

Proximity Labeling in Plants

Shou-Ling Xu,^{1,2} Ruben Shrestha,¹
Sumudu S. Karunadasa,¹ and Pei-Qiao Xie^{1,3}

¹Department of Plant Biology, Carnegie Institution for Science, Stanford, California, USA;
email: slxu@stanford.edu

²Carnegie Mass Spectrometry Facility, Carnegie Institution for Science, Stanford, California, USA

³Department of Molecular and Cell Biology, University of California, Berkeley, California, USA

Annu. Rev. Plant Biol. 2023. 74:285–312

First published as a Review in Advance on
February 28, 2023

The *Annual Review of Plant Biology* is online at
plant.annualreviews.org

<https://doi.org/10.1146/annurev-arplant-070522-052132>

Copyright © 2023 by the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See credit lines of images or other third-party material in this article for license information.

Keywords

proximity labeling, protein–protein interaction, protein–nucleotide interaction, subcellular proteome, cell type–specific proteome, quantification

Abstract

Proteins are workhorses in the cell; they form stable and more often dynamic, transient protein–protein interactions, assemblies, and networks and have an intimate interplay with DNA and RNA. These network interactions underlie fundamental biological processes and play essential roles in cellular function. The proximity-dependent biotinylation labeling approach combined with mass spectrometry (PL-MS) has recently emerged as a powerful technique to dissect the complex cellular network at the molecular level. In PL-MS, by fusing a genetically encoded proximity-labeling (PL) enzyme to a protein or a localization signal peptide, the enzyme is targeted to a protein complex of interest or to an organelle, allowing labeling of proximity proteins within a zoom radius. These biotinylated proteins can then be captured by streptavidin beads and identified and quantified by mass spectrometry. Recently engineered PL enzymes such as TurboID have a much-improved enzymatic activity, enabling spatiotemporal mapping with a dramatically increased signal-to-noise ratio. PL-MS has revolutionized the way we perform proteomics by overcoming several hurdles imposed by traditional technology, such as biochemical fractionation and affinity purification mass spectrometry. In this review, we focus on biotin ligase-based PL-MS applications that have been, or are likely to be, adopted by the plant field. We discuss the experimental designs and review the different choices for engineered biotin ligases, enrichment, and quantification strategies. Lastly, we review the validation and discuss future perspectives.

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

OPEN ACCESS 

Contents

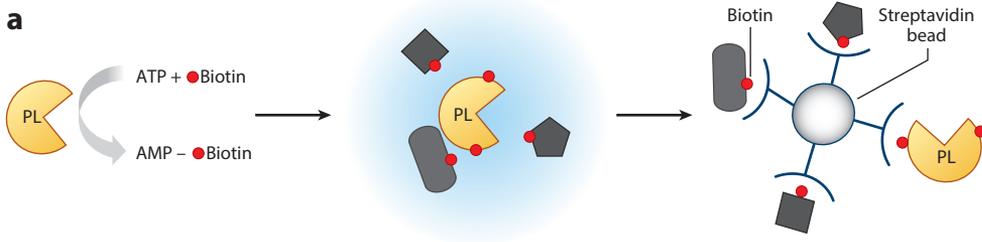
1. INTRODUCTION	286
2. PROXIMITY-LABELING MASS SPECTROMETRY APPLICATIONS TO STUDY PROTEIN-PROTEIN INTERACTIONS AND SUBCELLULAR/ORGANELLE AND CELL TYPE-SPECIFIC PROTEOMES	289
2.1. Using Proximity-Labeling Mass Spectrometry to Study Protein-Protein Interactions	289
2.2. Using Proximity-Labeling Mass Spectrometry to Study Subcellular Proteomes	290
2.3. Using Proximity Labeling to Study Cell Type-Specific Proteomes	292
3. MODULATED PROXIMITY-LABELING APPLICATIONS	292
3.1. Split Proximity Labeling to Map Conditional Protein-Protein Interactions and Localized Proteome	292
3.2. CRISPR-Driven Proximity-Labeling Mass Spectrometry to Study the Protein-Nucleotide Interactome	294
3.3. A Modular System Using a Conditionally Stable GBP-Proximity Labeling Method to Map Protein-Protein Interactions	295
4. CHOICES OF BIOTIN LIGASES	296
5. EFFECTIVE ENRICHMENT STRATEGIES FOR DETECTING BIOTINYLATED PROTEINS	299
5.1. Enrichment of Proximity-Labeled Proteins at the Protein Level	299
5.2. Enrichment of Proximity-Labeled Proteins at the Peptide Level	301
6. QUANTIFICATION STRATEGIES USED IN PROXIMITY LABELING-MASS SPECTROMETRY	301
6.1. Label-Free Quantification	301
6.2. Multiplexed Tandem Mass Tag Isobaric Labeling Quantification	303
6.3. ¹⁵ N Metabolic Labeling	303
6.4. Data Normalization	304
7. VALIDATION OF CANDIDATES FROM PROXIMITY LABELING-MASS SPECTROMETRY EXPERIMENTS	304
8. FUTURE PROSPECTS	304

1. INTRODUCTION

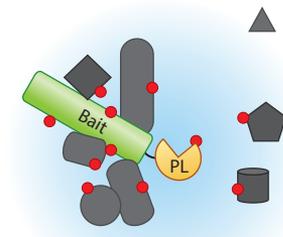
Plant cells have unique structural and functional features that evolved in response to environmental cues (59). Elucidating the complex cellular networks that underlie plant-specific functions at the molecular level is critical to understand plant growth or development and to dissect genotype-to-phenotype relationships. In particular, knowledge about protein abundance, localization, protein-protein interactions (PPIs), and the intimate interplay between protein and DNA/RNA will shed light on the functional and mechanical understanding of plants' unique molecular capabilities. However, obtaining such data remains challenging, especially for weak and transient interactions, low-abundant proteins, and hard-to-purify subcellular structures (97). In addition, significant proportions of genes and the proteins they encode in plants are yet to be functionally characterized (98).

Biotin ligase-based proximity-labeling (PL) technology has emerged as a powerful tool for studying complex cellular networks. In PL, a genetically encoded biotin ligase is targeted to a

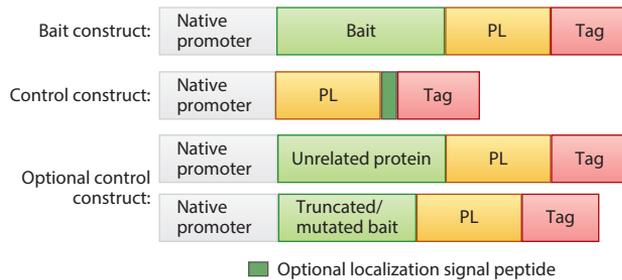
protein complex of interest or to an organelle by fusing to a protein or a localization signal peptide. In the presence of applied biotin, this engineered promiscuous biotin ligase converts biotin into short-lived, diffusible, and activated biotin adenylate intermediates, which are transferred to the ϵ -amino group of surface-exposed lysine residues of proximal proteins within a 10-nm radius (101) (**Figure 1a**). The biotinylated proteins are then enriched by streptavidin beads and subsequently



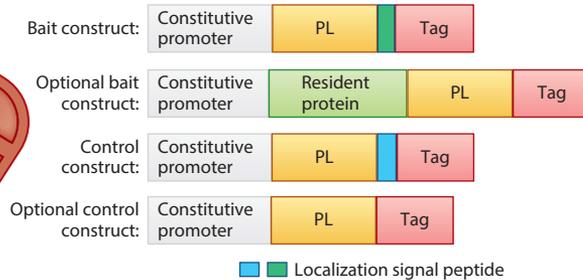
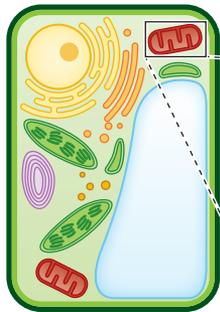
b **Principle**
Protein–protein interaction



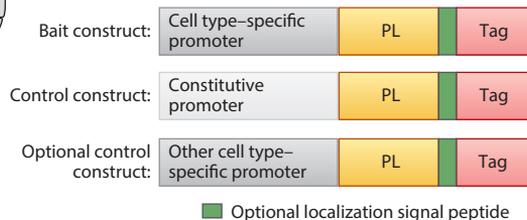
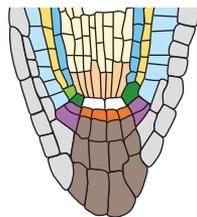
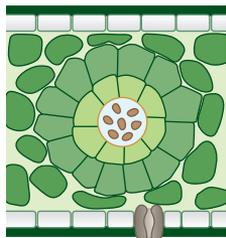
Suggested construct design



c Subcellular/organelle proteome



d Cell type–specific (subcellular) proteome



(Caption appears on following page)

Figure 1 (Figure appears on preceding page)

The general principle and application of PL in studying protein–protein interactions, subcellular and organelle proteomes, and cell type–specific proteomes. (a) Depiction of biotin ligase PL. The PL enzyme reacts with biotin and ATP to form a diffusible, reactive intermediate biotinoyl-5'-AMP that biotinylates proximal proteins. The blue region denotes the biotinylation radius. Biotinylated proteins are subsequently enriched using streptavidin beads followed by mass spectrometric analysis. (b–d) Principle (left) and suggested construct designs (right) of different types of PL-MS studies. (b) Depiction of PL-MS applications used to study protein–protein interactions. The PL enzyme is fused to the bait protein expressed at the endogenous level to target a protein complex. Upon the addition of biotin, the PL enzyme biotinylates proteins in its vicinity. (c) Depiction of a PL-MS application used to study subcellular/organelle proteomes. A schematic view of a plant cell is shown with its membrane-bound or membraneless compartments, exemplifying a PL application in the mitochondria matrix. (d) Depiction of PL-MS used to study cell type–specific (subcellular) proteomes. Example proteomes include bundle-sheath cells and mesophyll cells in *C4* plants and *Arabidopsis* root cells in the root tip. Tags (e.g., GFP/Myc/Flag) are often included to visualize protein localization and to assess protein expression levels. The positions of the tag, bait protein, PL enzyme, and localization signal peptide can be rearranged to allow the proper expression and function of constructs in PL. Abbreviations: AMP, adenosine monophosphate; ATP, adenosine triphosphate; GFP, green fluorescent protein; Myc, a peptide tag derived from the c-Myc protein; Flag, a short, hydrophilic protein tag; MS, mass spectrometry; PL, proximity labeling.

identified by mass spectrometry (MS). Multiple versions of biotin ligases have been engineered, such as BioID (BirA*) (55), BioID2, ultraID and microID2 (10, 49, 60), TurboID and miniTurboID (10), and AirID (54). In particular, the advent of TurboID (10) allows for highly efficient labeling in a timescale of minutes and a broader range of working temperatures. These improvements on PL have enabled proteomic mapping in a spatial and temporal manner and extended its use to many more organisms, including plants. Moreover, the recent coupling of PL with clustered regularly interspaced short palindromic repeats (CRISPR) has allowed more biological questions to be answered (78, 133).

PL has several major advantages. First, PL allows the tagging of proteins in a physiological context as the PL enzyme is genetically encoded and expressed *in vivo*. Maintaining intact protein complexes or organelles during lysis and purification is often challenging when using traditional methods such as affinity purification and biochemical fractionations. PL circumvents this hurdle by tagging both bait and prey proteins or by tagging constituent proteins within an organelle with biotin, then capturing them using the high-affinity streptavidin beads ($K_d = 10^{-14}$ mol/L) under very harsh conditions, including high concentrations of detergent, salt, or denaturing agents. Because of these major advantages over traditional methods, PL has been widely adopted to study PPIs, subcellular proteomes, cell type–specific proteomes, and protein–nucleotide interactions (29, 95, 120).

Another PL system, APEX or APEX2, engineered based on plant ascorbate peroxidases, has been mainly used in human cell lines and some dissected tissues (45, 63, 73, 96). The APEX system biotinylates proteins with the biotin-phenol substrate under oxidizing conditions (H_2O_2). This system prevails with its fast speed (the labeling can be done in 1 min) and its timely control upon H_2O_2 availability. In addition, the APEX system can also generate contrast in electron microscopy (73) and spatially map endogenous RNAs (32, 50). However, the APEX system has not been widely adopted in studies of plants due to low signal-to-noise ratio caused by the high background labeling from endogenous peroxidases. The poor membrane permeability of substrate biotin-phenol (which is much more hydrophobic than biotin) also restricts the use of APEX methods in complex tissues of plants (and animals). The stress response and potential toxicity induced by the substrate H_2O_2 further limit its use in plants. In this review, we focus on the biotin ligase–based PL system and its applications. A detailed review of the APEX system can be found in References 71 and 101.

CRISPR: clustered regularly interspaced short palindromic repeats

In this review, we cover the most common or adaptable PL applications and their experimental designs. We then discuss different biotin ligases and enrichment choices. We also explore different quantification strategies involving mass spectrometric data analysis and offer an overview of the validation methods of the candidates discovered from PL-MS experiments. Lastly, we discuss the future prospects of PL-MS studies in plants.

LC-MS/MS: liquid chromatography-coupled tandem mass spectrometry

2. PROXIMITY-LABELING MASS SPECTROMETRY APPLICATIONS TO STUDY PROTEIN-PROTEIN INTERACTIONS AND SUBCELLULAR/ORGANELLE AND CELL TYPE-SPECIFIC PROTEOMES

PL has emerged as a powerful technique, complementary to traditional methods (e.g., affinity purification and biochemical fractionation), to study protein interaction networks and to map subcellular and cell type-specific proteomes. In this section, we compare PL-MS to traditional methods, discuss the experimental designs, and demonstrate several application examples that can be potentially adapted by or directly applied to plant studies.

2.1. Using Proximity-Labeling Mass Spectrometry to Study Protein-Protein Interactions

PPIs underlie nearly every cellular process (106). Mapping PPIs in either protein assemblies or signaling pathways is essential for understanding the mechanisms of cellular processes and their roles in growth and development. While some of the interactions are stable, most of them are weak, transient, and restricted in a spatial and temporal manner. Accurate and comprehensive characterization of these dynamic interactions is critical in proteomic studies to annotate protein functions and to dissect the biological processes.

Affinity purification coupled with mass spectrometry (AP-MS) is most used in traditional methods for identifying PPIs. AP-MS has been used to map stable protein complexes successfully and, in some cases, to recover transient interactions using modified protocols (8, 42, 47, 81, 82). For instance, a quick and dirty procedure that includes quick binding and short and mild washes was performed to preserve the transient interactors (52). Alternatively, a low concentration of chemical cross-linkers was employed to detect the weak and transient interactions (111, 137). However, AP-MS experiments often result in the loss of weak and transient interactors during lysis and purification. Preserving protein interactions in AP-MS is even more challenging for less-soluble membrane-associated protein complexes or chromatin-associated protein complexes and for low-abundance protein complexes (39, 64).

PL-MS shows advantages over AP-MS in several aspects (39). Most prominently, the PL approach allows biotinylation labeling of both bait and prey in their native cellular environment such that the capture is independent of the affinity of the prey to the bait protein. Therefore, keeping the protein complexes in their native state during the affinity capture of biotinylated proteins is no longer necessary. Furthermore, compared to the affinity of antibody-mediated interactions used in most AP-MS such as the high-affinity tandem green fluorescent protein (GFP) nanobody ($K_d = 3.6 \times 10^{-11}$ mol/L) (34), a significantly higher affinity of streptavidin ($K_d = 10^{-14}$ mol/L) allows the capture of biotinylated proteins under very harsh conditions (with high concentrations of detergent, salt, or denaturing agents) with a much greater efficiency. These conditions allow the near-complete extraction of the proteome, which significantly improves the input for binding. The high affinity also allows thorough washes to reduce contaminants (e.g., nonspecific binding proteins) and to remove detergents used in the initial steps of extraction and binding. Removal of detergents is particularly important for downstream on-bead digestion and liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS) analysis as they are often contaminated with polymers that interfere with MS analysis.

A successful PL-MS experiment design for studying PPIs (**Figure 1b**) requires three critical parameters: (a) functional bait-fusion protein(s), (b) proper control(s), and (c) optimized labeling time. First, the bait protein is fused to a PL enzyme (**Figure 1b**) at either the N or C terminus and expressed under the bait protein's native promoter at the endogenous level. A functional complementation analysis is a prerequisite to determine whether the bait-PL fusion protein is functional in its physiological state. Second, the choice of control is experiment dependent and can include (a) the PL enzyme alone or (b) the PL enzyme fused to an optional localization peptide under the same native promoter. Optional control constructs may also include the PL enzyme fused to an unrelated protein that localizes to the same compartment as the bait or to a truncated or mutated bait. It is important to ascertain whether the control fusion protein(s) and the bait protein have similar expression levels and share a similar subcellular localization pattern. For instance, if the bait localizes to the nucleus, then the control should also target the nucleus. The controls will generate a filter list of proteins, including endogenously biotinylated proteins, stochastically labeled proteins, and bead-bound nonspecific proteins (72, 138). Filtering out the stochastically labeled proteins that reside by chance in proximity of the bait from bona fide interacting partners is of particular importance, since the PL enzymes cannot distinguish between these two types of proteins. In addition, extra tags, such as the fluorescence, Myc, and Flag tags, are often included to examine localization and expression of both bait and control. Lastly, control over the biotin labeling time is equally critical for generating high-quality data. Depending on the abundance of the bait proteins, the duration of labeling needs to be experimentally determined to obtain optimal results. For example, a shorter time of biotin treatment (≤ 1 h) is preferred to get a snapshot of PPIs for more abundant proteins such as some kinases (58) and splicing factors (R. Shrestha & S.-L. Xu, unpublished observations), while a longer biotin treatment (up to several hours) may be needed for less-abundant bait proteins (e.g., transcription factors) (72). In summary, the above three parameters should be easily met in well-established plant models. For plant species in which transgenesis and mutants are not available, researchers should at least examine the proper localization of the bait-PL fusion proteins and express them close to the endogenous level to obtain bona fide interactors.

PL-MS has been adopted in plants and extensively applied in studying PPIs and mapping protein networks. For example, PL-MS has enabled the mapping of low-abundant protein transcription factor FAMA interactomes (72), the hard-to-enrich plasma membrane protein complex (3), and the nuclear envelope membrane protein complex (44, 116). PL-MS has also been superior for dissecting weak, transient interactions of signaling protein complexes, such as kinase protein/kinase substrates (58) or E3 ligase/substrates (139). These applications of PL-MS not only confirmed known interactors but also discovered many novel protein interactors, demonstrating the efficacy and potential of PL over traditional approaches in elucidating protein networks in plants.

2.2. Using Proximity-Labeling Mass Spectrometry to Study Subcellular Proteomes

Protein functions are closely correlated with their localization within the cell. Eukaryotic cells contain highly specialized membrane-bound and membraneless compartments that are critical to regulate various molecular complexes and biological processes in living organisms. Mapping the subcellular/organelle proteomes has long been a major interest in molecular biology to study their distinct and/or overlapping functions in both plants and animals (1, 21).

Various techniques, including organelle-specific biochemical fractionation and cell-wide biochemical fractionation (e.g., correlation profiling), have been attempted to study the proteome

compositions of organelles and subcellular structures (21, 38). These methods take advantage of the differences in organelles' density, size, or membrane properties to separate organelles via multiple centrifugations or detergent steps. Using these approaches, plant proteomics studies have been conducted on numerous membrane-bound compartments, including the chloroplast (7), mitochondria (100, 107), peroxisome (86), vacuole (105), and nucleus (91), along with other subcellular structures such as the plasmodesma (33) and nucleolus (87). Proteomics studies have also probed proteomes of different membranes, including the plasma membrane (130), Golgi membrane (85), and chloroplast envelope membrane (110). These attempts have provided invaluable insights into subcellular proteomics, organelle biogenesis, and functionality analysis. However, these traditional proteomics methods often suffer contamination from nontarget organelles, incomplete coverage/undersampling, and inconsistent assignment of protein localization to subcellular fractions. In addition, challenges remain in studying subcellular compartments and structures including membraneless organelles, which cannot be easily separated by traditional biochemical fractionations.

PL-MS has unique advantages in studying subcellular proteomics (29, 71). By targeting the PL enzyme to a subcellular location, the constituents within that compartment or subcellular structure are biotin labeled in a physiological context. The enrichment of these biotin-labeled resident proteins in PL-MS is straightforward and highly efficient using the high-affinity streptavidin beads, bypassing the requirement to preserve the integrity of organelles/subcellular structures during sample purification as in the conventional biochemical fractionation method.

Experiment designs (**Figure 1c**) for PL subcellular proteomics include fusing the PL enzyme to well-validated sorting motifs or localization peptides. For instance, the PL enzyme can be fused to a nuclear localization signal (NLS) to target the nucleus, to a nuclear export signal (NES) to target the cytosol, or to a transit peptide (e.g., from the Rubisco small subunit) to target chloroplast stroma (79). Alternatively, the biotin ligases can be fused to known resident proteins from different subcellular locations (40), albeit with the potential risk of interfering with function or localizing to microdomains (18). To ascertain whether most constituents within the compartment of interest are labeled, the bait proteins are often expressed constitutively in every cell. Proper controls typically include stand-alone biotin ligases or biotin ligases fused to a different localization signal or to resident proteins from other subcellular locations.

The PL-MS approach has been extensively applied in animals to study subcellular structures (39). Recently, large-scale subcellular mapping using BioID fused to 192 resident markers localized to over 20 intracellular compartments has been performed in human HEK293 cells (40). Go et al. (40) exploited multiple baits that localize to the same subcellular compartment to generate orthogonal data sets and assigned over 4,000 intercellular proteins.

Similar large-scale subcellular profiling has not been attempted in plants yet. Recently, Mair et al. (72) probed the nuclear proteome in *Arabidopsis* using PL-MS and identified proteins that were significantly enriched in nuclear assignment. To profile different compartments/subcellular structures, plant biologists can design PL experiments based on published subcellular proteomes that have been categorized by the Subcellular Localization Database for *Arabidopsis* Proteins (*SUBA*; <https://suba.live>) (41, 43, 117). More importantly, the highly active TurboID opens more possibilities for further investigations of subcellular proteomes in a spatial and temporal manner and other compartments that have been historically challenging to study, including membraneless, phase-separated, cytosolic, and nuclear granules and cell wall-embedded plasmodesmata. By virtue of TurboID, plant biologists can also extend current subcellular proteome studies to a wider range of plant species, including woody plants and crops, to dissect functions of organelles across plants at a systems level.

2.3. Using Proximity Labeling to Study Cell Type-Specific Proteomes

To study the unique biological functions and properties of individual cell types, in-depth protein expression analysis in different types of cells, particularly specialized cells (e.g., guard cells or bundle-sheath cells), is fundamental for investigating the unique functions of tissues and organisms (80).

Various types of tools have been developed to study the cell type-specific proteomes in both animals and plants. These tools include laser capture microdissection or fluorescence-activated cell sorting (FACS) of labeled cells in a tissue (9, 25, 26, 36, 48, 84) and less common methods such as mechanical separation and sequential protoplast generation (114). Together, these approaches have enabled cell type-specific proteomics studies in multiple tissues, including anther pollen (140), root (90, 144), leaf (114), and trichome (5).

Generating sufficient data sets from traditional methods, however, comes at a cost of time and effort to attain sufficient quality and quantity of isolated cells. As a noninvasive, time-efficient, and highly sensitive complementary tool, PL-MS has gained traction in studying cell type-specific proteomes (72, 112). By expressing the PL enzyme in a specific cell type, proteins therein can be tagged and efficiently isolated, allowing the proteomic workflow to be executed at the tissue level with minimized labor and time. More importantly, PL methods circumvent the challenges of obtaining enough pure cells and the automation workflow required in traditional methods.

To design PL-MS experiments for studying the proteome of specific cell type(s), the PL enzyme is expressed under a cell type-specific promoter and, in some cases, fused to a localization signal to target specific subcellular compartments. The control can be a PL enzyme expressed under either a different cell type-specific promoter or a constitutive promoter. For example, to study the guard cell-specific nuclear proteome, Mair et al. (72) expressed TurboID-NLS under the guard cell-specific FAMA promoter and used a constitutive promoter for the same fusion construct as the control. PL-MS not only identified known guard cell-associated transcription factors but also discovered many additional components that were identified by previous proteomics studies using 300 million guard cell protoplasts (141, 142).

PL-MS can be extended to many more cell types that are of high interest to plant biologists, including different root cells and the bundle-sheath cells in C_4 plants (**Figure 1d**). Researchers can take advantage of cell type-specific promoters such as that of the nicotinamide adenine dinucleotide phosphate (NADP)-malic enzyme (bundle sheath-specific) and phosphoenolpyruvate (PEP) carboxylase (mesophyll-specific) (30, 128) to map the proteomes of these unique cell types and dissect their functions in the future.

3. MODULATED PROXIMITY-LABELING APPLICATIONS

Multiple modified PL approaches have been developed recently to tackle additional challenges and to increase the versatility of PL applications. In this section, we cover the split-PL system in studying conditional PPIs and hard-to-purify subcellular structures or compartments. We then discuss modified PL approaches that take advantage of CRISPR to study the protein-nucleotide interactome. Lastly, we will discuss a modular and versatile PL system controlled by a conditionally stable GFP-binding protein nanobody (GBP).

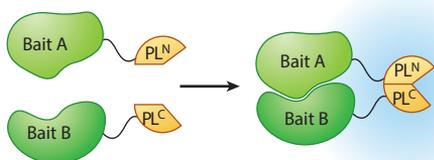
3.1. Split Proximity Labeling to Map Conditional Protein-Protein Interactions and Localized Proteome

Consisting of two inactive fragments of the biotin ligase (BioID or TurboID) that can be reconstituted through PPIs or organelle-organelle interactions, split-PL has enabled the study of

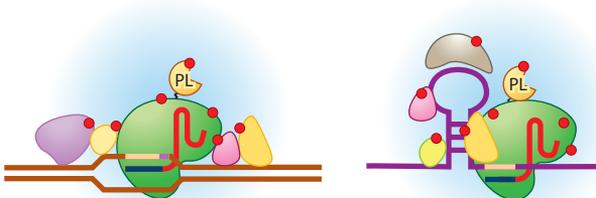
conditional interactomes and proteomes of subcellular structures that are difficult to access by conventional PL approaches (17, 27, 61, 104, 115) (**Figure 2a**).

In cells, an increasing number of proteins have been identified as moonlighting proteins that associate with multiple complexes. Deconvoluting their interactomes and assigning them to specific complexes will be key to dissecting their distinct and/or overlapping functions (46). Split-BioID has been applied to assign specific complexes to moonlighting proteins. For instance, Argonaute 2 (Ago2) is involved in microRNA (miRNA)-induced silencing and known to be part of at least two functionally distinct complexes, including the miRNA-induced silencing complex (miRISC)

a Split-TurboID



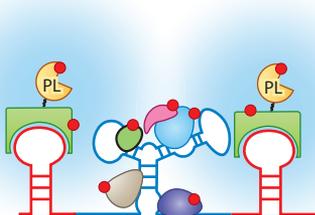
b dCAS9 gRNA-driven PL



Control: No gRNA cotransformation

■ Localization signal peptide

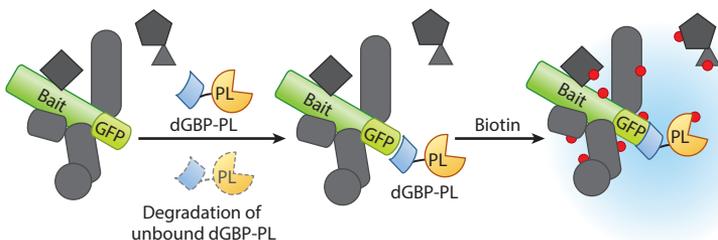
c BoxB/λN-directed PL



Control: No BoxB-RNA cotransformation

■ Localization signal peptide

d Conditionally stable anti-GFP nanobody-driven PL



Control: dGBP-PL alone

■ Optional localization signal peptide

(Caption appears on following page)

Figure 2 (Figure appears on preceding page)

Modulated PL applications. (a) Schematic of the split-PL approach and suggested design of the constructs. The PL enzyme is divided into two inactive fragments (PL^N and PL^C) that fail to bind unless brought together by either protein–protein interactions or organelle–organelle interactions to reconstitute biotinylation activity. Choice of promoters is experiment dependent (e.g., native promoters are preferred for conditional protein–protein interactions). (b) Diagram depicting CRISPR-driven PL for protein–DNA or protein–RNA interaction studies. The PL enzyme fused to the nuclease-deficient dCas9 protein (or dCas13 or dCasRx) is targeted to a specific genomic locus or RNA of interest via a gRNA and biotinylates proteins in its vicinity. (c) Scheme showing BoxB/ λ N-directed PL to detect the RNA-binding protein. Engineered RNA is composed of BoxB stem loops flanking the RNA motif of interest. The PL enzyme is fused to the λ N peptide and targeted to the engineered RNA via the λ N-mediated binding to the BoxB stem loop. Upon biotin treatment, the PL enzyme will biotinylate the RNA-binding proteins in its vicinity. Two BoxB stem loops are included in the engineered RNA to increase the sensitivity of biotinylation labeling. (d) Illustration of the conditionally stable dGBP-PL system. The dGBP-PL fusion protein transgenic lines can be crossed with existing transgenic GFP-tagged lines for PL studies. Proximity biotinylation is initiated when dGBP-PL is stabilized upon binding to the bait-GFP fusion protein in cells containing both transgenes. By contrast, the unbound dGBP-PL is unstable due to its rapid degradation by the ubiquitin proteasome system; thus, the background labeling is minimal. The dGBP-PL construct can contain an optional localization signal peptide and be expressed in the same tissue as the bait to obtain optimal results. Biotin is represented by red circles. The blue region denotes the biotinylation radius. Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; dGBP-PL, anti-GFP nanobody-driven PL; GBP, GFP-binding protein nanobody; GFP, green fluorescent protein; gRNA, guide RNA; PL, proximity labeling.

and the RISC-loading complex (RLC). Schopp et al. (104) generated constructs containing Ago2 paired with two different known binding partners, Dicer and TNRC6C from the miRISC and RLC, respectively (NBirA*-Dicer/CBirA*-Ago2 and NBirA*-TNRC6C/CBirA*-Ago2), and used NBirA*-GFP as the control. Split-BioID probed the proteomes of distinct functional complexes containing Ago2 and identified novel candidates. Similarly, split-BioID was employed to map interactomes for a specific heterodimeric protein phosphatase complex (27).

Split-PL has also been utilized to map some challenging localized proteomes (17, 61, 115) with greater specificity in the targeting of biotinylated locations. Some localized proteomes (e.g., endoplasmic reticulum–mitochondrial contact sites) are very hard to map selectively because available characterized resident proteins also localize to other regions, resulting in nonspecific and convoluted mapping when fused to full-length PL. Meanwhile, traditional methods such as sequential centrifugation often suffer from contamination from other organelles and lack reproducibility. By combining split-PL with the chemically inducible FRB-FKBP dimerization system, this contact site has been mapped with a higher specificity (17, 61). Similarly, split-PL has been adopted to study hard-to-reach astrocyte–neuron junctions (115).

In the future, Split-PL can be easily adapted to the plant system to study conditional PPIs and organelles that are yet to be mapped, including the photobody (126) and plasmodesmata (65). In addition, split-PL can be combined with an on/off switch to control targeted proteome mapping more tightly. For instance, split-PL can be combined with a light switch to probe the photobody compositions using the phytochrome–phytochrome-interacting factor (PIF) protein systems, given that phytochrome and PIF proteins move to the photobody upon their light-dependent interaction. A caveat of the split-PL system, however, is its lower kinetics than those of the full-length PL. For instance, Cho et al. (17) showed that split-TurboID required 4 h of biotin treatment to obtain comparable results with those of TurboID using 1 min of biotin treatment.

3.2. CRISPR-Driven Proximity-Labeling Mass Spectrometry to Study the Protein–Nucleotide Interactome

Coupling CRISPR technology (143) with PL-MS has recently emerged as a valuable complementary approach to comprehensively map local chromatin interactions and architecture in situ (37, 78, 103). Gene expression is controlled at specific regions of genomic DNA by the concerted action of myriad transcriptional regulatory proteins, which either activate or repress

the transcription of target genes. Tremendous effort has been made to map these transcriptional regulatory complexes on a particular chromatin locus of interest to study local chromatin composition and architecture. Chromatin immunoprecipitation (ChIP) (35), a conventional approach, has provided information on both single-locus and genome-wide distribution of transcription factors but is often limited to examining a single previously defined transcription factor in each experiment. Chemical cross-linking and precipitation with complementary DNA probes have been employed to directly analyze chromatin complexes but often suffer from the loss of cellular and/or chromatin context (28, 53).

PL-MS shows great potential in mapping protein–DNA interactomes. Guided by single-guide RNAs (sgRNAs), a nuclease-deficient dCas9-fused biotin ligase targets a specific genomic locus (**Figure 2b**) and biotinylates proteins in close proximity to the target locus (103). To ensure the specificity of dCas9-PL, ChIP–quantitative real-time polymerase chain reaction (qPCR) is a necessary quality-control step to examine the target locus enrichment by pulling down either dCas9-PL or biotinylated proteins followed by qPCR. The control experiment contains no sgRNA co-transformation for generating a filter-out list consisting of off-target chromatin interactors and other usual contaminants (e.g., nonspecific binding proteins and endogenous biotinylated proteins). Similar approaches using APEX enzymes also successfully mapped the subnuclear proteomic landscapes to predefined genomic loci (37, 78).

The CRISPR-driven PL approach can be easily adapted to plant systems to study chromatin interactors. Several reporter lines, such as the stably transgenic DR5 and TCSn reporter lines used for auxin and cytokinin responses, respectively (125, 145), can be utilized to test their efficacy; these reporters contain tandem regulatory elements that can boost the detection sensitivity. CRISPR-PL and specific gRNAs can be designed to probe the chromatin interactors at these loci, and we would expect to detect AUXIN RESPONSE TRANSCRIPTION FACTORS (ARFs) or type B *Arabidopsis* response regulators (ARRs). This CRISPR-driven PL-MS approach can be used to study the regulatory components of multiple plant genes of interest that are subjected to tight controls in response to different biotic or abiotic responses. However, whether all the identified chromatin interactors from PL-MS are specific to the genomic elements remains unknown but can be determined in the future with the use of large collections of data.

Using the same principle, CRISPR-coupled PL-MS has also been adapted to map the protein–RNA interactomes (**Figure 2b**). Yi et al. (133) developed the CRISPR-assisted RNA–protein interaction detection (CARPID) method to probe binding proteins of specific long noncoding RNA (lncRNA) X-inactive-specific transcript (XIST) in a native cellular context. To examine the targeting specificity guided by sgRNA, a functional CasRx fused to biotin ligase was utilized to confirm the cleavage in target RNA. This CARPID approach identified not only known validated XIST interactors but also many more novel interacting proteins (133).

Alternatively, to probe RNA-interacting proteins, some researchers have utilized BoxB- λ N or MS2-MCP systems to target PL enzymes to an engineered RNA of interest fused with a BoxB stem loop (or MS2 stem loop) to induce biotinylation of RNA-binding proteins (77, 92, 135) (**Figure 2c**). Such methods map proteins that bind to engineered RNA but not those that bind to endogenous RNA but may have advantages for studying RNA-binding proteins in more limited regions of RNA.

3.3. A Modular System Using a Conditionally Stable GBP-Proximity Labeling Method to Map Protein–Protein Interactions

To enhance the versatility of the PL system, a modular system has been developed that can take advantage of existing transgenic GFP lines instead of generating transgenic lines containing each

DR5: auxin-responsive promoter

TCSn: two-component signaling sensor/cytokinin-responsive promoter

XIST: X-inactive-specific transcript

λ N: a 22-amino acid RNA-binding domain of λ bacteriophage antiterminator protein N

BoxB: λ N's specific 19-nucleotide binding site

MS2: a stem-loop structure from the phage genome

MCP: MS2 bacteriophage coat protein that binds to the MS2 stem loop

bait-PL fusion (and control) for each protein of interest. To screen for Cavin-associated networks, the Hall group (129) combined TurboID with a conditionally stable GFP-binding protein [destabilized GBP (dGBP), a 14-KDa nanobody] to create a destructive GFP-directed PL line in zebrafish and outcrossed this line to existing transgenic lines that contain native promoter-driven Cavin bait proteins fused with a GFP variant. Notably, the dGBP-PL fusion protein is nearly undetectable due to its rapid degradation via the ubiquitin proteasome system, thereby generating minimal background labeling within target cells in vivo. Biotinylation only starts when dGBP-PL stabilizes upon binding to the bait-GFP fusion in cells expressing both bait-GFP and dGBP-PL (129) (**Figure 2d**). This system enhances the versatility of PL applications by allowing the proteomic mapping between different bait proteins with available GFP fusion transgenic lines.

It is tempting to adapt this system to plants for proteomic mapping between different proteins, given that many transgenic GFP lines are available within the plant community. Whether the conditionally stable nanobody, when not bound by the antigen, is also rapidly degraded by the ubiquitin proteasome system in plants remains to be tested. We predict, however, that it will work in a similar manner since multiple nanobodies have been tested in the plant system (127). To achieve optimal results, multiple transgenic lines containing dGBP-TurboID (or other biotin ligases) fused with different localization signals should be created and crossed with existing GFP-tagged lines that have similar subcellular localizations. The promoter used for dGBP-TurboID may also be optimized to achieve desirable results.

4. CHOICES OF BIOTIN LIGASES

We group the major available engineered biotin ligases based on their origins (**Figure 3**): (a) BioID, TurboID, and miniTurbo, derived from *Escherichia coli* BirA (10, 55, 99) (**Figure 3a**); (b) BioID2, ultraID, and microID2, derived from *Aquifex aeolicus* BirA (10, 49, 60) (**Figure 3b**); (c) AirID, a synthetic BirA* designed de novo using ancestral enzyme reconstruction and site-directed mutagenesis (54) (**Figure 3c**); and (d) BASU, derived from *Bacillus subtilis* (92) (**Figure 3d**). Mutation at the conserved residue arginine (R) in the catalytic domain is shared among all versions of all the engineered biotin ligases and is most critical as it engenders a 100-fold greater dissociation constant for biotin and an approximately 400-fold higher dissociation rate for the reactive biotin intermediate (biotinoyl-5'-AMP) compared to wild-type BirA (62). When biotinoyl-5'-AMP is no longer retained, the biotin ligases become promiscuous and label proteins in the vicinity (19). **Supplemental Table 1** summarizes the labeling time and working temperatures for each biotin ligase, tested working organelles in both animals and plants, and tested plant species. Significant engineering effort has been devoted to improving the catalytic efficiency and activity of the biotin ligases, reducing their size, and minimizing the background labeling.

As the first developed biotin ligase with promiscuous protein-labeling activity, BioID (BirA*) has been extensively applied in studies in a variety of subcellular organelles and in protein network mapping (13, 20, 40) (**Supplemental Table 1**). However, with an optimum working temperature at 37°C, BioID has a slow labeling kinetics with 18–24 h or more labeling durations, which affects its temporal resolution, as BioID labels all the proteins in proximity over time rather than giving a snapshot of protein interactors. BioID has a limited applicability in model organisms, for example, flies, worms, and plants (10, 72).

TurboID, engineered via a directed evolution of BioID by the Ting group (72), has the highest activity of all biotin ligases and the broadest working temperature for model species (3, 10). With its rapid labeling kinetics in a time scale of minutes, TurboID is suitable for studying spatial and temporal cellular dynamics and subcellular or organelle proteomes (3, 72, 94, 113). The drawback of TurboID, however, is its high background labeling from endogenous biotin due to its higher

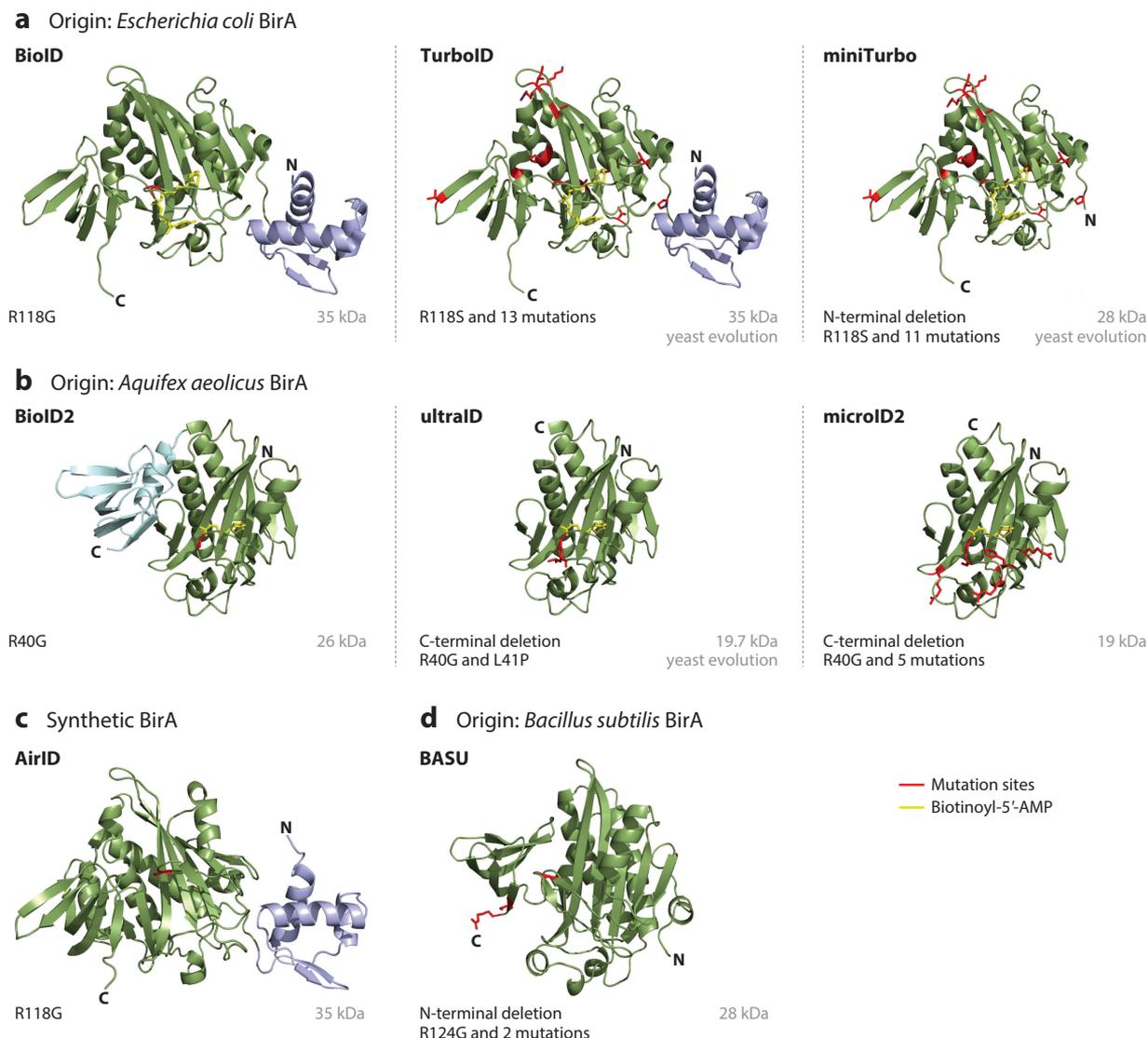


Figure 3

Choices of engineered biotin ligases. (a) Structure of biotin ligases engineered from *Escherichia coli* BirA, including BioID (BirA*) with an R118G mutation, TurboID with R118S and 13 additional mutations, and miniTurbo with R118S and 11 additional mutations and a truncated N terminus. (b) Structure of biotin ligase variants containing an R40G mutation engineered from *Aquifex aeolicus* BirA, including BioID2, ultraID with a truncated C terminus and an additional L41P mutation, and microID2 with both a C-terminal deletion and 5 additional mutations. This family differs from *E. coli* BirA due to the absence of the N-terminal domain. (c) Structure of AirID, a synthetic version of BirA with R118G mutation using ancestral enzyme reconstruction and site-directed mutagenesis. (d) Structure of BASU engineered from *Bacillus subtilis* BirA with R124G and 2 additional mutations as well as a truncated N terminus. Mutation sites are highlighted in red, and the biotinoyl-5'-AMP is depicted in yellow. Structures in panels a and b were generated based on Protein Data Base structures 2EWN and 3EFS, respectively; structures in panels c and d were generated by AlphaFold; visualization of structures was done in PyMol.

sensitivity to biotin. To address this issue for applications that specifically need lower background labeling, the Ting group (72) also engineered miniTurbo, which allows for more control over background labeling. Hence, miniTurbo would potentially be beneficial in studying plant tissues that have higher endogenous biotin levels than average (e.g., mature green siliques) (108). The overall activity of miniTurbo, however, is lower than that of TurboID and is less stable in some experimental systems (10, 129).

The second group includes BioID2, ultraID, and microID2 (**Figure 3b**), all of which were engineered from a biotin ligase of the thermophilic bacterium *A. aeolicus* (49, 57, 60). This family differs from *E. coli* BirA due to the absence of the N-terminal domain (**Figure 3b**). BioID2 has been utilized in mapping protein networks and subcellular proteomics but suffers from slow labeling kinetics (4, 11, 16, 57, 69, 134). Recently, a C-terminal truncation of BioID2 combined with directed evolution was used to create ultraID (60). Similarly, a C-terminal truncation of BioID combined with site-directed mutagenesis was used to create microID2 (49). Both ultraID and microID have higher activity levels than BioID2. Particularly, the smaller size of ultraID (19.7 kDa) and microID2 (19 kDa) may be advantageous to reduce the potential interference at the fusion protein for its proper folding, function, and localization. Notably, microID2 has been shown to be unstable in transgenic lines (49). Compared to other biotin ligases, this group shows optimum activity and thermal stability at higher temperatures, which may be suitable for applications in species such as desert plants or thermophilic algae that grow in high temperatures.

AirID (**Figure 3c**) is a new biotin ligase enzyme developed by the Sawasaki group (54) by using ancestral enzyme reconstruction and site-directed mutagenesis. AirID shows 82% sequence similarity to BioID but has a higher biotinylation activity. However, with a requirement of 3–6 h of labeling, AirID still exhibits lower enzyme kinetics than TurboID. The activity of AirID has been tested in mammalian cell cultures (131) but not in plants yet. BASU (**Figure 3d**) was developed by the Khavari group (92) from *B. subtilis* BirA by introducing a truncation of the N terminus and three mutations, including one R124G mutation, to its reactive biotinoyl-5'-AMP binding motif. BASU has a similar enzyme activity as BioID and BioID2 (10); its activity has been tested in mammalian cells (10) and in zebrafish (129) but not yet in plants.

Researchers need to determine which biotin ligases work best for their species as well as their specific experiments. TurboID and miniTurbo have been tested to work effectively in multiple plant species, including *Arabidopsis*, tobacco, tomato, liverwort, maize, and citrus (**Supplemental Table 1**). Whether TurboID and other biotin ligases are suitable for use in more plant species and organelles remains to be tested. Surprisingly, both TurboID and miniTurbo fail to work in the cytosol or nucleus of green algae, but a less-active intermediate G3 version (prior to TurboID) derived from yeast-directed evolution (10) works (Grossman lab and Onishi lab, unpublished observations); the exact reason for the failure of TurboID/miniTurbo in this context remains to be determined. Thus, it will be of interest to examine if other biotin ligases such as AirID, microID2, or ultraID work better than the G3 version in green algae. Similarly, for plant species in which TurboID has not been tested, researchers may initiate pilot experiments to compare different biotin ligase choices and determine which one works best. However, whether biotin ligases work in more organelles remains to be tested. For instance, TurboID has been extensively tested in several organelles and subcellular structures, but whether TurboID and other biotin ligases work in a vacuole with low pH or in the extracellular domain with low ATP concentration remains undetermined.

TurboID has the highest activity, so it is often the top choice for mapping proteomes in a temporal manner. Yet it may not always be the best choice since it has a higher propensity for using endogenous biotin than other biotin ligases. For instance, the high self-biotinylation of the bait-TurboID fusion protein may potentially inactivate the bait protein and/or interfere with the bait's

Supplemental Material >

Table 1 Comparison of the enrichment of biotin-labeled proteins at protein and peptide levels in plants

Enrichment	Biotin concentration	Evidence	Sensitivity	Reagents/equipment	Duration	Additional information
At protein level	Only allows a lower concentration of biotin treatment ($\leq 50 \mu\text{M}$)	Indirect evidence	More sensitive due to more abundance of nonbiotinylated peptides than of biotinylated peptides	Size-exclusion column (e.g., PD-10) to remove free biotin	4–5 days, including 2–3 days to perform enrichment (done by user) and 2 days on-bead digestion (done by MS facility)	Not applicable
At peptide level	Allows a higher concentration of biotin treatment, as biotin can be efficiently removed during protein precipitation	Direct and highly confident evidence	If biotinylated peptides are efficiently enriched, it can produce similar results due to reduced sample complexity for MS acquisition	Trypsin and desalting Zip Tips for digestion and desalting and SpeedVac for drying peptides	4–5 days (done by user)	Provides topology information and surface-exposed region of proteins.

function if surface-exposed biotin-modified lysine sites are essential to functionality. Additionally, if the turnover rate for some prey proteins is very low, then a relatively high background labeling introduced by TurboID may be problematic for temporal mapping, as the labeling may already be saturated before biotin and/or stimulus treatment. Note that biotin is a critical catalytic factor for several key carboxylases in plants (15); reducing the endogenous biotin level without causing stress to the plants can therefore be a daunting task. Thus, other biotin ligases, despite being less active than TurboID, may in some cases be a better choice with lower background labeling or less toxicity.

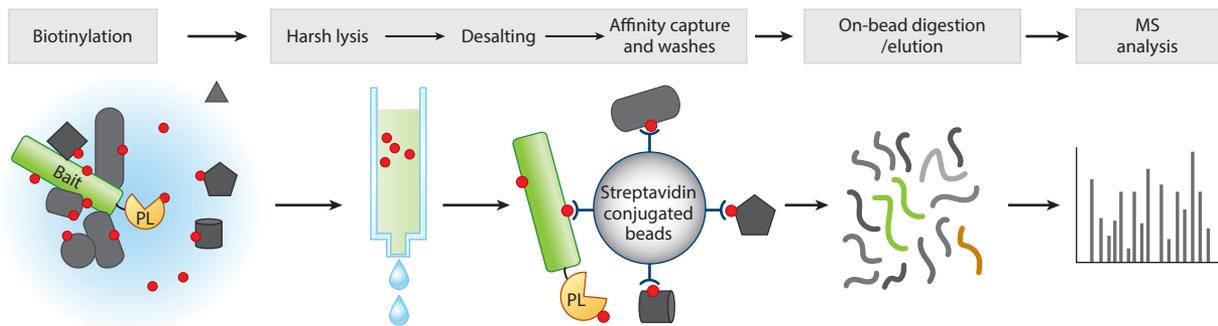
5. EFFECTIVE ENRICHMENT STRATEGIES FOR DETECTING BIOTINYLATED PROTEINS

Effective enrichment of biotinylated proteins is key for the success of PL experiments. The enrichment can be done at the protein or peptide level. Overall, the protein-level enrichment prevails with its simplicity; the peptide-level enrichment excels at the reliability of its data. Here, we describe the workflows of the two strategies and compare their respective strengths and weaknesses (Table 1).

5.1. Enrichment of Proximity-Labeled Proteins at the Protein Level

Most of the published PL experiments have adopted enrichment at the protein level, in which streptavidin beads capture intact proximity-labeled proteins in the binding step (67, 72) (Figure 4a). Because the streptavidin has an extraordinarily high binding affinity to biotin ($K_d = \sim 10^{-14}$ mol/L), the lysis, binding, and washes can be done in very harsh conditions (with detergent, salt, or denaturing reagents). However, the high affinity of streptavidin to biotin also poses a challenge to elution. To procure an efficient elution and to enable highly biotinylated proteins to be eluted at the same rate as less-biotinylated proteins, most labs perform on-bead digestions for elution in which the biotinylated proteins are eluted from the beads using a short (~ 3 -h) trypsin digestion, followed by overnight full digestion and analysis via LC-MS/MS (Figure 4a). Overall, the protein level enrichment is simple and straightforward. The drawback to this approach, however, is that the biotinylated peptides are rarely detected, mainly because the samples are overwhelmed by nonbiotinylated peptides and the trypsin digestion elution is too mild to break the biotinylated peptides' binding to the high-affinity streptavidin beads. To recover the remaining biotinylated peptides on the beads, Arora et al. (3) attempted a second-step elution

a Enrichment at protein level



b Enrichment at peptide level

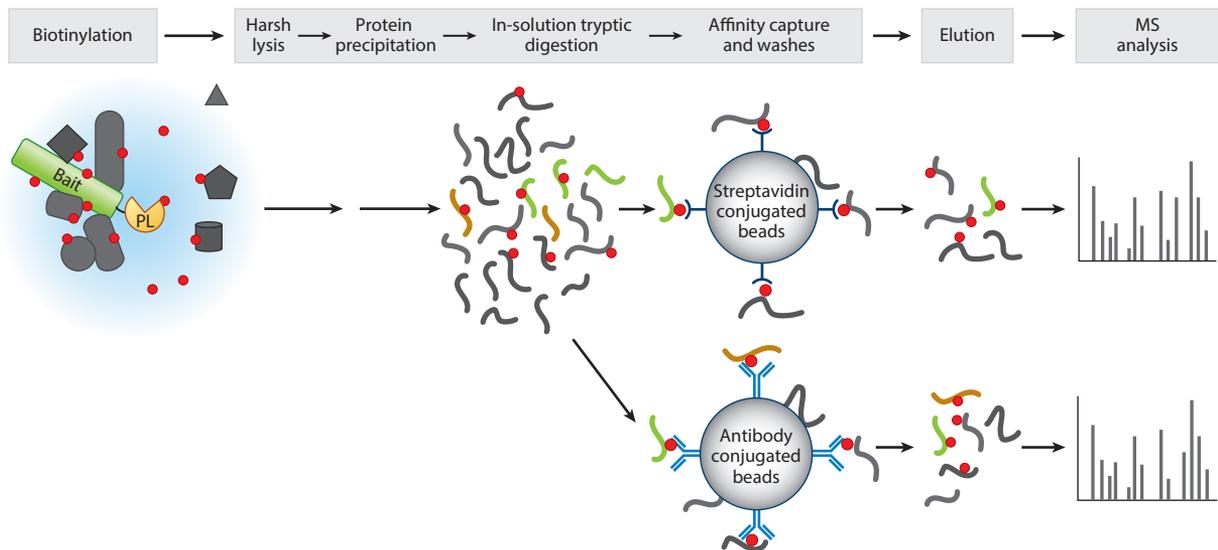


Figure 4

Comparison of protein- and peptide-level enrichment for the capture and detection of proximity-labeled biotinylated proteins.

(a) Depiction of the workflow of protein-level enrichment of biotinylated proteins in plant systems. Biotinylated samples are subjected to harsh lysis and the removal of free biotin by a size-exclusion column, followed by the affinity capture of biotinylated proteins using streptavidin beads and elution by on-bead tryptic digestion. The eluate is further digested by trypsin prior to LC-MS/MS analysis.

(b) For peptide-level enrichment of biotinylated proteins, biotinylated samples are extracted after harsh lysis, followed by protein precipitation. The proteins are then subjected to in-solution tryptic digestion to peptides. Biotinylated peptides are subsequently captured by either streptavidin beads followed by harsh elution or antibody-coated beads followed by mild elution prior to LC-MS/MS analysis. Biotin is represented by red circles. The blue region denotes the biotinylation radius. Abbreviations: LC-MS/MS, liquid chromatography-coupled mass spectrometry; MS, mass spectrometry; PL, proximity labeling.

using biotin and boiling with strong acid and were able to detect some biotinylated peptides from a small proportion of enriched proteins.

When using the protein-level enrichment strategy in plants, avoiding high concentration biotin treatment is important because free biotin competes with binding to streptavidin beads, leading to a significantly reduced capture (3, 58, 72). Meanwhile, conventional washes to remove free biotin used for human cell lines (22) are often not sufficient for plants. Hence, a desalting step using size-exclusion columns such as PD-10 to remove the free biotin and less than 50 μM of biotin treatment

are strongly recommended for experiments (3, 72). Additionally, removing free biotin through protein precipitation before enrichment has also been reported (83) but has remained challenging for the downstream streptavidin capture due to incomplete or inconsistent resolubilization of precipitated proteins.

5.2. Enrichment of Proximity-Labeled Proteins at the Peptide Level

Several labs, including ours, have attempted to enrich proximity-labeled proteins at the peptide level in a process in which streptavidin beads capture biotinylated peptides (instead of biotinylated proteins) in the binding step (**Figure 4b**). Compared to protein-level enrichment, this approach has several advantages. First, detection of biotinylation sites provides direct and highly confident evidence of PL (56, 123). Second, it allows for a higher concentration of biotin treatment as biotin can be removed during the protein precipitation step. Third, in some cases such as for the sub-cellular proteome, the biotinylation site information offers insights on the topology of proteins. The peptide-level enrichment approach uses conditions similar to those mentioned above to lyse the tissues, but the extracted proteins are first precipitated and subjected to trypsin digestion. The resulting peptides are then used as input for streptavidin bead capture (102) or antibody enrichment (56, 123), followed by LC-MS/MS analysis. Alternatively, biotinylated peptides are enriched using an engineered avidin-like protein that has reversible biotin-binding capacity (76).

Peptide-level enrichment is considered to be less sensitive than protein-level enrichment due to its much lower abundance compared to nonbiotinylated peptides (56, 123). In MS, identifying any peptide from a protein is sufficient to determine the protein accession number. However, our lab found that the two enrichment strategies produce nearly comparable results after optimizing the peptide enrichment protocols (>50% of peptides were identified as biotinylated peptides). We noticed that the overall complexity of the samples is dramatically reduced after peptide-level enrichment, which in turn increases the chance for detection of biotinylated peptides (R. Shrestha & S.-L. Xu, unpublished observations).

6. QUANTIFICATION STRATEGIES USED IN PROXIMITY LABELING–MASS SPECTROMETRY

Protein quantification is an integral step in the PL-MS workflow to identify enriched targets in PL-MS samples. The specific choice of quantification strategy depends on the experimental aims and designs, the studied species, and the cost. Here, we review the three most used quantification methods, including label-free quantification (LFQ), multiplexed tandem mass tag (TMT) isobaric labeling, and ¹⁵N metabolic labeling (**Figure 5**). We compare the strengths and weaknesses of each method in **Table 2**. We also briefly discuss the normalization method used in PL-MS analysis to achieve accurate and high-throughput quantification. To achieve better results, all quantification data should be acquired on high-resolution/high-accuracy instruments.

6.1. Label-Free Quantification

LFQ is the method most used in PL-MS quantification (44, 72, 138) due to its simplicity, low cost, and unlimited numbers of samples being compared (23). In addition, the downstream data processing following data acquisition is straightforward, using the publicly available software Maxquant/Perseus for data search and statistical analysis, and can be performed by most users with minimum training (24, 121, 122).

The LFQ method does not restrict the type of plant species, growth conditions, and growth duration that may be used. Typically, after trypsin digestion, each PL sample is directly injected for MS analysis (**Figure 5a**). The LFQs depend on the accurate determination of extracted ion

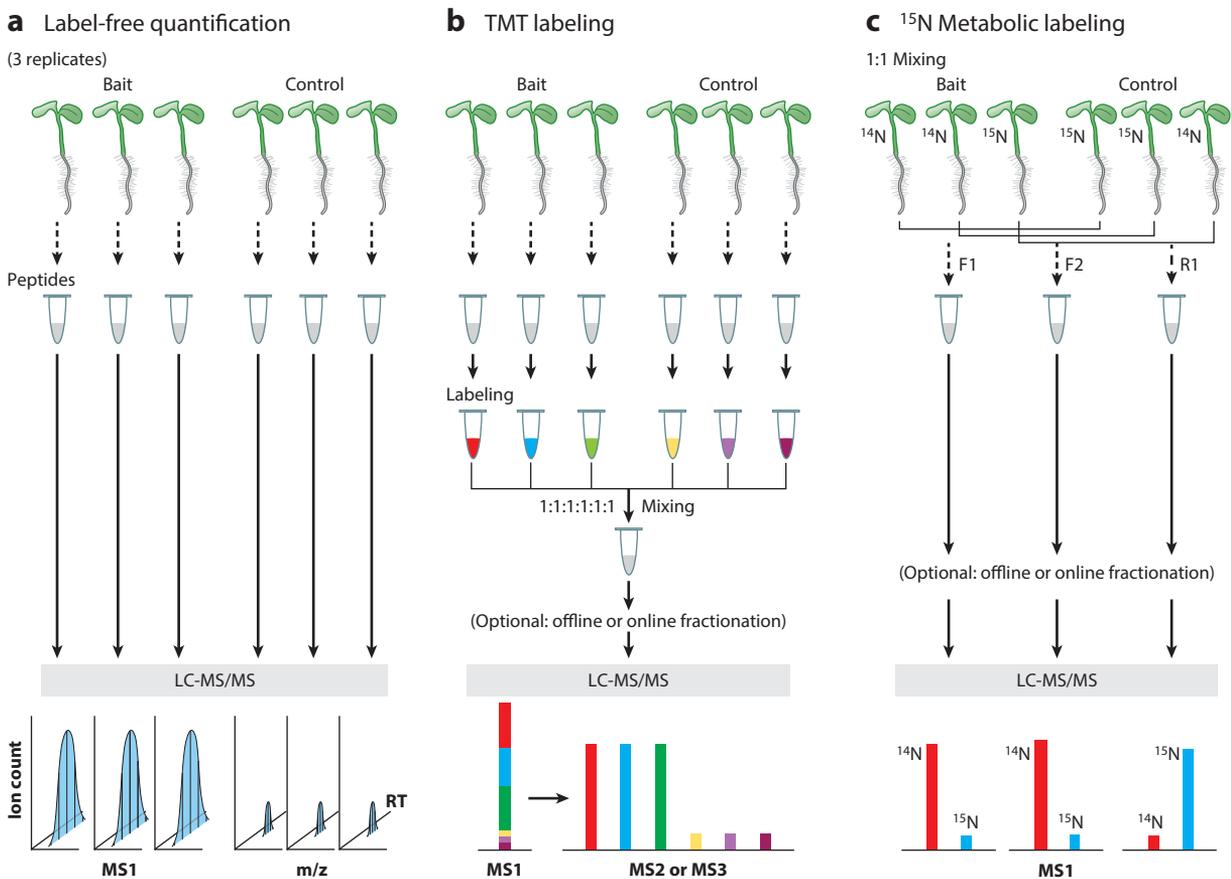


Figure 5

Schematic overview of the experimental workflow of three major quantification approaches in PL-MS experiments. Three biological replicates are required in each approach. (a) For label-free quantification, peptides from PL samples are directly measured by MS analysis. Protein abundance is quantified using the sum of XICs of peptides at the MS1 level. (b) For TMT isobaric labeling multiplexing quantification, each PL sample is labeled at the peptide level with the unique isobaric tag of the reagent, and the resulting barcoded samples are mixed in equal parts for simultaneous MS measurement. Quantification is based on the intensities of the yielded reporter ions at either the MS2 or MS3 levels. (c) For ^{15}N metabolic labeling, plants are grown in ^{14}N and ^{15}N media and mixed to generate two forward samples, F1 and F2 (^{14}N -bait/ ^{15}N -control), and one reverse sample, R1 (^{15}N -bait/ ^{14}N -control), at the beginning of sample processing. Quantification is done by comparing the intensities of ^{14}N and ^{15}N peaks using XICs at the MS1 level. In panels b and c, either offline or online fractionations can be incorporated to increase proteome coverage prior to MS analysis. Abbreviations: LC-MS/MS, liquid chromatography–coupled mass spectrometry; MS, mass spectrometry; m/z, mass to charge; PL, proximity labeling; RT, retention time; TMT, tandem mass tag; XIC, extracted ion chromatogram.

MS1: precursor ion scan or the first stage of tandem mass spectrometry

chromatograms (XICs) of peptides at the MS1 (precursor ion) level between the runs. Peptide ion signals from each protein are summed to increase the sensitivity of quantification and are compared between baits and controls to determine enrichment. A good LFQ relies heavily on a high-standard MS instrumentation condition, demanding similar sensitivity in mass spectrometers and minimum shifts of retention time in liquid chromatography across all data acquisitions within the same experiment. LFQ generally has a higher amount of variance than TMTs or ^{15}N metabolic labeling and does not allow fractionations for complex samples due to quantification complications (31, 93).

Table 2 Comparison of quantification methods used in proximity-labeling mass spectrometry

Labeling	Applied plant species	Multiplexing samples	Growth duration	Special reagents/kit needed	Fractionation to increase coverage	Quantification	Most used software for data analysis
Label-free quantification (LFQ)	All species	None	No restriction	No	No	MS1(precursor), ratio of summed peptide signals	Maxquant/Perseus (24, 121, 122)
Multiplexed TMT isobaric labeling	All species	Up to 18	No restriction	TMT isobaric tagging kit	Yes	Ratios from MS2 or MS3 fragments	Maxquant/Perseus (24, 121, 122), Protein Discoverer, Protein Prospector (109)
¹⁵ N metabolic labeling	<i>Arabidopsis</i> , <i>Cblamydomonas</i>	2	≥ 14 days, unless seeds are metabolic labeled for one generation	¹⁵ N salt (K ¹⁵ NO ₃ ; ¹⁵ NH ₄ ¹⁵ NO ₃)	Yes	MS1(precursor), median of ratios of peptides	Protein Prospector (109), pFIND (66)

6.2. Multiplexed Tandem Mass Tag Isobaric Labeling Quantification

Isobaric tagging reagents TMTs (31, 93) have been employed for relative quantification of a large number of samples used in PL-MS experiments (17, 138). For concurrent MS analysis of multiple samples, each PL sample is barcoded by a unique isobaric tag of the reagent at the peptide level, and the resulting labeled samples are mixed for MS analysis (**Figure 5b**). Up to 18 samples can be simultaneously analyzed using the most recent TMT18pro kits (68). Within a TMT labeling set, all mass tagging reagents (barcodes) have an identical nominal mass and chemical structure composed of a mass reporter, a spacer arm, and an amine-reactive *N*-hydroxysuccinimide (NHS) ester group. This NHS ester group conjugates and tags the peptide via a covalent reaction to the peptide's N terminus and lysine residue(s). Since different tags only vary in terms of distribution of the heavy isotopes (¹³C and ¹⁵N) within the reporter ions and mass normalization group, the same peptide from different samples will possess the same precursor mass after labeling and coelute in chromatography, appearing as a single peak in the MS1 (precursor ion) scan. After MS/MS (MS2) fragmentation of the precursor, each TMT will generate a unique reporter mass (e.g., 126 to 131 Da from TMTsixplex kits) in the low-mass region of the high-resolution MS/MS spectrum (**Figure 5b**). The abundance of these reporter ions is used for the relative quantification of peptide/protein abundance from each sample. Alternatively, the reporter ions can be generated and reported at the MS/MS/MS (MS3) level (in which multiple MS2 fragment ions are co-isolated and cofragmented to generate the MS3 spectrum). Quantifying at the MS2 level provides more sensitivity, while quantifying at the MS3 level provides more accuracy (118).

Like LFQ, TMT isotope labeling quantification can be applied to different species without restrictions on growth conditions or duration. Because samples are mixed and measured together, TMT labeling provides more accurate quantification than LFQ and enables a higher coverage (112) because it allows extensive fractionations, such as strong cation exchange or high pH reverse-phase high-performance liquid chromatography.

6.3. ¹⁵N Metabolic Labeling

¹⁵N metabolic labeling quantification is another method for PL-MS quantification in plants (58). In this workflow (**Figure 5c**), ¹⁵N-containing salts are incorporated into the plant proteins *in vivo*, contributing to a difference in mass measured at the MS1 (precursor) level for relative protein quantification (109). Samples are typically reciprocally labeled in replicates to reduce systematic error. ¹⁴N and ¹⁵N pairs are quantified at the MS1 (precursor) level using extraction ion chromatography, and then the median ratio of all quantified peptide pairs from each protein is reported

MS/MS (MS2): fragment ion scan after fragmenting the isolated precursor ions from the MS1 level, or the second stage of tandem mass spectrometry

MS/MS/MS (MS3): fragment ion scan after fragmenting the isolated ions from the MS2 level, or the third stage of tandem mass spectrometry

(109). As the samples can be mixed at the beginning of sample processing, the ^{15}N metabolic labeling significantly reduces preparative and analytical variabilities, thereby enabling more accurate quantifications compared to both LFQ and TMTs. Like TMTs, ^{15}N metabolic labeling allows for extensive fractionation of the complex samples to achieve a high coverage. The drawback of the ^{15}N metabolic labeling, however, is its restrictions on the type of plant species and the duration of plant growth (109).

6.4. Data Normalization

Normalization is a critical step in all quantifications. In most of the PL experiments in which a large collection of background proteins is detected, global normalization is sufficient to generate accurate quantification as it assumes the majority of all the detected proteins are similar across samples. However, for sample types that have overall much higher and more distinct signals than others, local normalization has to be applied to achieve better quantification. For instance, the majority of proteins identified from a ubiquitously expressed TurboID experiment are differentially abundant from those identified from a guard cell-specific TurboID (72), and a global normalization results in an overall reduction of signals and biased normalization and must therefore be corrected using a local normalization method such as LOESS-R (14).

7. VALIDATION OF CANDIDATES FROM PROXIMITY LABELING–MASS SPECTROMETRY EXPERIMENTS

PL-MS experiments, like most other discovery proteomics studies, generate a wealth of data. Identifying the known interactors or well-established markers often assures the quality of the data, and validating novel candidates is a necessary and exciting step to prove their biological relevance. Genetic perturbations (e.g., knockdowns/knockouts and overexpression) of candidate proteins from PL-MS experiments are common approaches to validate their biological functions. These genetic manipulations are expected to cause a phenotype that is similar to or opposite of that of the bait protein from PPI mapping (139), alter the organelle functions from subcellular proteomes (17, 124), or induce chromosome activation or inactivation from protein–DNA interaction mapping (133). To validate direct PPIs, bimolecular fluorescence complementation (BiFC) and yeast two-hybrid assays are commonly used (72, 139). Alternatively, in AP-MS or reciprocal PL-MS, switching the identified candidate to bait can also be utilized to validate the interactions. Less common approaches such as proximity ligation assay (6) can also be used to confirm PPIs. In addition, fluorescence imaging of newly identified candidates to verify their localization is often recommended for validation in subcellular and cell type-specific proteome studies. For candidates discovered from protein–DNA/RNA interactome PL-MS experiments, orthogonal approaches such as fluorescence in situ hybridization (e.g., immunoFISH), cross-linking immunoprecipitation sequencing (CLIP-seq), and RNA immunoprecipitation (RIP)-qPCR will serve the purpose for validation (133).

8. FUTURE PROSPECTS

As the development of PL-MS gains traction in plant research, mapping more dynamic interactomes and subcellular proteomes, particularly those that occur in response to environmental cues, will help address more biological questions. For instance, what is the protein interface for driving the light intensity-dependent movement of chloroplasts inside of cells (51) or for providing the precise guidance of pollen tube growth by the ovule during pollination (12)? We envision that elegant PL-MS experiments can be designed to answer these questions.

Coupling the power of PL-MS with posttranslational modification (PTM) studies will help to elucidate the regulatory mechanisms for dynamic changes (29). PTMs such as phosphorylation and *O*-GlcNAcylation have been shown to affect protein localization, PPIs, and protein–DNA interactions (132). Correlating and linking the PTM changes with protein changes discovered from PL-MS studies will be important. Recently, PL-MS has been coupled with a PTM reader to allow PTM-driven labeling in subcellular space and to monitor the dynamic changes of protein complexes. For example, a GlycoID-coupling miniTurboID with an *O*-GlcNAc reader (GafD, an *O*-GlcNAc binding protein) was created to track the dynamic changes of protein complexes associated with *O*-GlcNAc-modified targets (70).

Coupling PL-MS with cross-linking mass spectrometry (XL-MS) will also be a promising approach to reveal the architecture of protein complexes. While PL-MS provides a parts list for interactomes and subcellular proteomes, XL-MS offers connectivity and structures of the protein complexes (136). In XL-MS, after interacting proteins react with bifunctional cross-linking reagents to become cross-linked, proteins are extracted and subsequently digested to peptides and subjected to MS analysis. The identification of the cross-linked peptides will not only present direct evidence of connectivity between proteins but also map interacting domains as well as the structure of protein complexes. These PL-MS and XL-MS data sets can also be further combined with large-scale binary interactions based on yeast two-hybrid and cofractionation MS to build a more complete interaction map (2, 74, 119).

Another emerging research interest is to couple cell type-specific PL-MS with single-cell RNA-sequencing and proteomics. A quantitative atlas of the *Arabidopsis* proteome has been achieved at the tissue level (75) but not at cell-type nor single-cell level. While single-cell proteomics is still challenging (88, 89), it will gain momentum soon. With a joint effort of the community, generating such a wealth of data at different scales with a myriad of approaches will pave the way toward creating a plant cell atlas in the future (97).

SUMMARY POINTS

1. Proximity-labeling mass spectrometry (PL-MS) is a powerful tool used to dissect protein–protein interactions (PPIs), subcellular proteomes, and cell type-specific proteomes.
2. Engineered high-activity TurboID has extended its applications to multiple species and to answer more biological questions.
3. Modular designs of PL-MS have enabled the mapping of conditional PPIs and hard-to-purify subcellular structures and have improved the versatility of PL applications.
4. Coupling CRISPR with PL has opened a new avenue for probing the landscape of defined genome loci and for mapping RNA-binding protein complexes.
5. PL-MS can be combined with posttranslational modification studies to dissect regulatory mechanisms in the future.
6. PL-MS can be coupled with cross-linking MS and single-cell proteomics to build a plant cell atlas that will be crucial to the plant community.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Andres V. Reyes for helpful discussion and critical reading of the manuscript. We thank Danbi Byun for help in making the figures. This work was funded by the National Institutes of Health grants R01GM135706 and S10OD030441 to S.-L.X. and by the Carnegie Endowment Fund to the Carnegie Mass Spectrometry Facility.

LITERATURE CITED

1. Agrawal GK, Bourguignon J, Rolland N, Ephritikhine G, Ferro M, et al. 2011. Plant organelle proteomics: collaborating for optimal cell function. *Mass Spectrom. Rev.* 30:772–853
2. *Arabidopsis* Interactome Mapp. Consort. 2011. Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333:601–7
3. Arora D, Abel NB, Liu C, Van Damme P, Yperman K, et al. 2020. Establishment of proximity-dependent biotinylation approaches in different plant model systems. *Plant Cell* 32:3388–407
4. Bagchi P, Torres M, Qi L, Tsai B. 2020. Selective EMC subunits act as molecular tethers of intracellular organelles exploited during viral entry. *Nat. Commun.* 11:1127
5. Barkla BJ, Vera-Estrella R, Raymond C. 2016. Single-cell-type quantitative proteomic and ionic analysis of epidermal bladder cells from the halophyte model plant *Mesembryanthemum crystallinum* to identify salt-responsive proteins. *BMC Plant Biol.* 16:110
6. Bauer N, Škiljaica A, Malenica N, Razdorov G, Klasić M, et al. 2019. The MATH-BTB protein TaMAB2 accumulates in ubiquitin-containing foci and interacts with the translation initiation machinery in *Arabidopsis*. *Front. Plant Sci.* 10:1469
7. Bhattacharya O, Ortiz I, Walling LL. 2020. Methodology: an optimized, high-yield tomato leaf chloroplast isolation and stroma extraction protocol for proteomics analyses and identification of chloroplast co-localizing proteins. *Plant Methods* 16:131
8. Bi Y, Deng Z, Ni W, Shrestha R, Savage D, et al. 2021. Arabidopsis ACINUS is O-glycosylated and regulates transcription and alternative splicing of regulators of reproductive transitions. *Nat. Commun.* 12:945
9. Birnbaum K, Jung JW, Wang JY, Lambert GM, Hirst JA, et al. 2005. Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nat. Methods* 2:615–19
10. **Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, et al. 2018. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* 36:880–87**
11. Callegari S, Müller T, Schulz C, Lenz C, Jans DC, et al. 2019. A MICOS–TIM22 association promotes carrier import into human mitochondria. *J. Mol. Biol.* 431:2835–51
12. Chae K, Lord EM. 2011. Pollen tube growth and guidance: roles of small, secreted proteins. *Ann. Bot.* 108:627–36
13. Chan CJ, Le R, Burns K, Ahmed K, Coyaud E, et al. 2019. BioID performed on Golgi enriched fractions identify C10orf76 as a GBF1 binding protein essential for Golgi maintenance and secretion. *Mol. Cell. Proteom.* 18:2285–97
14. Chawade A, Alexandersson E, Levander F. 2014. Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets. *J. Proteome Res.* 13:3114–20
15. Che P, Weaver LM, Wurtele ES, Nikolau BJ. 2003. The role of biotin in regulating 3-methylcrotonyl-coenzyme A carboxylase expression in *Arabidopsis*. *Plant Physiol.* 131:1479–86
16. Chen Z, Lei C, Wang C, Li N, Srivastava M, et al. 2019. Global phosphoproteomic analysis reveals ARMC10 as an AMPK substrate that regulates mitochondrial dynamics. *Nat. Commun.* 10:104
17. **Cho KF, Branon TC, Rajeev S, Svinkina T, Udeshi ND, et al. 2020. Split-TurboID enables contact-dependent proximity labeling in cells. *PNAS* 117:12143–54**
18. Cho KF, Branon TC, Udeshi ND, Myers SA, Carr SA, Ting AY. 2020. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* 15:3971–99
19. Choi-Rhee E, Schulman H, Cronan JE. 2008. Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci.* 13:3043–50

10. Describes the engineering of high-activity TurboID.

17. Describes the engineering of split-TurboID.

20. Chojnowski A, Ong PF, Wong ESM, Lim JSY, Mutalif RA, et al. 2015. Progerin reduces LAP2 α -telomere association in Hutchinson-Gilford progeria. *eLife* 4:e07759
21. Christopher JA, Geladaki A, Dawson CS, Vennard OL, Lilley KS. 2022. Subcellular transcriptomics and proteomics: a comparative methods review. *Mol. Cell. Proteom.* 21:100186
22. Chua XY, Aballo T, Elnemer W, Tran M, Salomon A. 2021. Quantitative interactomics of Lck-TurboID in living human T cells unveils T cell receptor stimulation-induced proximal Lck interactors. *J. Proteome Res.* 20:715–26
23. Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. 2014. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteom.* 13:2513–26
24. Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26:1367–72
25. Dai S, Chen S. 2012. Single-cell-type proteomics: toward a holistic understanding of plant function. *Mol. Cell. Proteom.* 11:1622–30
26. Datta S, Malhotra L, Dickerson R, Chaffee S, Sen CK, Roy S. 2015. Laser capture microdissection: big data from small samples. *Histol. Histopathol.* 30:1255–69
27. De Munter S, Görmann J, Derua R, Lesage B, Qian J, et al. 2017. Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Lett.* 591:415–24
28. Déjardin J, Kingston RE. 2009. Purification of proteins associated with specific genomic loci. *Cell* 136:175–86
29. Dionne U, Gingras A-C. 2022. Proximity-dependent biotinylation approaches to explore the dynamic compartmentalized proteome. *Front. Mol. Biosci.* 9:852911
30. Döring F, Streubel M, Bräutigam A, Gowik U. 2016. Most photorespiratory genes are preferentially expressed in the bundle sheath cells of the C₄ grass *Sorghum bicolor*. *J. Exp. Bot.* 67:3053–64
31. Dowell JA, Wright LJ, Armstrong EA, Denu JM. 2021. Benchmarking quantitative performance in label-free proteomics. *ACS Omega* 6:2494–504
32. Fazal FM, Han S, Parker KR, Kaewsapsak P, Xu J, et al. 2019. Atlas of subcellular RNA localization revealed by APEX-seq. *Cell* 178:473–90.e26
33. Fernandez-Calvino L, Faulkner C, Walshaw J, Saalbach G, Bayer E, et al. 2011. Arabidopsis plasmodesmal proteome. *PLOS ONE* 6:e18880
34. Fridy PC, Li Y, Keegan S, Thompson MK, Nudelman I, et al. 2014. A robust pipeline for rapid production of versatile nanobody repertoires. *Nat. Methods* 11:1253–60
35. Furey TS. 2012. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions. *Nat. Rev. Genet.* 13:840–52
36. Gallardo VE, Behra M. 2013. Fluorescent activated cell sorting (FACS) combined with gene expression microarrays for transcription enrichment profiling of zebrafish lateral line cells. *Methods* 62:226–31
37. Gao XD, Tu L-C, Mir A, Rodriguez T, Ding Y, et al. 2018. C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9–APEX2. *Nat. Methods* 15:433–36
38. Geladaki A, Kočevár Britovšek N, Breckels LM, Smith TS, Vennard OL, et al. 2019. Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics. *Nat. Commun.* 10:331
39. Gingras A-C, Abe KT, Raught B. 2019. Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Curr. Opin. Chem. Biol.* 48:44–54
40. Go CD, Knight JDR, Rajasekharan A, Rathod B, Hesketh GG, et al. 2021. A proximity-dependent biotinylation map of a human cell. *Nature* 595:120–24
41. Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH. 2007. SUBA: the Arabidopsis Subcellular Database. *Nucleic Acids Res.* 35:D213–18
42. Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, et al. 2015. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* 163:712–23
43. Hooper CM, Castleden IR, Tanz SK, Aryamanesh N, Millar AH. 2017. SUBA4: the interactive data analysis centre for Arabidopsis subcellular protein locations. *Nucleic Acids Res.* 45:D1064–74
44. Huang A, Tang Y, Shi X, Jia M, Zhu J, et al. 2020. Proximity labeling proteomics reveals critical regulators for inner nuclear membrane protein degradation in plants. *Nat. Commun.* 11:3284

40. Assigns large-scale intercellular proteins using BioID.

45. Huang M-S, Lin W-C, Chang J-H, Cheng C-H, Wang HY, Mou KY. 2019. The cysteine-free single mutant C32S of APEX2 is a highly expressed and active fusion tag for proximity labeling applications. *Protein Sci.* 28:1703–12
46. Huberts DH, van der Klei IJ. 2010. Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta Mol. Cell Res.* 1803:520–25
47. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, et al. 2015. The BioPlex network: a systematic exploration of the human interactome. *Cell* 162:425–40
48. Ip P-L, Birnbaum KD. 2015. Tissue-specific gene expression profiling by cell sorting. *Methods Mol. Biol.* 1284:175–83
49. Johnson BS, Chafin L, Farkas D, Adair J, Elhance A, et al. 2022. MicroID2: a novel biotin ligase enables rapid proximity-dependent proteomics. *Mol. Cell Proteom.* 21:100256
50. Kaewsapsak P, Shechner DM, Mallard W, Rinn JL, Ting AY. 2017. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *eLife* 6:e29224
51. Kagawa T, Sakai T, Suetsugu N, Oikawa K, Ishiguro S, et al. 2001. *Arabidopsis* NPL1: a phototropin homolog controlling the chloroplast high-light avoidance response. *Science* 291:2138–41
52. Keilhauer EC, Hein MY, Mann M. 2015. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol. Cell. Proteom.* 14:120–35
53. Kennedy-Darling J, Guillen-Ahlers H, Shortreed MR, Scalf M, Frey BL, et al. 2014. Discovery of chromatin-associated proteins via sequence-specific capture and mass spectrometric protein identification in *Saccharomyces cerevisiae*. *J. Proteome Res.* 13:3810–25
54. Kido K, Yamanaka S, Nakano S, Motani K, Shinohara S, et al. 2020. AirID, a novel proximity biotinylation enzyme, for analysis of protein–protein interactions. *eLife* 9:e54983
55. Kim DI, Birendra KC, Zhu W, Motamedchaboki K, Doye V, Roux KJ. 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *PNAS* 111:E2453–61
56. Kim DI, Cutler JA, Na CH, Reckel S, Renuse S, et al. 2018. BioSITE: a method for direct detection and quantitation of site-specific biotinylation. *J. Proteome Res.* 17:759–69
57. Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, et al. 2016. An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* 27:1188–96
58. Kim T-W, Park CH, Hsu C-C, Kim Y-W, Ko Y-W, et al. 2023. Mapping the signaling network of BIN2 kinase using TurboID-mediated biotin labeling and phosphoproteomics. *Plant Cell* 35(3):975–93
59. Krämer U. 2015. Planting molecular functions in an ecological context with *Arabidopsis thaliana*. *eLife* 4:e06100
60. Kubitz L, Bitsch S, Zhao X, Schmitt K, Deweid L, et al. 2022. Engineering of ultraID, a compact and hyperactive enzyme for proximity-dependent biotinylation in living cells. *Commun. Biol.* 5:657
61. Kwak C, Shin S, Park J-S, Jung M, Nhung TTM, et al. 2020. Contact-ID, a tool for profiling organelle contact sites, reveals regulatory proteins of mitochondrial-associated membrane formation. *PNAS* 117:12109–20
62. Kwon K, Beckett D. 2000. Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci.* 9:1530–39
63. Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, et al. 2015. Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* 12:51–54
64. Lambert J-P, Tucholska M, Go C, Knight JDR, Gingras A-C. 2015. Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J. Proteom.* 118:81–94
65. Leijon F, Melzer M, Zhou Q, Srivastava V, Bulone V. 2018. Proteomic analysis of plasmodesmata from *Populus* cell suspension cultures in relation with callose biosynthesis. *Front. Plant Sci.* 9:1681
66. Li D, Fu Y, Sun R, Ling CX, Wei Y, et al. 2005. pFind: a novel database-searching software system for automated peptide and protein identification via tandem mass spectrometry. *Bioinformatics* 21:3049–50
67. Li H, Frankenfield AM, Houston R, Sekine S, Hao L. 2021. Thiol-cleavable biotin for chemical and enzymatic biotinylation and its application to mitochondrial TurboID proteomics. *J. Am. Soc. Mass Spectrom.* 32:2358–65

68. Li J, Cai Z, Bomgardner RD, Pike I, Kuhn K, et al. 2021. TMTpro-18plex: the expanded and complete set of TMTpro reagents for sample multiplexing. *J. Proteome Res.* 20:2964–72
69. Liu L, Doray B, Kornfeld S. 2018. Recycling of Golgi glycosyltransferases requires direct binding to coatamer. *PNAS* 115:8984–89
70. Liu Y, Nelson ZM, Reda A, Fehl C. 2022. Spatiotemporal proximity labeling tools to track GlcNAc sugar-modified functional protein hubs during cellular signaling. *ACS Chem. Biol.* 17:2153–64
71. Mair A, Bergmann DC. 2022. Advances in enzyme-mediated proximity labeling and its potential for plant research. *Plant Physiol.* 188:756–68
72. Mair A, Xu S-L, Branon TC, Ting AY, Bergmann DC. 2019. Proximity labeling of protein complexes and cell-type-specific organellar proteomes in *Arabidopsis* enabled by TurboID. *eLife* 8:e47864
73. Martell JD, Deerinck TJ, Sancak Y, Poulos TL, Mootha VK, et al. 2012. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* 30:1143–48
74. McWhite CD, Papoulas O, Drew K, Cox RM, June V, et al. 2020. A pan-plant protein complex map reveals deep conservation and novel assemblies. *Cell* 181:460–74.e14
75. Mergner J, Frejno M, List M, Papacek M, Chen X, et al. 2020. Mass-spectrometry-based draft of the *Arabidopsis* proteome. *Nature* 579:409–14
76. Motani K, Kosako H. 2020. BioID screening of biotinylation sites using the avidin-like protein Tama-vidin 2-REV identifies global interactors of stimulator of interferon genes (STING). *J. Biol. Chem.* 295:11174–83
77. Mukherjee J, Hermesh O, Eliscovich C, Nalpas N, Franz-Wachtel M, et al. 2019. β -Actin mRNA interactome mapping by proximity biotinylation. *PNAS* 116:12863–72
78. 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* 15:437–39
79. Nelson BK, Cai X, Nebenführ A. 2007. A multicolored set of *in vivo* organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* 51:1126–36
80. Nelson T, Gandotra N, Tausta SL. 2008. Plant cell types: reporting and sampling with new technologies. *Curr. Opin. Plant Biol.* 11:567–73
81. Ni W, Xu S-L, González-Grandío E, Chalkley RJ, Huhmer AFR, et al. 2017. PPKs mediate direct signal transfer from phytochrome photoreceptors to transcription factor PIF3. *Nat. Commun.* 8:15236
82. Ni W, Xu S-L, Tepperman JM, Stanley DJ, Maltby DA, et al. 2014. A mutually assured destruction mechanism attenuates light signaling in *Arabidopsis*. *Science* 344:1160–64
83. Niu L, Zhang H, Wu Z, Wang Y, Liu H, et al. 2018. Modified TCA/acetone precipitation of plant proteins for proteomic analysis. *PLOS ONE* 13:e0202238
84. Ohtsu K, Takahashi H, Schnable PS, Nakazono M. 2006. Cell type-specific gene expression profiling in plants by using a combination of laser microdissection and high-throughput technologies. *Plant Cell Physiol.* 48:3–7
85. Okekeogbu IO, Pattathil S, González Fernández-Niño SM, Aryal UK, Penning BW, et al. 2019. Glycome and proteome components of Golgi membranes are common between two angiosperms with distinct cell-wall structures. *Plant Cell* 31:1094–112
86. Pan R, Liu J, Wang S, Hu J. 2020. Peroxisomes: versatile organelles with diverse roles in plants. *New Phytol.* 225:1410–27
87. Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, et al. 2005. Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol. Biol. Cell* 16:260–69
88. Perkel JM. 2021. Single-cell proteomics takes centre stage. *Nature* 597:580–82
89. Petelski AA, Emmott E, Leduc A, Huffman RG, Specht H, et al. 2021. Multiplexed single-cell proteomics using SCoPE2. *Nat. Protoc.* 16:5398–425
90. Petricka JJ, Schauer MA, Megraw M, Breakfield NW, Thompson JW, et al. 2012. The protein expression landscape of the *Arabidopsis* root. *PNAS* 109:6811–18
91. Petrovská B, Jeřábková H, Chamrád I, Vrána J, Lenobel R, et al. 2014. Proteomic analysis of barley cell nuclei purified by flow sorting. *Cytogenet. Genome Res.* 143:78–86
92. Ramanathan M, Majzoub K, Rao DS, Neela PH, Zarnegar BJ, et al. 2018. RNA–protein interaction detection in living cells. *Nat. Methods* 15:207–12

72. Investigates TurboID applications in plants.

92. Investigates protein–RNA interactions using proximity labeling.

103. Investigates protein–DNA interactions using proximity labeling.

93. Rauniyar N, Yates JR 3rd. 2014. Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome Res.* 13:5293–309
94. Rayaprolu S, Bitarafan S, Santiago JV, Betarbet R, Sunna S, et al. 2022. Cell type-specific biotin labeling in vivo resolves regional neuronal and astrocyte proteomic differences in mouse brain. *Nat. Commun.* 13:2927
95. Rees JS, Li X-W, Perrett S, Lilley KS, Jackson AP. 2015. Protein neighbors and proximity proteomics. *Mol. Cell. Proteom.* 14:2848–56
96. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, et al. 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339:1328–31
97. Rhee SY, Birnbaum KD, Ehrhardt DW. 2019. Towards building a Plant Cell Atlas. *Trends Plant Sci.* 24:303–10
98. Rhee SY, Mutwil M. 2014. Towards revealing the functions of all genes in plants. *Trends Plant Sci.* 19:212–21
99. 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.* 196:801–10
100. Rugen N, Schaarschmidt F, Eirich J, Finkemeier I, Braun H-P, Eubel H. 2021. Protein interaction patterns in *Arabidopsis thaliana* leaf mitochondria change in dependence to light. *Biochim. Biophys. Acta Bioenerg.* 1862:148443
101. Samavarchi-Tehrani P, Samson R, Gingras A-C. 2020. Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches. *Mol. Cell. Proteom.* 19:757–73
102. Schiapparelli LM, McClatchy DB, Liu H-H, Sharma P, Yates JR, Cline HT 3rd. 2014. Direct detection of biotinylated proteins by mass spectrometry. *J. Proteome Res.* 13:3966–78
103. **Schmidtman E, Anton T, Rombaut P, Herzog F, Leonhardt H. 2016. Determination of local chromatin composition by CasID. *Nucleus* 7:476–84**
104. Schopp IM, Amaya Ramirez CC, Debeljak J, Kreibich E, Skribbe M, et al. 2017. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* 8:15690
105. Schulze WX, Schneider T, Starck S, Martinoia E, Trentmann O. 2012. Cold acclimation induces changes in Arabidopsis tonoplast protein abundance and activity and alters phosphorylation of tonoplast monosaccharide transporters. *Plant J.* 69:529–41
106. Scott JD, Pawson T. 2009. Cell signaling in space and time: where proteins come together and when they're apart. *Science* 326:1220–24
107. Senkler J, Senkler M, Eubel H, Hildebrandt T, Lengwenus C, et al. 2017. The mitochondrial complexome of *Arabidopsis thaliana*. *Plant J.* 89:1079–92
108. Shellhammer J, Meinke D. 1990. Arrested embryos from the *bio1* auxotroph of *Arabidopsis thaliana* contain reduced levels of biotin. *Plant Physiol.* 93:1162–67
109. Shrestha R, Reyes AV, Baker PR, Wang Z-Y, Chalkley RJ, Xu S-L. 2022. ¹⁵N metabolic labeling quantification workflow in Arabidopsis using Protein Prospector. *Front. Plant Sci.* 13:832562
110. Simm S, Papatirou DG, Ibrahim M, Leisegang MS, Müller B, et al. 2013. Defining the core proteome of the chloroplast envelope membranes. *Front. Plant Sci.* 4:11
111. Subbotin RI, Chait BT. 2014. A pipeline for determining protein–protein interactions and proximities in the cellular milieu. *Mol. Cell. Proteom.* 13:2824–35
112. Sun X, Sun H, Han X, Chen P-C, Jiao Y, et al. 2022. Deep single-cell-type proteome profiling of mouse brain by nonsurgical AAV-mediated proximity labeling. *Anal. Chem.* 94:5325–34
113. Sunna SN, Rayaprolu S, Bowen CA, Bagchi P, Seyfried NT, Rangaraju S. 2021. Cellular proteomic profiling using proximity labeling by TurboID in microglial and neuronal cell lines. *Alzheimer's Dement.* 17:e058534
114. Svozil J, Gruissem W, Baerenfaller K. 2015. Proteasome targeting of proteins in Arabidopsis leaf mesophyll, epidermal and vascular tissues. *Front. Plant Sci.* 6:376
115. Takano T, Wallace JT, Baldwin KT, Purkey AM, Uezu A, et al. 2020. Chemico-genetic discovery of astrocytic control of inhibition in vivo. *Nature* 588:296–302
116. Tang Y, Huang A, Gu Y. 2020. Global profiling of plant nuclear membrane proteome in *Arabidopsis*. *Nat. Plants* 6:838–47

117. Tanz SK, Castleden I, Hooper CM, Vacher M, Small I, Millar HA. 2013. SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.* 41:D1185–91
118. Ting L, Rad R, Gygi SP, Haas W. 2011. MS3 eliminates ratio distortion in isobaric labeling-based multiplexed quantitative proteomics. *Nat. Methods* 8:937–40
119. Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, et al. 2017. CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nat. Methods* 14:819–25
120. Trinkle-Mulcahy L. 2019. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res.* 8(F1000 Faculty Rev.):135
121. Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* 11:2301–19
122. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, et al. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13:731–40
123. Udeshi ND, Pedram K, Svinikina T, Fereshetian S, Myers SA, et al. 2017. Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Methods* 14:1167–70
124. Uezu A, Kanak DJ, Bradshaw TWA, Soderblom EJ, Catavero CM, et al. 2016. Identification of an elaborate complex mediating postsynaptic inhibition. *Science* 353:1123–29
125. Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ. 1997. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* 9:1963–71
126. Van Buskirk EK, Decker PV, Chen M. 2012. Photobodies in light signaling. *Plant Physiol.* 158:52–60
127. Wang W, Yuan J, Jiang C. 2021. Applications of nanobodies in plant science and biotechnology. *Plant Mol. Biol.* 105:43–53
128. Wyrich R, Dressen U, Brockmann S, Streubel M, Chang C, et al. 1998. The molecular basis of C₄ photosynthesis in sorghum: isolation, characterization and RFLP mapping of mesophyll- and bundle-sheath-specific cDNAs obtained by differential screening. *Plant Mol. Biol.* 37:319–35
129. **Xiong Z, Lo HP, McMahon K-A, Martel N, Jones A, et al. 2021. In vivo proteomic mapping through GFP-directed proximity-dependent biotin labelling in zebrafish. *eLife* 10:e64631**
130. Yadeta KA, Elmore JM, Coaker G. 2013. Advancements in the analysis of the *Arabidopsis* plasma membrane proteome. *Front. Plant Sci.* 4:86
131. Yamanaka S, Horiuchi Y, Matsuoka S, Kido K, Nishino K, et al. 2022. A proximity biotinylation-based approach to identify protein-E3 ligase interactions induced by PROTACs and molecular glues. *Nat. Commun.* 13:183
132. Yang X, Qian K. 2017. Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* 18:452–65
133. Yi W, Li J, Zhu X, Wang X, Fan L, et al. 2020. CRISPR-assisted detection of RNA–protein interactions in living cells. *Nat. Methods* 17:685–88
134. Yoshinaka T, Kosako H, Yoshizumi T, Furukawa R, Hirano Y, et al. 2019. Structural basis of mitochondrial scaffolds by prohibitin complexes: insight into a role of the coiled-coil region. *iScience* 19:1065–78
135. Youn J-Y, Dunham WH, Hong SJ, Knight JDR, Bashkurov M, et al. 2018. High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. *Mol. Cell* 69:517–32.e11
136. Yu C, Huang L. 2018. Cross-linking mass spectrometry: an emerging technology for interactomics and structural biology. *Anal. Chem.* 90:144–65
137. Yu C, Yang Y, Wang X, Guan S, Fang L, et al. 2016. Characterization of dynamic UbR-proteasome subcomplexes by *in vivo* cross-linking (X) assisted bimolecular tandem affinity purification (XBAP) and label-free quantitation. *Mol. Cell. Proteom.* 15:2279–92
138. Zhang Y, Li Y, Yang X, Wen Z, Nagalakshmi U, Dinesh-Kumar SP. 2020. TurboID-based proximity labeling for in planta identification of protein–protein interaction networks. *J. Vis. Exp.* 159:e60728
139. Zhang Y, Song G, Lal NK, Nagalakshmi U, Li Y, et al. 2019. TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. *Nat. Commun.* 10:3252
140. Zhang Z, Hu M, Feng X, Gong A, Cheng L, Yuan H. 2017. Proteomes and phosphoproteomes of anther and pollen: availability and progress. *Proteomics* 17:1600458

129. Describes a modular system to increase the versatility of proximity labeling.

141. Zhao Z, Stanley BA, Zhang W, Assmann SM. 2010. ABA-regulated G protein signaling in *Arabidopsis* guard cells: a proteomic perspective. *J. Proteome Res.* 9:1637–47
142. Zhao Z, Zhang W, Stanley BA, Assmann SM. 2008. Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20:3210–26
143. Zhu H, Li C, Gao C. 2020. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* 21:661–77
144. Zhu Y, Li H, Bhatti S, Zhou S, Yang Y, et al. 2016. Development of a laser capture microscope-based single-cell-type proteomics tool for studying proteomes of individual cell layers of plant roots. *Hortic. Res.* 3:16026
145. Zürcher E, Tavor-Deslex D, Lituiev D, Enkerli K, Tarr PT, Müller B. 2013. A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant Physiol.* 161:1066–75