8. Add 1 mL 0.1 M Na_2CO_3 solution. Wash the beads for 4 min.

Plate 9. Add 1 mL freshly prepared 2 M urea in 50 mM Tris-HCl (pH 8.0). Wash the beads for 4 min.

Plate 10-11. Add 1 mL 50 mM Tris-HCl (pH 8.0). Wash the beads twice for 5 min.

Plate 12. Add 200 µL of freshly prepared 2 M urea in 50 mM Tris-HCl buffer (pH 8.0). The beads with bound biotinylated proteins are collected in this plate for on-bead digestion.

Sample preparation for MS analysis

Samples were reduced with 5 mM TCEP at 37 ^{o}C for 30 min by shaking at 1000 rpm, followed by alkylation with 5 mM iodoacetamide (IAA) at room temperature (RT) for 30 min. Extra IAA was quenched by 5 mM dithiothreitol (DTT) for 10 min at RT. Samples were then digested with 1 µg Lys-C and 1 µg trypsin at 37 ^{o}C for 6 hrs and 25 ^{o}C for 13 hrs. 0.5 µg trypsin was added to each sample for an additional 2 hrs incubation at 37 ^{o}C. Following digestion, the supernatant was transferred to a new 96-well plate and 10 µL of 10% trifluoroacetic acid (TFA) was added to each sample to a final pH 2-3. The peptide samples were desalted using C18 96-well plate (BioPureSPE, HNS S18V-20mg, the Nest group) by centrifugation at 1,600 rpm for 2 min at each step. The plate was activated with 100 µL methanol, washed three times with 100 µL 80% acetonitrile (ACN)/0.1% TFA, and equilibrated three times with 100 µL 2% ACN/0.1% TFA. After loading samples, the plate was washed with 3 x 100 µL 2% ACN/0.1% TFA and eluted with 2 x

100 µL 50% ACN/0.25% formic acid (FA). Samples were dried by vacuum centrifugation prior to MS analysis.

Mass spectrometric data acquisition

Digested peptide mixtures were analyzed on an Orbitrap Exploris 480 MS system equipped with an Easy nLC 1200 ultra-high pressure liquid chromatography system. Samples were injected on a C18 reverse phase column (15 cm x 75 µm I.D. packed with BEH C18 1.7 µm particles, Waters) in 0.1% formic acid (FA). Mobile phase A consisted of 0.1% FA and mobile phase B consisted of 80% acetonitrile (ACN)/0.1% FA. Peptide mixtures were separated by mobile phase B ranging from 4% to 16% over 40 min, followed by an increase to 28% B over 26 min and 44% B over 4 min, then held at 95% B for 10 min at a flow rate of 300 nL/min.

To generate a spectral library, one of three biological replicates in each experiment was acquired using data-dependent acquisition (DDA). DDA analysis consisted of one full scan over a *m*/*z* range of 350-1050 in the Orbitrap at a resolving power (RP) of 120K with an RF lens of 40%, a normalized automatic gain control (AGC) target of 300%. The 20 most intense precursors at charge states of 2-6 from the full scan were selected for higher energy collisional dissociation (HCD) $FTMS^2$ analysis at RP 15K with an isolation width of 1.6 Da, a normalized collision energy (NCE) of 30, an AGC target of 200%, and a maximum injection time of 22 ms. Dynamic exclusion was enabled for 30 seconds with a repeat count of 1.

The parameters evaluated for data-independent acquisition (DIA) method include mass scan range of 350-1050 and 390-1010, MS^2 scans with a normalized HCD (higher-energy collisional dissociation) collision energy of 27%, 30% and 33%, maximum injection time of 22 ms, 32 ms, 42 ms, dynamic 6-point, dynamic 8-point, and Auto, and window overlap width of 0, 1, and 2 *m*/*z*. In the end, all samples were analyzed using DIA-MS by collecting one full scan over a range of 350-1050 *m*/*z* in the Orbitrap at RP 120K with an AGC target of 300% and an automatic maximum injection time, followed by DIA MS^2 scans over 350-1050 *m*/*z* at RP 15K using an isolation window of 20 *m*/*z*, an overlapping isolation window of 2 *m*/*z*, a normalized HCD collision energy of 30%, an AGC target of 200%, and a maximum injection time of 32 ms.

Protein identification and quantification

The DDA raw data for automated enrichment combined with data-dependent acquisition MS (Auto_DDA) were analyzed using MaxQuant (version 1.6.12.0) for identification of peptides and proteins. The DIA raw data for manual enrichment protocol combined with data-independent acquisition MS (Manual_DIA), automated enrichment protocol with data-independent analysis (Auto_DIA), and DIA-MS method optimization were analyzed using Spectronaut (version 16.0) with direct DIA search. All the raw files were searched against the SwissProt Human database (downloaded 10/2020) using default settings with variable modification of methionine oxidation and protein N-termini acetylation, and fixed modification of cysteine carbamidomethylation. The data were filtered to obtain a false discovery rate of 1% at the peptide spectrum match and the protein level.

In the data analysis of mapping subcellular proteomes or protein interaction dynamics, the DDA datasets were collected to construct a spectral library. All the raw data files were searched

against the SwissProt Human database (downloaded 10/2020) by using the Pulsar search engine integrated into Spectronaut (version 16.0). The raw mass spectra files were processed using the default BGS settings with digestion enzyme of trypsin, fixed modification of cysteine carbamidomethylation, variable modification of methionine oxidation and protein N-termini acetylation, and filtered to a 1% false discovery rate (FDR) at the peptide spectrum match (PSM), peptide, and protein level. The DIA datasets were searched against the generated DDA spectral libraries in Spectronaut with default BGS setting, but the cross-run normalization was disabled.

Statistical analysis of proteomic data and data visualization

Proteomic statistical analyses were completed using the R programming language, version 4.2.0. Peptide ion intensities from Spectronaut were summarized to protein intensities using the package MSstats (version 4.4.0) and its function dataProcess enabling options to turn off model-based imputation, and turning on removal of uninformative features and outliers (Tsai *et al*, 2020). All downstream analyses were performed on log2 transformed protein intensities.

Subcellular proteomes were compared for each protein or protein group using the built-in R function lm to estimate additive effects for both subcellular location and replicate batch, and the emmeans function from package emmeans (version 1.7.3) to calculate pairwise subcellular location $\log_2$ fold changes and t-test statistics including p values. Adjusted p values were calculated as false discovery rates using the Benjamini-Hochberg method available in R built-in function p.adjust. Selection of location-specific proteins was made by identifying proteins with significant labeling increases when compared to the cytosolic NES-APEX2 sample ($\log_2$FC > 1, unadjusted p.value < 0.005), and requiring the increase to be exclusive to GalT, Lyn11, or to the combined 2xFYVE, LAMTOR1, LAMP1 set. Gene ontology enrichment analysis of location-specific protein sets was performed using the enricher function from package clusterProfiler (version 4.4.1) and human gene ontology annotations in the Bioconductor annotation package org.Hs.eg.db (version 3.15.0). The Human Cell Map data were downloaded on 2022-July-06.

Proteomic labeling changes over 5 time points (0, 1, 5, 10, and 30 minutes) were scored per protein by fitting a cubic polynomial curve through the data using R built in functions lm and poly. A categorical term for replicate batch was included in the model, and time was continuous but transformed from minutes to ranks (1 through 5). P values were calculated using F-tests comparing the models with and without the polynomial time terms. The time point with the largest deviation of mean intensity from time point 0 was used to calculate the log2FC as the difference in mean log2 intensities. Time courses of protein labeling changes were clustered by starting with a protein-by-time matrix of differences in mean log2 intensity from time 0, calculating protein-protein distances as $1 - R$, where R is Pearson's correlation coefficient, and defining clusters with the pam function in package cluster (version 2.1.3), with option nstart set to 20.

Networks for proteins represented in the transient and sustained cluster were generated extracting interactions between proteins in each cluster with a confidence above 0.4 derived from STRING (https://string-db.org/) (Szklarczyk *et al*, 2011). Interactions were important into Cytoscape (version 3.8.1) for visualization (Shannon *et al*, 2003). The maximum log2-fold

change compared to time point 0 for each protein was visualized in the network as node color. Enrichment of Reactome pathways (https://reactome.org/) (Gillespie *et al*, 2022) was performed using Enrichr (https://maayanlab.cloud/Enrichr/) (Kuleshov *et al*, 2016; Xie *et al*, 2021; Chen *et al*, 2013).

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (https://www.ebi.ac.uk/pride/) (Deutsch *et al*, 2023) partner repository with the dataset identifier PXD040762. The reviewer account details are shown below.
Username: reviewer_pxd040762@ebi.ac.uk
Password: poggC9R8

Acknowledgements

This work was supported by funding from the Defense Advanced Research Projects Agency (DARPA) under the Cooperative Agreements HR0011-19-2-0020 (to N.J.K., R.H., B.L.R.) and HR0011-20-2-0029 (to N.J.K., R.H., B.L.R.). The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This work further received funding from the NIH (R01DA056354 to R.H. and M.v.Z.; P01HL146366 to N.J.K.). The work was carried out in the Thermo Fisher Scientific Mass Spectrometry Facility for Disease Target Discovery at the J. David Gladstone Institutes and the UCSF Center for Advanced Technology. The imaging data for this study were acquired at the Center for Advanced Light Microscopy-Nikon Imaging Center at UCSF on instruments obtained using grants from the UCSF Program for Breakthrough Biomedical Research funded in part by the Sandler Foundation and the UCSF Research Resource Fund Award. We thank Ajda Rojc, Prachi Khare, and Michael McGregor for the help with flow cytometry analysis and members of Krogan lab for helpful advice and comments.

Author Contributions

X.Z., Q.L., and R.H. conceived and directed the study with input from N.J.K., M.v.Z., and B.L.R. Proximity proteomics: X.Z. and Q.L.; Data analysis: B.J.P., R.H., X.Z., and Q.L.; Cell line development: T.P., J.F.D., Q.L., R.H., and J.X.; Confocal fluorescence imaging: R.V., X.Z., A.M., H.F., and Q.L.; Flow cytometry: T.P. and R.H.; BRET assay: J.F.D.; X.Z., R.H., Q.L., and M.E. wrote the manuscript with input from all authors.

Competing Interest Statement

The Krogan Laboratory has received research support from Vir Biotechnology and F. Hoffmann-La Roche. Nevan Krogan has consulting agreements with the Icahn School of Medicine at Mount Sinai, New York, Maze Therapeutics and Interline Therapeutics. He is a shareholder in Tenaya Therapeutics, Maze Therapeutics and Interline Therapeutics, and a financially compensated Scientific Advisory Board Member for GEn1E Lifesciences, Inc.

Figures

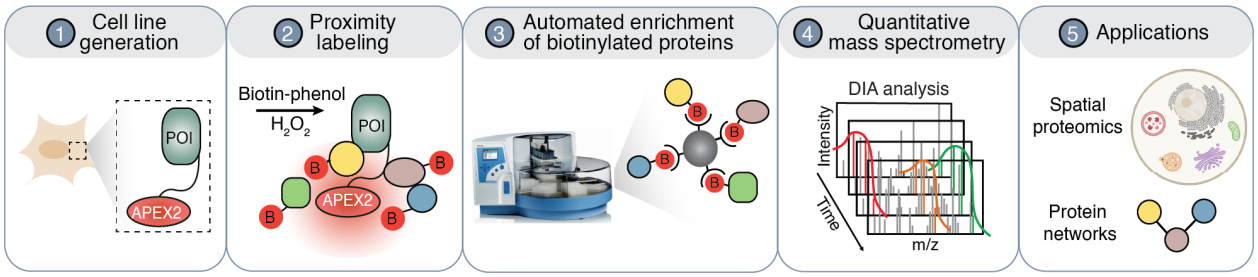


Figure 1. A schematic overview of universal workflow for proximity proteomics.

This workflow includes (1) PL construct design as well as generation and characterization of monoclonal cell lines, (2) proximity labeling, (3) automated enrichment of biotinylated proteins, (4) quantitative mass spectrometry data acquisition, and (5) the application of this workflow to map subcellular proteomics as well as protein-protein interactions and their dynamics.

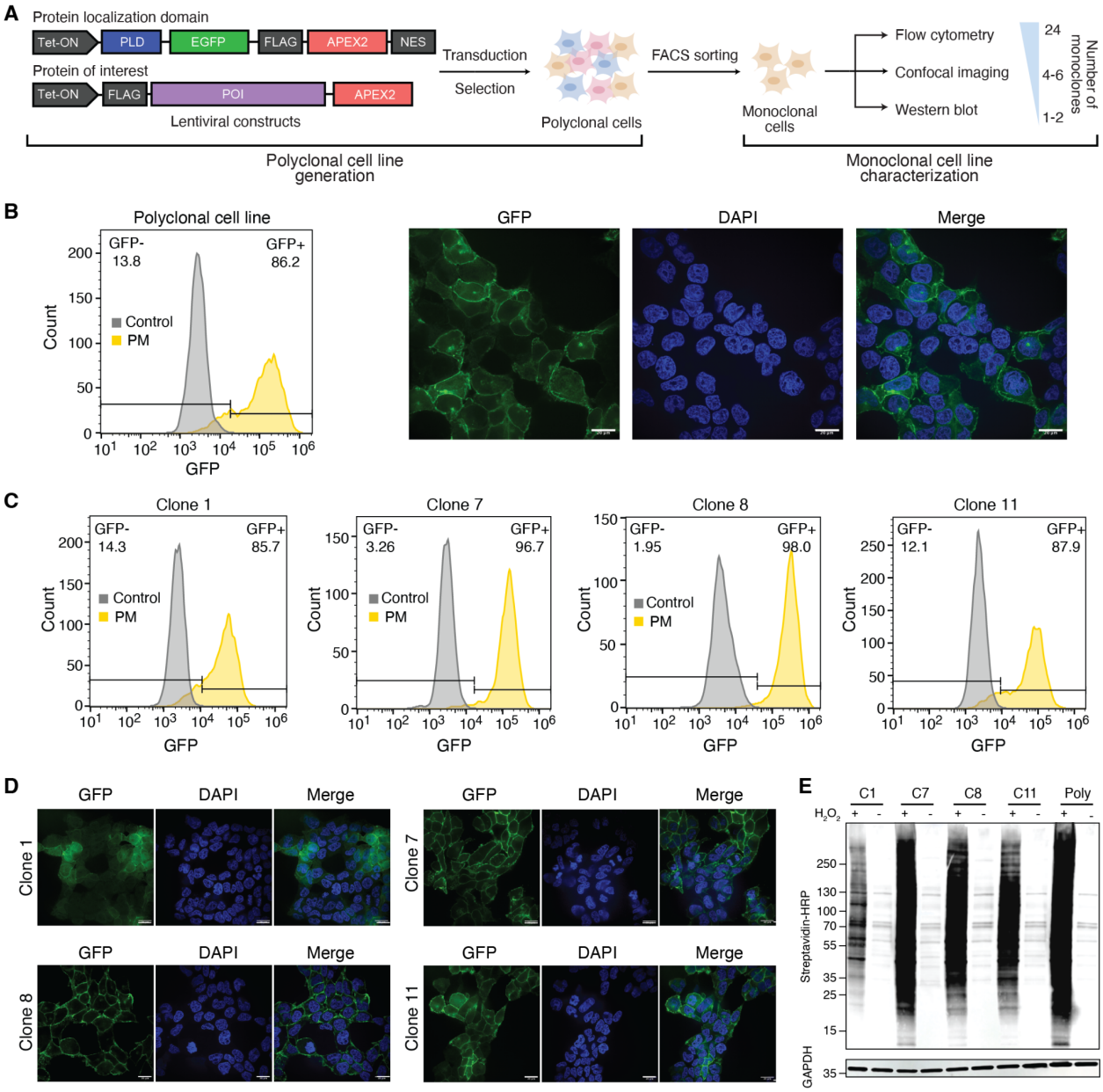


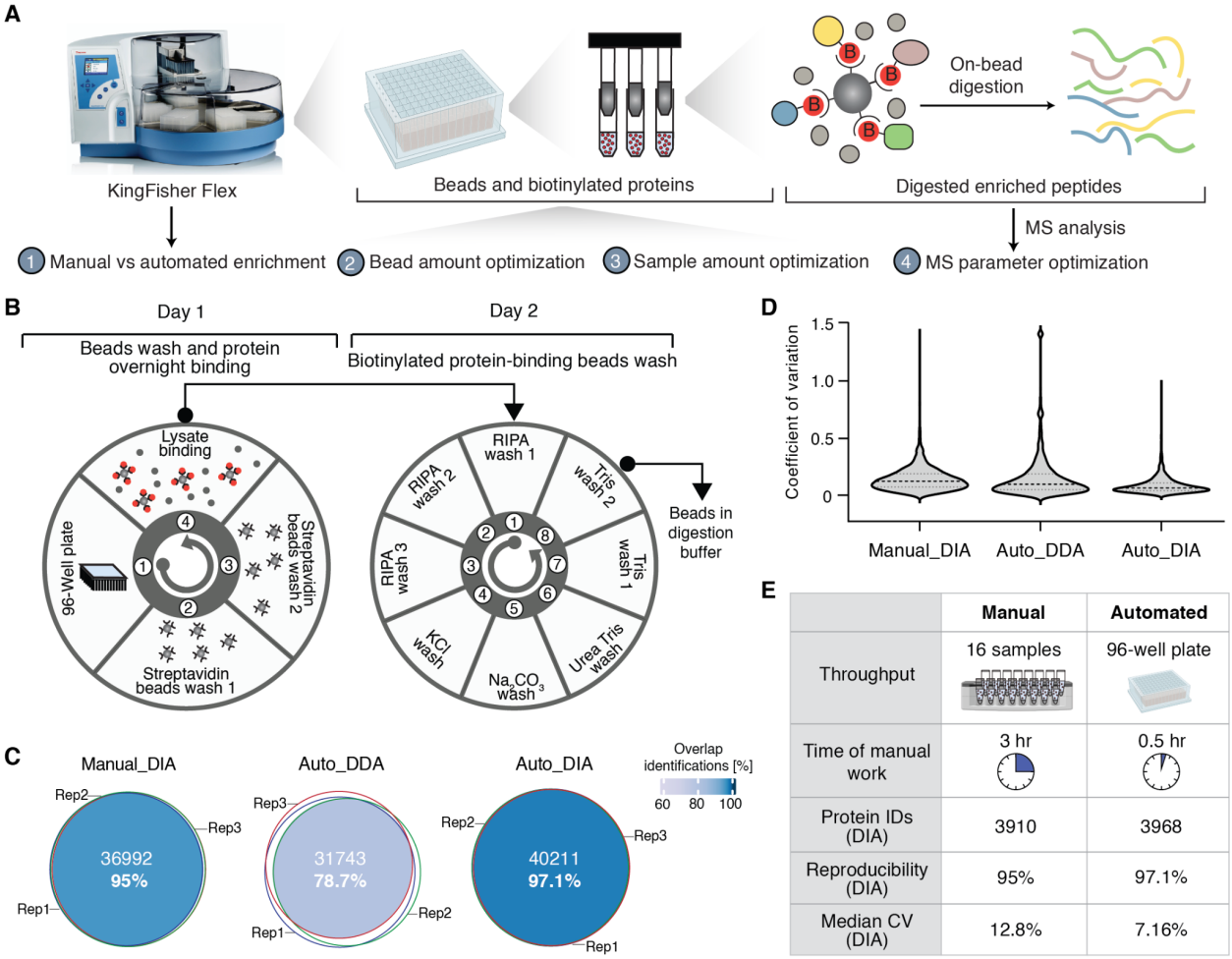
Figure 2. Step-by-step guide for generating monoclonal cell lines for proximity labeling.

A. Schematic overview of PL construct design and generation and stepwise characterization of monoclonal cell lines for PL (PLD refers to protein localization domain and POI to protein of interest).

B. Characterization of polyclonal proximity labeling cell lines by flow cytometry (left) and confocal imaging (right). The APEX2 enzyme was targeted to the plasma membrane using a Lyn11 protein localization domain (PM-APEX2). GFP was used to evaluate expression and localization of the APEX2 construct. The non-Doxycyline induced cell line was used as control for flow cytometry. Scale bar represents 20μm.

C. Flow cytometry analysis of four monoclonal PM-APEX2 cell lines.

D. Confocal imaging of four monoclonal PM-APEX2 cell lines. Cells were stained with DAPI and Hoechst 33342 for the nucleus and the APEX2 construct location was indicated by GFP. Scale bar represents 20μm.
E. Western blot analysis of whole cell lysate derived from four monoclonal PM-APEX2 cell lines to evaluate the biotin labeling efficiency of APEX2 using streptavidin antibody. GAPDH is used as loading control.



	Manual	Automated
Throughput	16 samples	96-well plate
Time of manual work	3 hr	0.5 hr
Protein IDs (DIA)	3910	3968
Reproducibility (DIA)	95%	97.1%
Median CV (DIA)	12.8%	7.16%

Figure 3. An automated, scalable biotin enrichment method coupled to quantitative MS.

A. Schematic overview of the development and optimization of the automated enrichment method for biotinylated proteins and quantitative mass spectrometry method.
B. Overview of the automated biotin enrichment protocol established on the KingFisher Flex.
C. Performance comparison of manual enrichment protocol combined with data-independent acquisition MS (Manual_DIA), automated enrichment combined with data-dependent acquisition MS (Auto_DDA), and automated enrichment protocol with data-independent analysis (Auto_DIA). Performance was evaluated comparing the protein identifications from three biological replicates.
D. Distribution of the coefficient of variation of three biological replicates comparing Manual_DIA, Auto_DDA, and Auto_DIA approaches.
E. Summary comparison of the manual and automated enrichment method for biotinylated proteins.

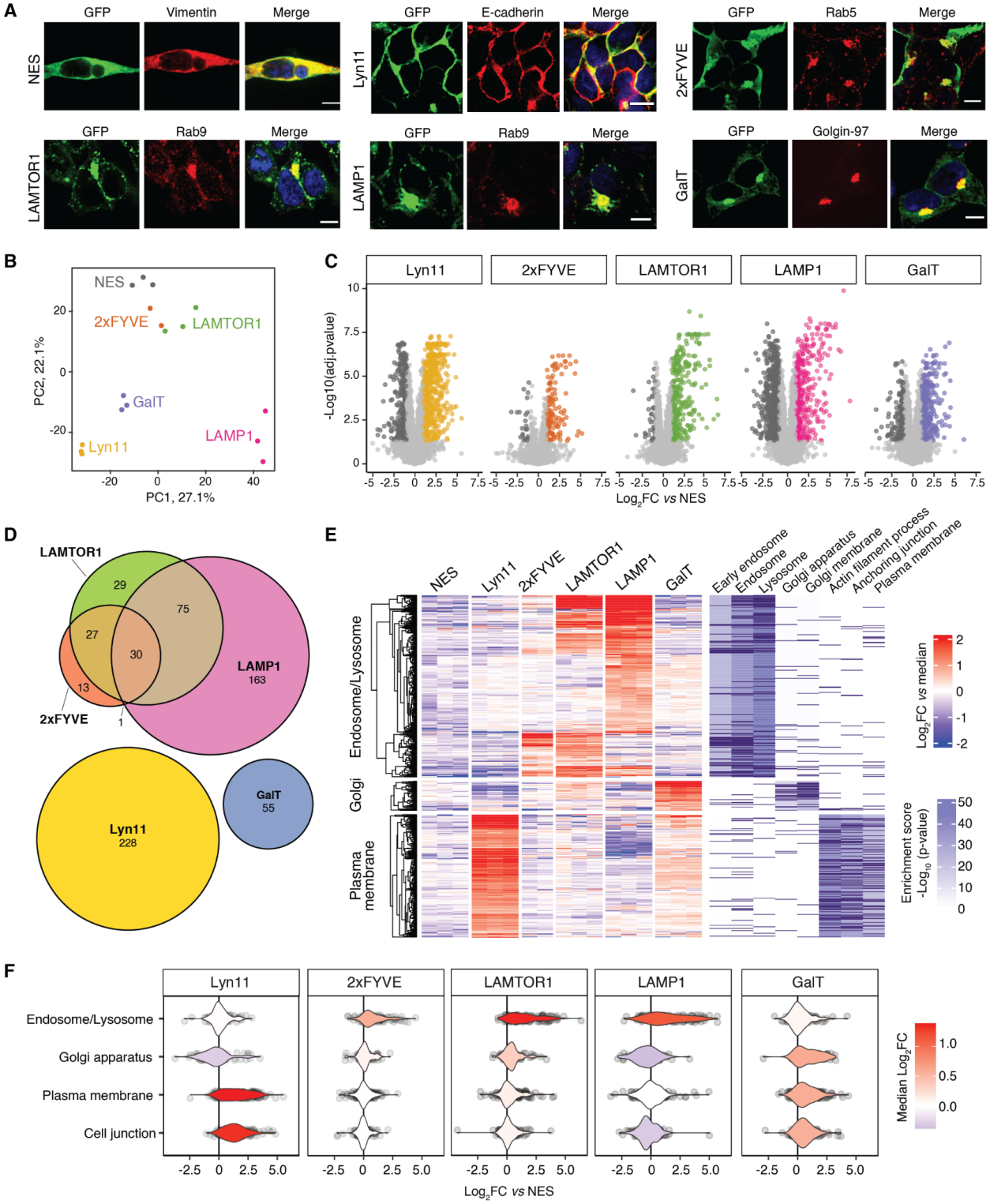

A
GFP
Vimentin
Merge
NES
E-cadherin
Lyn11
Rab5
2xFYVE
Rab9
LAMTOR1
LAMP1
Golgin-97
GalT
B
NES
2xFYVE
LAMTOR1
GalT
LAMP1
Lyn11
PC2, 22.1%
PC1, 27.1%
C
Lyn11
2xFYVE
LAMTOR1
LAMP1
GalT
-Log10(adj.pvalue)
Log2FC vs NES
D
LAMTOR1
29
75
27
30
13
1
LAMP1
163
2xFYVE
Lyn11
228
GalT
55
E
NES
Lyn11
2xFYVE
LAMTOR1
LAMP1
GalT
Early endosome
Endosome
Lysosome
Golgi apparatus
Golgi membrane
Actin filament process
Anchoring junction
Plasma membrane
Endosome/Lysosome
Golgi
Plasma membrane
Log2FC vs median
Enrichment score
-Log10 (p-value)
F
Lyn11
2xFYVE
LAMTOR1
LAMP1
GalT
Endosome/Lysosome
Golgi apparatus
Plasma membrane
Cell junction
Log2FC vs NES
Median Log2FC

Figure 4. Mapping subcellular proteomes using the automated proximity proteomics strategy.

A. Confocal fluorescence imaging of monoclonal cell lines of all spatial APEX2 constructs to co-localize with a subcellular location marker. APEX2-fusion constructs were stained using a GFP antibody. Location markers of vimentin, E-cadherin, Rab5, Rab9, and Golgin-97 were used for NES (cytosol), Lyn11 (plasma membrane), 2xFYVE (endosome), LAMTOR1 and LAMP1 (late endosome/lysosome), and GalT (Golgi apparatus), respectively. (Scale bars, 10 μm.)
B. Principal-component analysis (PCA) comparing proximal proteomes of each replicate of the spatial APEX2 constructs.
C. Volcano plots comparing proximal proteome of each location specific construct to the control (NES). Proteins with a $\log_2$ fold change >1 and adjusted p-value <0.05 were considered significant and colored. Data were collected from three independent biological replicates (except 2xFYVE).
D. Venn diagram comparing the location specific proteins.
E. Heatmap of location-specific proteins was highlighted as three groups based on the median normalized $\log_2$ fold change. Gene ontology (GO) enrichment analysis was performed for each cluster and the matched genes for each GO term are indicated in dark purple.
F. Comparison of the location proteins candidates identified in this study with the protein localizations defined in Human Cell Map (Go *et al*, 2021).

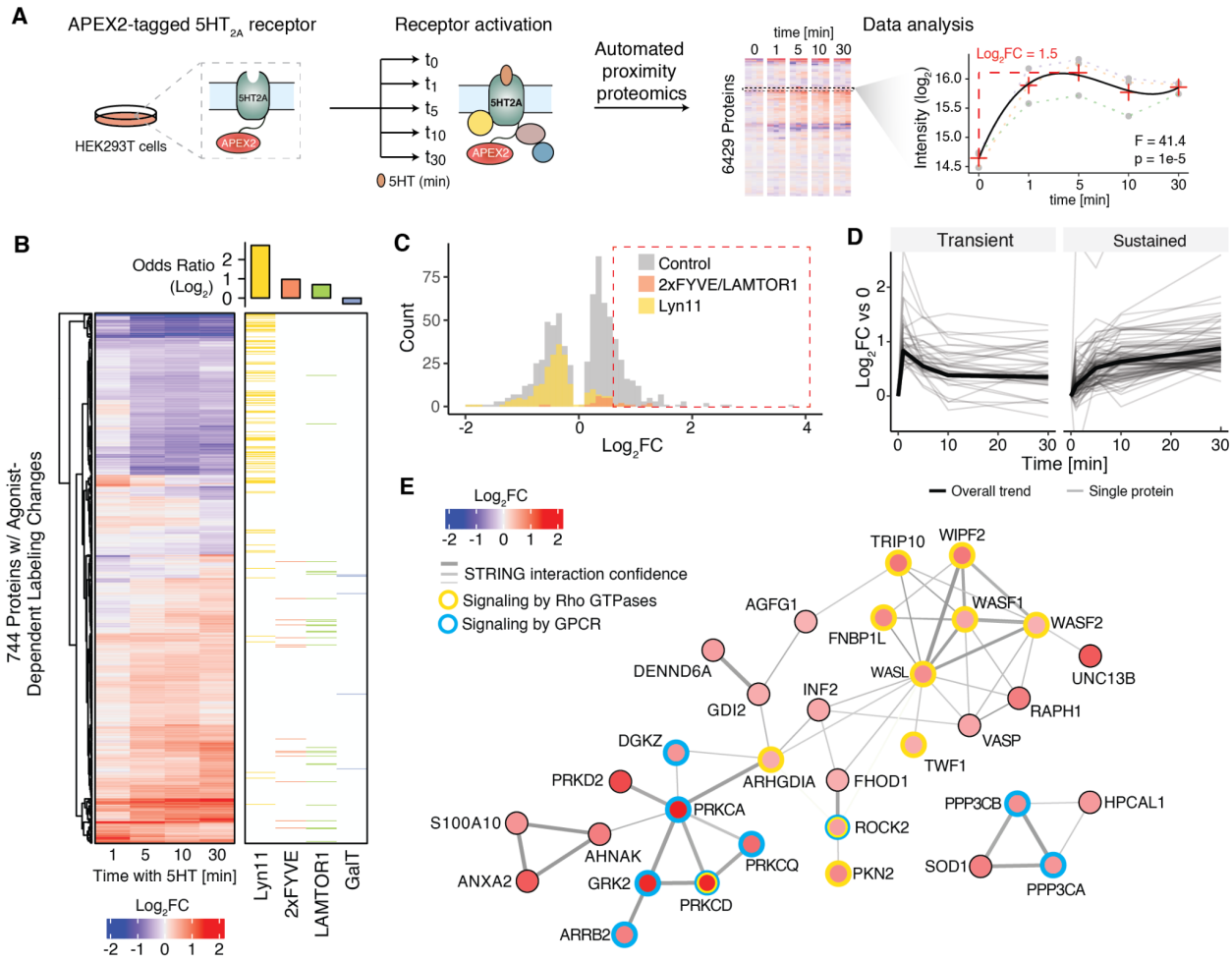


Figure 5. Mapping ligand-dependent proximal interaction network changes of $5HT_{2A}$.

A. Experimental design to map proximal interaction networks of $5HT_{2A}$. APEX2-fused $5HT_{2A}$ receptor was stably expressed in HEK293T cells, activated with 5HT and at indicated time points after agonist treatment proteins proximal of the receptor were biotinylated by addition of H_2O_2. Following automated enrichment of biotinylated proteins, the samples were analyzed using DIA.

B. Heatmap depicting proteins that significantly change in biotinylation after agonist addition. Overrepresentation of location-specific proteins for plasma membrane (Lyn 11), endosome (2xFYVE), late endosome/lysosome (LAMTOR1), and Golgi (GalT) was shown in the right heartmap. Data were collected from three independent biological replicates.

C. Log_2 fold change distribution of proteins in Figure B (Control: Grey). Endosomal/lysosomal proteins (2xFYVE/LAMTOR1) and plasma membrane proteins (Lyn11) were highlighted in orange and yellow, respectively.

D. Line charts of proteins filtered out from Figure C with a fold change cutoff of 1.5 clustered into proteins with transient and sustained responses to 5HT treatment.

E. Protein interaction network connecting proteins with transient agonist-dependent changes in the proximity of $5HT_{2A}$. Proteins shown as nodes were colored according to

their $\log_2$ fold change. The edges connecting the proteins were derived from STRING (Szklarczyk *et al*, 2011) and the edge width was scaled according to the interaction confidence. Proteins corresponding to Reactome pathways (Jassal *et al*, 2020) that were enriched within the transient cluster were indicated with different colored node borders.

Supplemental Figures

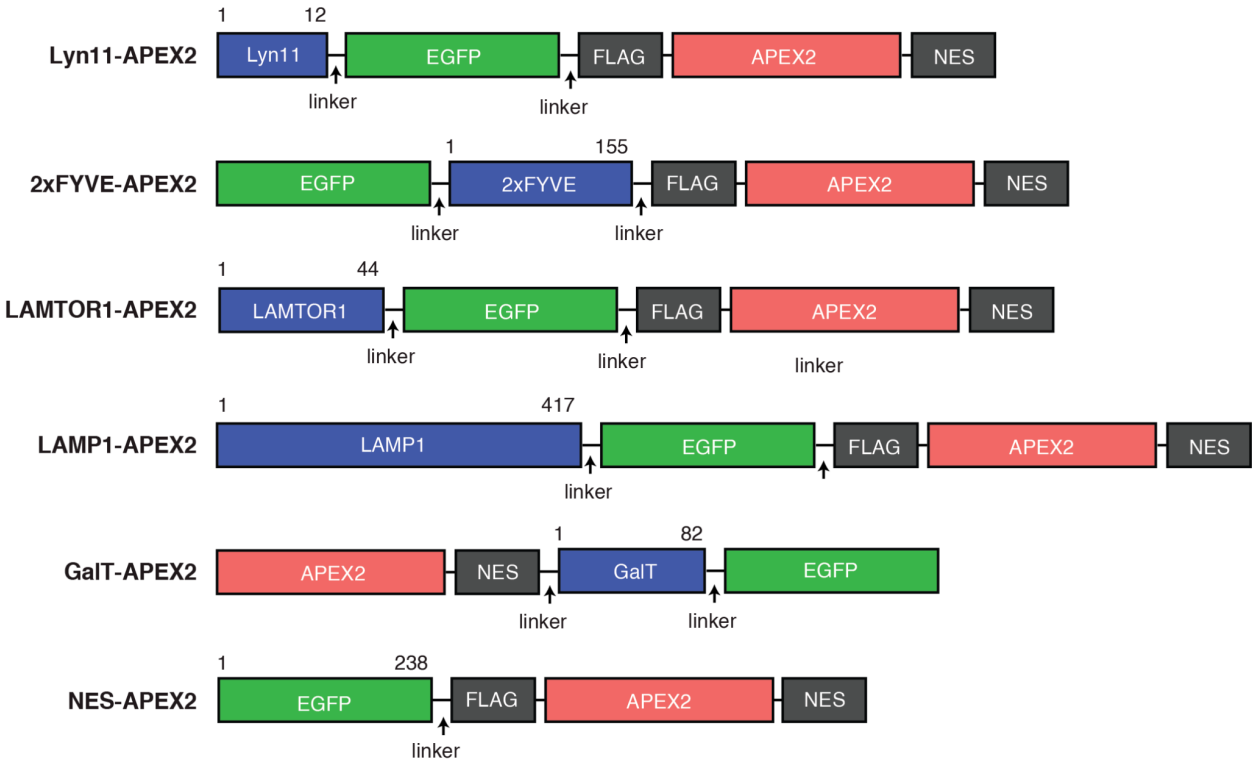


Figure S1. Scheme of APEX2-tagged cellular compartments.

In our study, we selected domain Lyn11 for plasma membrane, 2xFYVE for endosome, LAMTOR1 and LAMP1 targeting sequence for late endosome/lysosome, and β-1,4 galactosyltransferase (GalT) for Golgi apparatus. A non-location specific construct at cytosol was used as a localization control.

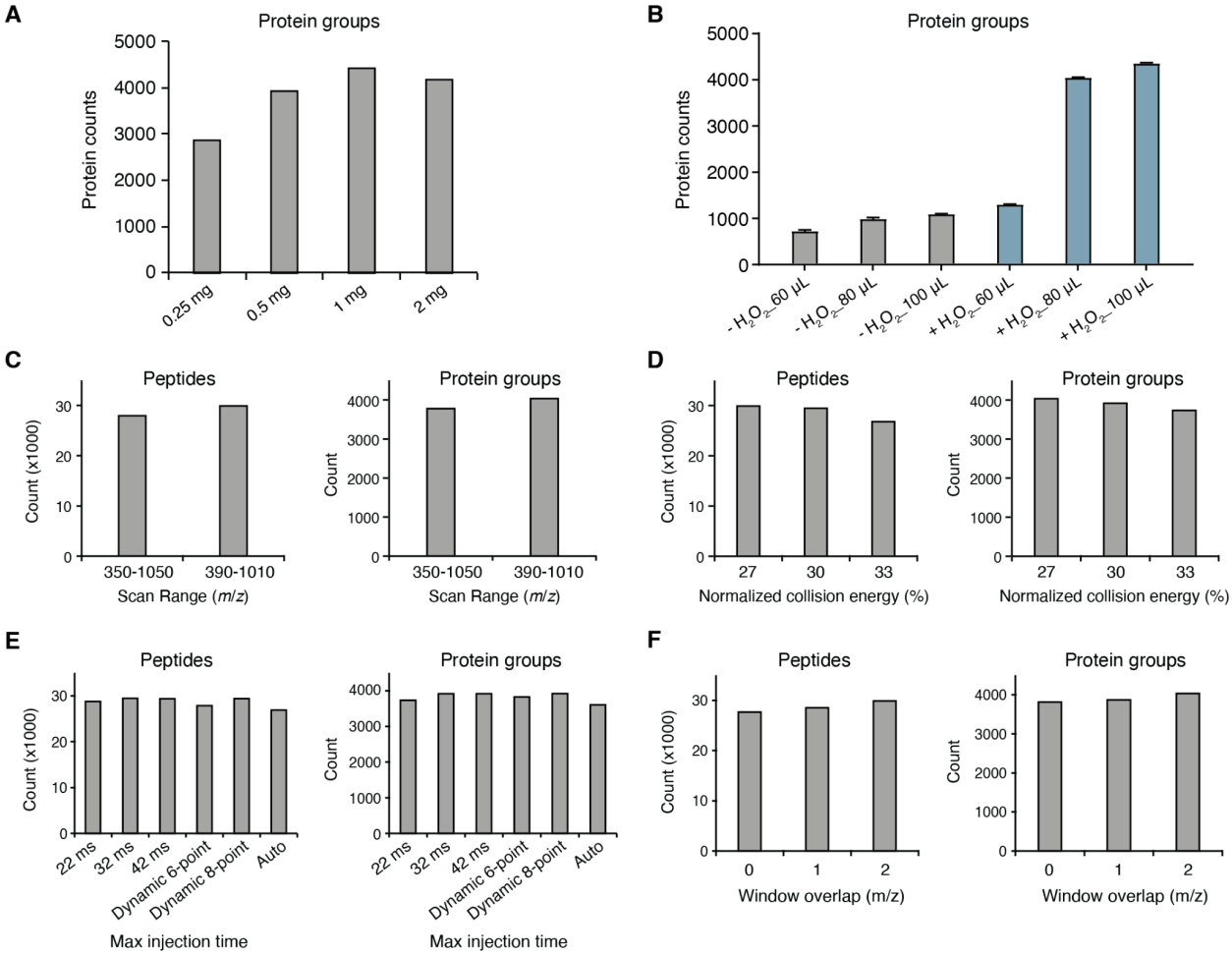


Figure S2. Sample enrichment and DIA-mass spectrometry methods optimization.

A. Different protein loading amounts with 100 µL beads. B. With vs. without H_2O_2 and different streptavidin beads amount with 1 mg protein. Three replicates for each condition. Optimization of DIA-mass spectrometry parameters including: C. Scan range. D. Normalized collision energy. E. Maximum injection time. F. Width of overlapping window.

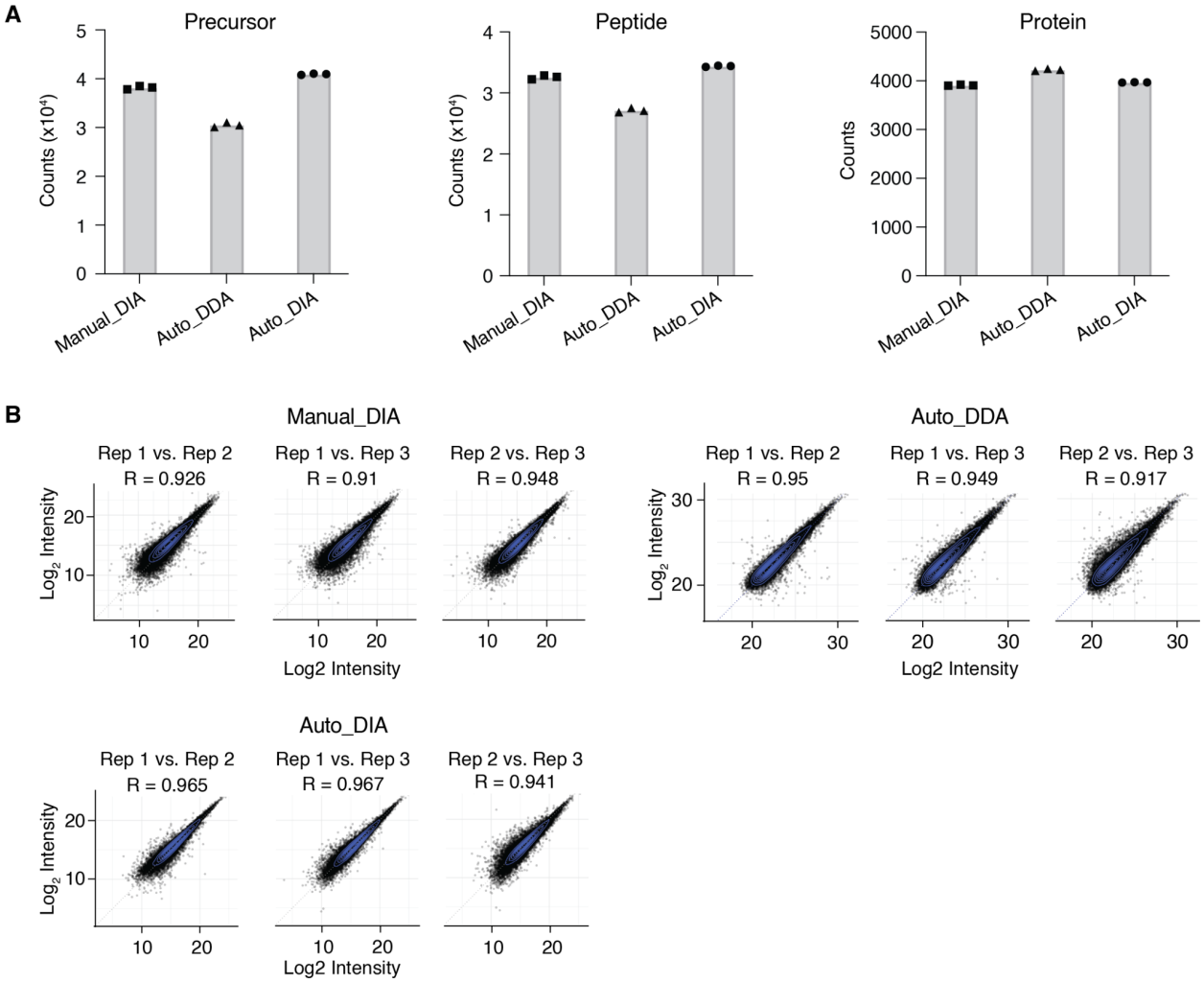


Figure S3. Comparison of manual- and automated- enrichment methods coupled to DDA- and DIA-mass spectrometry methods.

A. Identified numbers of precursors, peptides, and proteins. B. Correlation of 3 replicates using different methods.

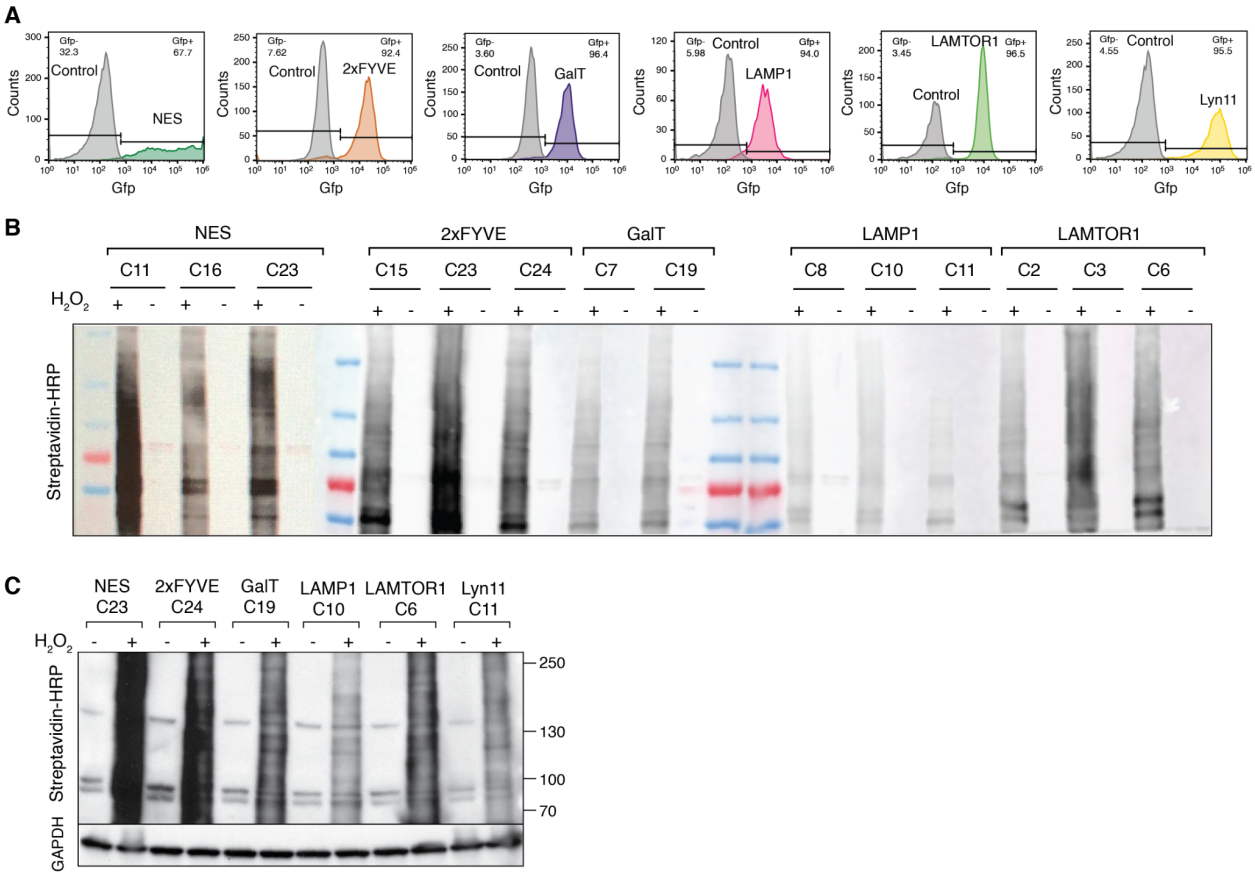


Figure S4. Characterization of all selected monoclonal APEX2-tagged location domain cell lines.

A. Flow cytometry analysis of APEX2-tagged location domain cell lines. GFP was used to evaluate expression and localization of the APEX2 construct. The non-Doxycyline induced cell line was used as control. B. Western blot analysis of whole cell lysate derived from monoclonal APEX2 cell lines. C. Western blot analysis of final monoclonal APEX2-tagged location domain cell lines with or without H_2O_2.

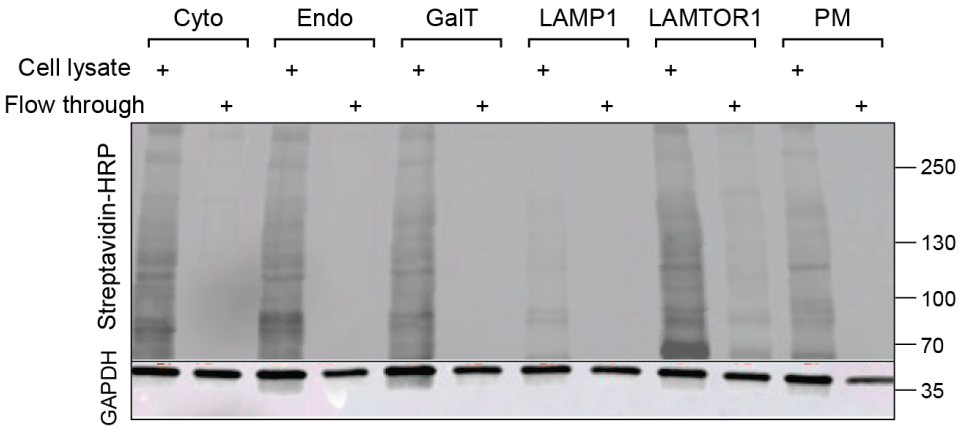


Figure S5. APEX-labeling proteomics quality control.

Western blot analysis of biotinylation and enrichment efficiency of 6 monoclonal APEX2-tagged location domain cell lines.

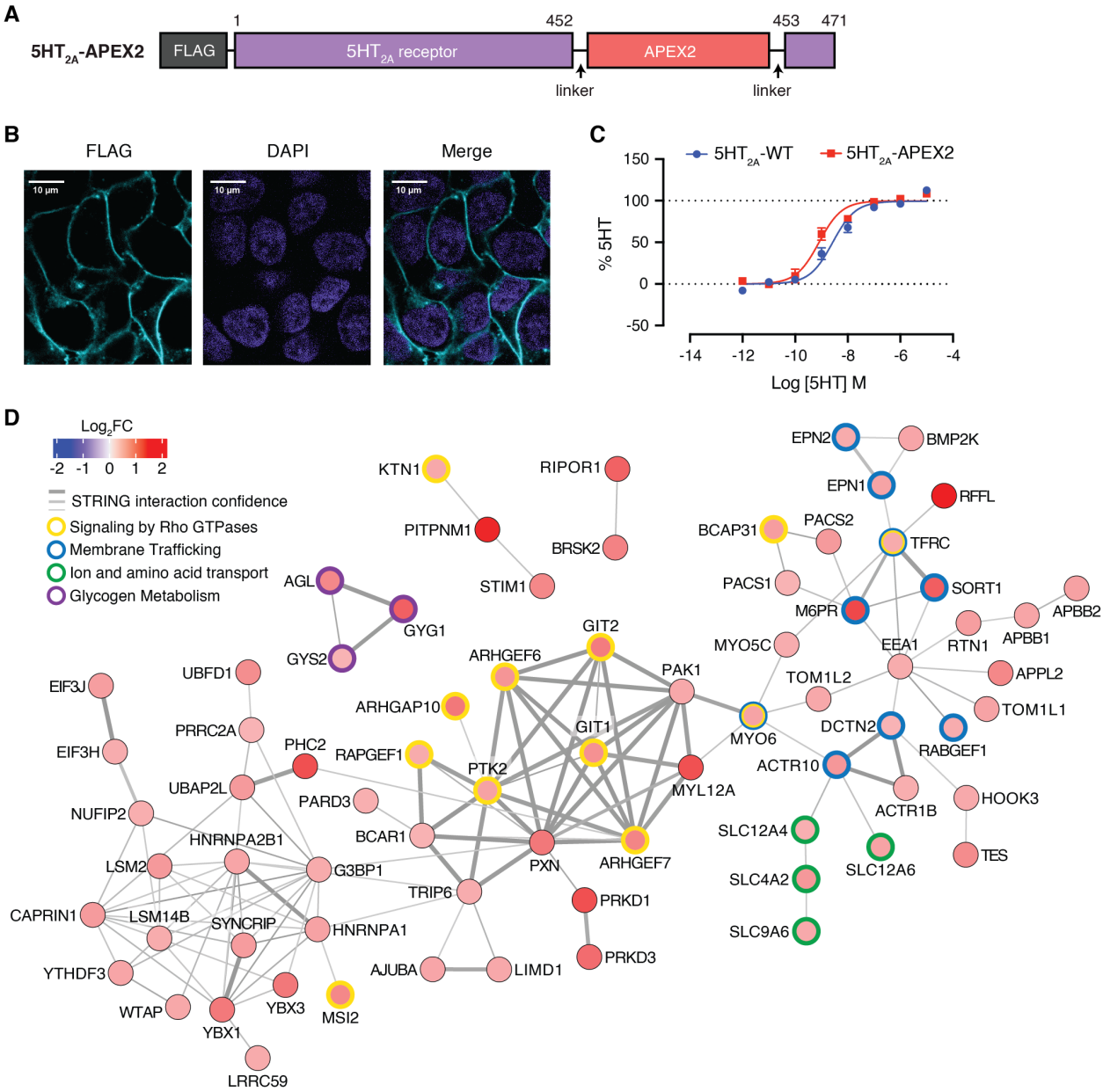


Figure S6. Mapping ligand-dependent proximal interaction network changes of 5HT$_{2A}$.

A. Construct design of 5HT$_{2A}$ receptor.

B. Confocal imaging of monoclonal 5HT$_{2A}$-APEX2 cell lines. Cells were stained with DAPI for the nucleus and the APEX2 construct location was indicated by FLAG. Scale bar represents 10 µm.

C. BRET validation of 5HT mediated Gq recruitment to APEX2-tagged vs. wild-type 5HT$_{2A}$.

D. Protein interaction network connecting proteins with sustained agonist-dependent changes in the proximity of 5HT$_{2A}$. Proteins shown as nodes were colored according to their $\log_2$ fold change. The edges connecting the proteins were derived from STRING (Szklarczyk *et al*, 2011) and the edge width was scaled according to the interaction confidence. Proteins corresponding

to Reactome pathways (Jassal *et al*, 2020) that were enriched within the sustained cluster were indicated with different colored node borders.

Supplemental Tables

Table S1. Table of plasmids used in this study.

Plasmid name	Plasmid vector	Promoter	Features	Details	Used for
Cyto-APEX2	pLVX	Tet-On	EcoRI - EGFP - 17aa linker - Flag tag - APEX2 - NES sequence - Stop - BamHI	17aa linker: SGLRTRAQASNSAAATM Flag tag: DYKDDDDK NES sequence: LQLPPLERLTLD	Stable expression in cytoplasm
Endo-APEX2	pLVX	Tet-On	EcoRI - EGFP - 16aa linker - 2xFYVE - 18aa linker - Flag tag - APEX2 - NES sequence - Stop - BamHI	16aa linker: YSDLELKLRIEFESDA 18aa linker: YVDGTAGPGSTGSAAATM Flag tag: DYKDDDDK NES sequence: LQLPPLERLTLD	Stable expression in endosome
GALT-APEX2	pLVX	Tet-On	EcoRI - APEX2 - NES sequence - 9aa linker - GalT - 7aa linker - EGFP - Stop - BamHI	NES sequence: LQLPPLERLTLD 9aa linker: DVELKLAAT 7aa linker: KDPPVAT	Stable expression in Golgi

LAMP1-APEX2	pLVX	Tet-On	EcoRI - LAMP1 - 8aa linker - EGFP - 7aa linker - Flag tag - APEX2 - NES sequence - Stop - BamHI	8aa linker: GGSGGSGG 7aa linker: SGAAATM Flag tag: DYKDDDDK	Stable expression in lysosome
LAMTOR1-APEX2	pLVX	Tet-On	EcoRI - LAMTOR1 - 8aa linker - EGFP - 7aa linker - Flag tag - APEX2 - NES sequence - Stop - BamHI	8aa linker: GGSGGSGG 7aa linker: SGAAATM Flag tag: DYKDDDDK	Stable expression in lysosome
PM-APEX2	pLVX	Tet-On	EcoRI - Lyn11 - 10aa linker - EGFP - 17aa linker - Flag tag - APEX2 - NES sequence - Stop - BamHI	10aa linker: DSAGSAGSAG 17aa linker: SGLRTRAQASNSAAATM Flag tag: DYKDDDDK NES sequence: LQLPPLERLTLD	Stable expression in plasma membrane
$5HT_{2A}$-APEX2	pcDNA3	CMV	EcoRI - Signal sequence - Flag tag - $5HT_{2A}$(1-452) - 8aa linker - APEX2 - 8aa linker - $5HT_{2A}$ (453-471) - Stop - BamHI	Signal sequence: KTIIALSYIFCLVFA Flag-tag: DYKDDDDAS 8aa linker:GGSGGGSG	Stable expression in whole cell

Table S2. Automated enrichment protocol for biotinylated proteins implemented on the KingFisher Flex platform.

Table S3. Performance comparison of the automated PL strategy combined with DIA-based MS.

Details of the venn diagrams depicted in Figure 3D.

	Rep1	**Rep2**	**Rep3**	Rep1∩Rep2	Rep2∩Rep3	Rep3∩Rep1	Rep1∩Rep2∩Rep3
Manual_DIA	163	49	55	502	959	210	36992 (95%)
Auto_DDA	604	946	954	2373	1556	2155	31743 (78.7%)
Auto_DIA	45	74	55	489	217	313	40211 (97.1%)

Table S4. Subcellular proteomes.

The table lists proteins quantified across all subcellular constructs and the statistical analysis from the comparison to the control (NES-APEX2). The location-specific proteins determined in this study are annotated as Lyn11 (plasma membrane), 2xFYVE (endosome), LAMP1 or LAMTOR1 (lysosome/late endosome), and GalT (Golgi), along with the proteins defined for respective compartments in Human Cell Map (Go *et al*, 2021).

Table S5. Proximal interaction networks of $5HT_{2A}$.

The table lists all proteins quantified in $5HT_{2A}$-APEX2 samples and the statistical analysis over the time course of 5HT activation. The proteins selected for transient cluster, sustained cluster, and location markers are annotated.

References

Adachi N, Hess DT, McLaughlin P & Stamler JS (2016) S-Palmitoylation of a Novel Site in the β2-Adrenergic Receptor Associated with a Novel Intracellular Itinerary. *J Biol Chem* 291: 20232–20246

Aebersold R & Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422: 198–207

Allen JA, Yadav PN & Roth BL (2008) Insights into the regulation of 5-HT2A serotonin receptors by scaffolding proteins and kinases. *Neuropharmacology* 55: 961–968 doi:10.1016/j.neuropharm.2008.06.048 [PREPRINT]

Antonicka H, Lin Z-Y, Janer A, Aaltonen MJ, Weraarpachai W, Gingras A-C & Shoubridge EA (2020) A High-Density Human Mitochondrial Proximity Interaction Network. *Cell Metab* 32: 479–497.e9

Bhatnagar A, Willins DL, Gray JA, Woods J, Benovic JL & Roth BL (2001) The dynamin-dependent, arrestin-independent internalization of 5-hydroxytryptamine 2A (5-HT2A) serotonin receptors reveals differential sorting of arrestins and 5-HT2A receptors during endocytosis. *J Biol Chem* 276: 8269–8277

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N & Ting AY (2018) Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* 36: 880–887

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR & Ma'ayan A (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14: 128

Cho H & Kehrl JH (2007) Localization of Gi alpha proteins in the centrosomes and at the midbody: implication for their role in cell division. *J Cell Biol* 178: 245–255

Choi M, Chang C-Y, Clough T, Broudy D, Killeen T, MacLean B & Vitek O (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* 30: 2524–2526 doi:10.1093/bioinformatics/btu305 [PREPRINT]

Collins BC, Hunter CL, Liu Y, Schilling B, Rosenberger G, Bader SL, Chan DW, Gibson BW, Gingras A-C, Held JM, *et al* (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun* 8: 291

Deutsch EW, Bandeira N, Perez-Riverol Y, Sharma V, Carver JJ, Mendoza L, Kundu DJ, Wang S, Bandla C, Kamatchinathan S, *et al* (2023) The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res* 51: D1539–D1548

DiBerto JF, Olsen RHJ & Roth BL (2022) TRUPATH: An Open-Source Biosensor Platform for Interrogating the GPCR Transducerome. *Bioluminescence*: 185–195

Eisenberg-Bord M, Shai N, Schuldiner M & Bohnert M (2016) A Tether Is a Tether Is a Tether:

Tethering at Membrane Contact Sites. *Developmental Cell* 39: 395–409 doi:10.1016/j.devcel.2016.10.022 [PREPRINT]

Gao XD, Tu L-C, Mir A, Rodriguez T, Ding Y, Leszyk J, Dekker J, Shaffer SA, Zhu LJ, Wolfe SA, *et al* (2018) C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. *Nat Methods* 15: 433–436

Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, Griss J, Sevilla C, Matthews L, Gong C, *et al* (2022) The reactome pathway knowledgebase 2022. *Nucleic Acids Res* 50: D687–D692

Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R & Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11: O111.016717

Go CD, Knight JDR, Rajasekharan A, Rathod B, Hesketh GG, Abe KT, Youn J-Y, Samavarchi-Tehrani P, Zhang H, Zhu LY, *et al* (2021) A proximity-dependent biotinylation map of a human cell. *Nature* 595: 120–124

Göös H, Kinnunen M, Salokas K, Tan Z, Liu X, Yadav L, Zhang Q, Wei G-H & Varjosalo M (2022) Human transcription factor protein interaction networks. *Nat Commun* 13: 766

Hu A, Noble WS & Wolf-Yadlin A (2016) Technical advances in proteomics: new developments in data-independent acquisition. *F1000Res* 5

Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, *et al* (2020) The reactome pathway knowledgebase. *Nucleic Acids Res* 48: D498–D503

Kim K, Che T, Panova O, DiBerto JF, Lyu J, Krumm BE, Wacker D, Robertson MJ, Seven AB, Nichols DE, *et al* (2020) Structure of a Hallucinogen-Activated Gq-Coupled 5-HT2A Serotonin Receptor. *Cell* 182: 1574–1588.e19

Kubitz L, Bitsch S, Zhao X, Schmitt K, Deweid L, Roehrig A, Barazzone EC, Valerius O, Kolmar H & Béthune J (2022) Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Commun Biol* 5: 657

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, *et al* (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44: W90–7

Kwan AC, Olson DE, Preller KH & Roth BL (2022) The neural basis of psychedelic action. *Nat Neurosci* 25: 1407–1419

Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, Mootha VK & Ting AY (2015) Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat Methods* 12: 51–54

Li X-W, Rees JS, Xue P, Zhang H, Hamaia SW, Sanderson B, Funk PE, Farndale RW, Lilley KS, Perrett S, *et al* (2014) New insights into the DT40 B cell receptor cluster using a proteomic proximity labeling assay. *J Biol Chem* 289: 14434–14447

Lobingier BT, Hüttenhain R, Eichel K, Miller KB, Ting AY, von Zastrow M & Krogan NJ (2017) An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell* 169: 350–360.e12

Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC & Aebersold R (2018) Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* 14: e8126

Martell JD, Yamagata M, Deerinck TJ, Phan S, Kwa CG, Ellisman MH, Sanes JR & Ting AY (2016) A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. *Nat Biotechnol* 34: 774–780

McClure-Begley TD & Roth BL (2022) The promises and perils of psychedelic pharmacology for psychiatry. *Nat Rev Drug Discov* 21: 463–473

Michalski A, Cox J & Mann M (2011) More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC−MS/MS. *Journal of Proteome Research* 10: 1785–1793 doi:10.1021/pr101060v [PREPRINT]

Myers SA, Wright J, Peckner R, Kalish BT, Zhang F & Carr SA (2018) Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat Methods* 15: 437–439

Olsen RHJ, DiBerto JF, English JG, Glaudin AM, Krumm BE, Slocum ST, Che T, Gavin AC, McCorvy JD, Roth BL, *et al* (2020) TRUPATH, an open-source biosensor platform for interrogating the GPCR transducerome. *Nat Chem Biol* 16: 841–849

Olsten MEK, Canton DA, Zhang C, Walton PA & Litchfield DW (2004) The Pleckstrin homology domain of CK2 interacting protein-1 is required for interactions and recruitment of protein kinase CK2 to the plasma membrane. *J Biol Chem* 279: 42114–42127

Paek J, Kalocsay M, Staus DP, Wingler L, Pascolutti R, Paulo JA, Gygi SP & Kruse AC (2017) Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. *Cell* 169: 338–349.e11

Piette BL, Alerasool N, Lin Z-Y, Lacoste J, Lam MHY, Qian WW, Tran S, Larsen B, Campos E, Peng J, *et al* (2021) Comprehensive interactome profiling of the human Hsp70 network highlights functional differentiation of J domains. *Mol Cell* 81: 2549–2565.e8

Qin W, Cho KF, Cavanagh PE & Ting AY (2021a) Deciphering molecular interactions by proximity labeling. *Nat Methods* 18: 133–143

Qin W, Myers SA, Carey DK, Carr SA & Ting AY (2021b) Spatiotemporally-resolved mapping of RNA binding proteins via functional proximity labeling reveals a mitochondrial mRNA anchor promoting stress recovery. *Nat Commun* 12: 4980

Ramanathan M, Majzoub K, Rao DS, Neela PH, Zarnegar BJ, Mondal S, Roth JG, Gai H, Kovalski JR, Siprashvili Z, *et al* (2018) RNA-protein interaction detection in living cells. *Nat Methods* 15: 207–212

Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA & Ting AY (2013) Proteomic

mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339: 1328–1331

Roth BL & Chuang DM (1987) Multiple mechanisms of serotonergic signal transduction. *Life Sci* 41: 1051–1064

Roux KJ, Kim DI, Raida M & Burke B (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196: 801–810

Samavarchi-Tehrani P, Abdouni H, Samson R & Gingras A-C (2018) A Versatile Lentiviral Delivery Toolkit for Proximity-dependent Biotinylation in Diverse Cell Types. *Mol Cell Proteomics* 17: 2256–2269

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

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, *et al* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9: 676–682

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B & Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498–2504

Short B, Preisinger C, Körner R, Kopajtich R, Byron O & Barr FA (2001) A GRASP55-rab2 effector complex linking Golgi structure to membrane traffic. *J Cell Biol* 155: 877–883

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, *et al* (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39: D561–8

Tsai T-H, Choi M, Banfai B, Liu Y, MacLean BX, Dunkley T & Vitek O (2020) Selection of Features with Consistent Profiles Improves Relative Protein Quantification in Mass Spectrometry Experiments. *Mol Cell Proteomics* 19: 944–959

Villani M, Subathra M, Im Y-B, Choi Y, Signorelli P, Del Poeta M & Luberto C (2008) Sphingomyelin synthases regulate production of diacylglycerol at the Golgi. *Biochem J* 414: 31–41

Xia Z, Gray JA, Compton-Toth BA & Roth BL (2003) A direct interaction of PSD-95 with 5-HT2A serotonin receptors regulates receptor trafficking and signal transduction. *J Biol Chem* 278: 21901–21908

Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, *et al* (2021) Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* 1: e90

Yamamoto H, Tobisawa Y, Inubushi T, Irie F, Ohyama C & Yamaguchi Y (2017) A mammalian homolog of the zebrafish transmembrane protein 2 (TMEM2) is the long-sought-after cell-surface hyaluronidase. *J Biol Chem* 292: 7304–7313