Published in final edited form as:

Nat Methods. 2021 February ; 18(2): 133–143. doi:10.1038/s41592-020-01010-5.

Deciphering molecular interactions by proximity labeling

Wei Qin[1,*], **Kelvin F. Cho**[1,2,*], **Peter E. Cavanagh**[3,*], **Alice. Y. Ting**[1,4,5,6]

[1]Department of Genetics, Stanford University, Stanford, CA, USA

[2]Cancer Biology Program, Stanford University, Stanford, CA, USA

[3]Department of Biochemistry, Stanford University, Stanford, CA, USA

[4]Department of Biology, Stanford University, Stanford, CA, USA

[5]Department of Chemistry, Stanford University, Stanford, CA, USA

[6]Chan Zuckerberg Biohub, San Francisco, CA, USA

Abstract

Many biological processes are executed and regulated through the molecular interactions of proteins and nucleic acids. Proximity labeling (PL) is a technology for tagging the endogenous interaction partners of specific protein “baits”, via genetic fusion to promiscuous enzymes that catalyze the generation of diffusible reactive species in living cells. Tagged molecules that interact with baits can then be enriched and identified by mass spectrometry or nucleic acid sequencing. Here we review the development of PL technologies and highlight studies that have applied PL to discover and analyze molecular interactions. In particular, we focus on the use of PL for mapping protein-protein, protein-RNA, and protein-DNA interactions in living cells and organisms.

Cellular functions are tightly regulated by proteins, nucleic acids, and their interactions, including protein-protein interactions (PPIs), protein-RNA interactions, and protein-DNA interactions[1,2]. Such molecular interaction networks are central to most biological processes, while their dysregulation has been linked to a variety of human diseases including cancers, immune disorders, and neurodegeneration. Methods enabling the large-scale discovery of molecular interactions in living cells have provided valuable insights for biological exploration and therapeutic intervention.

The traditional approaches of affinity purification (AP) and yeast two-hybrid (Y2H) have been widely applied to discover potential molecular interactions[3,4]. Antibody-based AP, in combination with mass spectrometry (MS)-based proteomics, allows the enrichment and identification of stable interaction partners of specific proteins of interest. Such efforts have expanded our understanding of protein interaction networks in a variety of systems, including yeast, flies, and human cells. AP can also be combined with crosslinking and

Correspondence should be addressed to Alice Y. Ting (ayting@stanford.edu).
*These authors contributed equally to this work.

Author Contributions

All authors contributed to the writing and editing of the paper.

Competing Interests

None.

nucleic acid sequencing to interrogate protein-nucleic acid interactions, such as in chromatin immunoprecipitation sequencing (ChIP-seq) and RNA immunoprecipitation sequencing (RIP-seq)[5,6]. The major limitation of AP, however, is that weak or transient interactions are often lost during cell lysis and the subsequent washing steps. To overcome this, AP can be combined with crosslinking[7], however this increases the rate of false positives. Moreover, AP is challenging to apply to insoluble targets or protein baits lacking high-affinity antibodies.

Y2H and other protein complementation assays (PCAs) represent another approach for mapping protein-protein, as well as protein-RNA, and protein-DNA interactions in living cells[3]. These approaches are often high throughput, enabling the screening of thousands to millions of potential molecular interactions[8]. However, many PCAs have cell-type and organelle-type restrictions (for example, Y2H does not work on membrane proteins), false positives due to overexpression and tagging of both bait and prey, and false negatives due to steric interference or geometric constraints of the required tags.

Proximity labeling (PL) was developed to provide a complementary approach to these traditional methods for molecular interaction mapping in living cells. PL uses engineered enzymes, such as peroxidases (APEX2[9], HRP[10]) or biotin ligases (BioID[11,12], BioID2[13], BASU[14], TurboID[15], miniTurbo[15]), that are genetically tagged to a protein of interest (POI) (Table 1). The PL enzyme converts an inert small-molecule substrate into a short-lived reactive species, such as a radical in the case of APEX[16] or an activated ester in the case of BioID/TurboID[17], that diffuses out from the enzyme active site to covalently tag neighboring endogenous species (Fig. 1a, b). The labeling radius is determined by both the small-molecule half-life and the concentration of quenchers in the environment, such as glutathione for APEX and amines for BioID/TurboID. The experimentally-determined labeling radii for HRP, APEX, BioID, and TurboID enzymes fall in the range of 1-10 nm in living cells[16,17]. However, instead of a fixed radius, it is more accurate to think of labeling by PL enzymes as a “contour map” in which the reactant concentration is highest at the PL enzyme and falls off nanometer by nanometer from the source[17,18]. Peroxidase- and biotin-ligase generated reactive species are also membrane impermeant[19], and thus the contour map ends at membrane boundaries. The substrate molecule typically contains a biotin handle to enable subsequent enrichment of tagged species using streptavidin beads and their identification by mass spectrometry (for proteins) or nucleic acid sequencing (for RNA) (Fig. 1c). Depending on the localization and expression of the PL enzyme, PL can be used to interrogate spatial proteomes on several different length scales – from entire cells[20], to organelles and subcellular compartments[18,19,21-25], to macromolecular complexes[26,27]. In this review, we will focus on the application of PL to studying molecular interactions, which has more published examples than organelle or cellular mapping[28].

For this review, we define PL as labeling catalyzed by genetically-encoded enzymes (as opposed to chemical catalysts) that generate diffusible reactive species in living systems. Many other conceptually-related technologies have been described but fall outside the scope of this review. These include proximity ligation assays (PLA) on fixed cells with antibody[29] or nucleic acid probes[30], photocrosslinking with unnatural amino acids[31], promiscuous

enzymes that label with non-diffusible substrates (e.g., PUP-IT)[32], and light-activated chemical catalysts[33].

Proximity labeling for profiling protein-protein interactions

PL has been applied to a wide range of protein-protein interaction (PPI) mapping problems, from signal transduction networks (MAPK[34,35], Hippo[36], Adrenergic[37], GPCR[38,39]) to enzyme-substrate interactions (E3 ligases[40,41], kinases[42]). These studies have been conducted in a variety of cell types (2D/3D culture[43], endothelial cells[44], neuronal cells[22,45,46], etc.) and organisms (bacteria[15,47], yeast[9,15,48,49], flies[15,41,50,51], worms[15], plants[52,53], mice[37,54-56], and primary human tissue[57]). In this section, we highlight some areas where proximity labeling has offered advantages over traditional methods in identifying molecular interactions and enabling biological discovery. These include characterizing the architecture of insoluble protein complexes (e.g. the nuclear envelope), capturing transient PPIs (e.g. enzyme-substrate interactions), dissecting dynamic processes (e.g. GPCR signaling), an enabling the specific interrogation of interactomes in live organisms (Table 2).

PL has enabled the study of insoluble baits that are difficult to analyze by AP, such as lamin A/C, a nuclear envelope resident protein critical for maintaining nuclear envelope structure[11]. To map lamin A/C's interaction partners by PL, Roux et al. used BioID, a promiscuous mutant of the *E. coli* biotin ligase BirA[12], and fused it directly to lamin A/C in HeLa cells[11]. Residents of the nuclear membrane and other previously unknown interactors, such as the nuclear pore complex were identified[11]. Subsequent work using PL enzymes have further built on this work by additional interactome mapping of other lamins, nuclear envelope proteins[17,58,59], and nuclear transporters[60].

Protein aggregates are extreme examples of insoluble baits. Chou et al. used BioID to identify interactors of TDP43 aggregates, a common histopathological marker of neurodegenerative disease, including ALS and frontotemporal dementia disease[61]. By fusing BioID to TDP43 to perform PL, the authors identified nucleocytoplasmic transport machinery and follow-up studies implicated TDP43 aggregates' disruption of nucleoporin and transport factor functions as a mechanism for pathology[61].

PL has proven especially useful in dissecting signaling pathways, where upstream and downstream effectors often interact only transiently. For example, Amber et al. used BioID to probe interactors along the Hippo pathway[36], a highly conserved signaling cascade that controls cell proliferation and apoptosis to dictate organ size. By mapping the interactomes of 19 pathway proteins using BioID, the authors generated protein interaction networks for the Hippo pathway and identified numerous putative regulators and kinase substrates[36]. PL-based interactome mapping has also been successfully used to map other signaling processes, such as NFκB[62], Ras[63,64], MAPK[34,35], and Hedgehog[65] pathways. PL-based interactome mapping can also uncover the remodeling of signaling pathways in the context of disease[66] and upon pathway activation to discover critical mediators of signal transduction[37,39]. In addition to intracellular interactome mapping, PL has been used to identify extracellular ligand – receptor interactions[32,67,68].

PL-based PPI mapping has also been informative for the study of enzyme-substrate interactions, where interactions are intrinsically transient due to substrate turnover[40-42]. E3 ubiquitin ligases in particular, which influence many aspects of cellular biology by controlling protein ubiquitination and degradation, each have numerous adapter proteins and substrates[69]. Etienne et al. used BioID in conjunction with pharmacological proteasome inhibition to probe interactors of SCF E3 ligases β-TrCP1 and β-TrCP2[40]. Using this approach, the authors validated twelve new substrates, including proteins involved in nuclear membrane integrity and translation control. PL has also been used to interrogate substrates of protein kinases. For instance, Cutler et al. fused BioID to p190 and p210 BCR-ABL tyrosine kinases, oncogenic protein fusions that result from chromosomal translocations[42]. Using PL, the authors identified distinct interactomes of each fusion and revealed that the Src family kinase Lyn, critical for transformation and drug resistance, is a preferential substrate of the p190 BCR-ABL fusion.

The short time-frame of APEX labeling (<1 minute) has been leveraged to capture temporally-resolved snapshots of changing interactomes of proteins involved in dynamic cellular processes, such as in Wnt[70] and GPCR signaling[38,39]. Paek et al. applied APEX-based PL to AT1R and β2AR GPCR signaling, in response to agonist activation[38], and proteomic analysis of the changing interactome supported the role of endocytosis in secluding GPCRs from G-proteins and demonstrated differing endocytosis kinetics for different GPCRs. Additionally, APEX has been used to capture snapshots of the δ-opioid receptor (DOR) interactome following treatment with agonist[39]. By identifying a time-course of protein interactions and using a set of spatial references to increase specificity in the context of receptor internalization and trafficking, Lobingier et al. identified two ubiquitin-pathway proteins implicated as mediators of DOR endosomal trafficking to the lysosome[39]. APEX has also been used to dissect the specificity of Wnt signaling. After demonstrating that Wnt9a signals by binding the Fzd9b receptor through an unknown factor, Grainger et al. leveraged the rapidity of APEX labeling to map the proteome specifically during receptor activation and identified EGFR as a key mediator of Wnt9a-Fzd9b interactions[70]. Overall, these studies and others have capitalized on the rapid labeling of APEX to dissect their respective pathways on a minute time scale, demonstrating the full potential of PL to probe dynamic interactions.

Many proteins participate in multiple distinct protein complexes that each carry out different cellular functions, but fusing PL enzymes directly to the bait in these scenarios would result in labeling proximal interactors of each complex, thereby reducing the confidence for those of a certain subpopulation. To overcome this, PL tools have been further adapted using various strategies for mapping interactomes of specific subcellular pools of a particular protein of interest, with the potential to dramatically improve specificity. For example, James et al. developed a strategy to probe only the inner nuclear membrane-localized pool of VAPB, which is localized to both the ER membrane and nuclear membrane. By taking advantage of the chemically inducible dimerization FRB-FKBP system, the authors employed rapamycin-dependent recruitment of nuclear-targeted APEX2-FKBP to inner nuclear membrane-localized FRB-VAPB but not ER membrane-localized FRB-VAPB[71]. More generalizable PL approaches for increasing spatial specificity have been developed in the form of split PL enzymes. Split PL enzymes consist of two inactive fragments

that can be brought together by protein-protein interactions or membrane apposition to reconstitute enzymatic activity[20,72-76]. Split-APEX2[75], split-HRP[20], various versions of split-BioIDs[72-74], and split-TurboID[76] have all been developed, with advantages and disadvantages mirroring their full-length counterparts. While not yet widely adapted for PPI mapping, the application of split enzymes for PL could drastically improve spatial specificity for mapping certain protein-protein interactions. For example, Schopp et al. successfully used split-BioID to probe interactors of the miRNA-induced silencing complex (miRISC)[73]. During complex maturation, the protein subunit Ago2 participates in two distinct subcomplexes containing either Dicer or TNRC6. By fusing fragments of split-BioID to Ago2 and TNRC6, and then to Ago2 and Dicer, the authors were able to differentiate distinct interactomes of each of the respective subcomplexes[73].

The development of biotin ligase-based PL approaches has also enabled PL studies *in vivo* across multiple organisms. While peroxidase-based approaches have been applied in various *ex vivo* studies[37,50], the requirement for hydrogen peroxide limits their use *in vivo*. Furthermore, peroxidase-based PL in plants is problematic because of background activity from endogenous plant peroxidases. BioID has been applied for proteomic mapping in *A. thaliana* and *N. benthamiana*[52,77], two key plant models. The development of more active TurboID and miniTurbo has improved these approaches[53]. Zhang et al. utilized TurboID to identify interactors of a plant immune receptor called N[78]. By using TurboID to perform biotin labeling in live *N. benthamiana* plants, the authors identified the interactor UBR7, a putative E3 ligase that downregulates N and mediates plant immunity against plant pathogens.

Studies in many model and non-model organisms have benefited from the simple and non-toxic labeling conditions of biotin ligase-based PL. PL has been carried out in live bacteria[15,47], yeast[15,48,49], slime molds[79,80], various parasites[81-85], worms[15], flies[15,86,87], and mice[55,56]. In the first *in vivo* mouse PL study, Dingar et al. fused BioID to the oncogene c-Myc, expressed this fusion construct in xenografted cells, and performed biotin labeling over the course of two days before proteomic analysis[56]. In a subsequent mouse study, Uezu et al. used BioID to map the inhibitory postsynaptic density over a course of seven days of biotin labeling before proteomic analysis[55]. These long labeling times were likely required for generating sufficient biotinylated material for mass spectrometry due to the low activity of BioID. The application of the more active TurboID or miniTurbo enzymes in future *in vivo* PL studies may offer increased temporal control for mapping dynamic processes in live organisms.

Proximity labeling for profiling protein-RNA interactions

The interactions between proteins and RNA are critical for a wide range of cellular functions, from transcription and translation to innate immunity and stress response[2], and a number of approaches have been developed to study these interactions[88]. Existing methods can be broadly classified as protein-centric or RNA-centric (Table 3). In protein-centric methods, the RNA interaction partners of a specific protein bait of interest can be identified by RNA sequencing. In RNA-centric methods, the protein partners of a specific RNA bait are identified[88]. There are many more protein-centric methods for mapping protein-RNA

interactions due to the availability of antibodies for protein pull-down and the ease of RNA sequencing.

In protein-centric methods, the addition of a chemical or UV crosslinking step prior to protein bait immunoprecipitation improves the efficiency of RNA capture. CLIP-seq and related methods have been widely applied to the detection of RNAs associated with a particular protein, and generally these methods are highly specific and can be carried out without the exogenous expression of any components[89-94]. However, existing approaches are limited by antibody quality, and UV crosslinking has low efficiency. Furthermore, these methods query RNA-protein interactions across the entire cell, while there may exist compartment-specific variability; for instance, a specific protein bait may localize to both the nucleus and cytosol and interact with different RNA partners in each location[95].

PL has been combined with RNA-protein crosslinking to discover RNAs proximal to protein baits in specific subcellular locales. APEX-RIP[96] uses formaldehyde, while Proximity-CLIP[97] uses UV, to crosslink APEX-biotinylated proteins to RNA just prior to cell lysis, enabling streptavidin-based enrichment of RNA-protein complexes (Fig. 2a). The methods were applied to the ER membrane[96], nuclear lamina[96], and cell-cell interfaces[97]. Using Proximity-CLIP, Benhalevy et al. observed the enrichment of CUG repeats in the RBP-protected footprints of mRNA 3'UTRs localized to cell-cell interfaces, among other functional insights of protein-RNA occupancy[97].

In APEX-RIP, the use of formaldehyde adds time, complexity, and degrades spatial specificity. In a more direct approach, APEX-seq bypasses the need for RNA-protein crosslinking altogether, and uses an APEX fusion protein to directly biotinylate proximal endogenous RNAs (Fig. 2b) [98,99]. After 1 minute labeling in live cells, streptavidin is used to enrich tagged RNAs for RNA-seq. An improved variation of APEX-seq uses a more efficient substrate, biotin-aniline, which improves RNA capture efficiency[100]. Fazal et al. used APEX-seq to generate a transcriptome-wide subcellular RNA atlas in HEK293T cells, uncovering numerous functional insights and correlating RNA-transcript location with genome architecture and protein localization[98]. By taking advantage of APEX's rapid kinetics, we used APEX-seq to quantify RNA dynamics at the outer mitochondrial membrane in response to drug perturbations and identified two distinct pathways for mRNA localization to the outer mitochondrial membrane[98]. APEX-seq has also been used to study stress granules, providing insights into the organization of translation initiation complexes on active mRNAs[99].

An alternative protein-centric PL method, CAP-seq, incorporates light-activated miniSOG for proximity-dependent photo-oxidation of RNA nucleobases, which can be subsequently captured by amine probes and identified by high-throughput sequencing[101] (Fig. 2c). Although the temporal resolution of CAP-seq (~20 min) is lower than that of APEX-seq (<1 min), both approaches offer distinct mechanisms of RNA labeling and may be complementary. Compared to traditional protein-centric sequencing methods, PL-based APEX-seq and CAP-seq do not require antibodies or crosslinking steps and can be easily adapted for identifying interacting or proximal RNAs of specific RBPs.

In contrast to protein-centric methods, RNA-centric methods target an RNA of interest to identify its protein binding partners. Traditionally, these approaches involve crosslinking and RNA capture using biotinylated oligonucleotide probes or MS2 tags[102-107]. However, the development of RNA-centric PL offers an alternative that does not require crosslinking. The first RNA-centric PL method reported, RaPID (RNA-protein interaction detection), allows for the biotinylation of RNA binding proteins (RBPs) by tagging an RNA of interest with a BoxB aptamer to recruit a fusion protein of λ-N and the biotin ligase BASU[14] (Fig. 2d). RaPID was used to discover host proteins that interact with Zika virus RNA[14]. In similar approaches, the MS2 coat protein has been fused to BioID[108] and to APEX2[109] to recruit these PL enzymes to MS2-tagged RNAs (Fig. 2e). However, these methods map proteins that interact with exogenously expressed tagged RNA, which may not accurately reflect the interactome of native transcripts.

The development of RNA-directed CRISPR systems offers the opportunity to target endogenous RNAs. For example, Han et al. targeted catalytically inactive RfxCas13d fused with APEX2 and a double-stranded RNA binding domain (dsRBD) (to enhance its binding affinity) to human telomerase RNA (hTR)[109] (Fig. 2f). Using this approach, the authors discovered a previously unknown interaction between hTR and the *N*6-methyladenosine (m6A) demethylase, ALKBH5, and subsequent studies showed that post-transcriptional regulation by ALKBH5 affects both telomerase complex assembly and activity. Alternative methods have been developed that combine inactive dCas13 orthologs with BioID2[110], APEX2[111], PUP-IT[112], or BASU[113] labeling for RBP profiling. These approaches vary in their benefits and drawbacks; for example, different dCas13 orthologs may exhibit differential binding to the accessible regions of the target RNA, and the chosen PL enzyme will have corresponding benefits and limitations, as previously discussed (Table 1). A potential limitation of these approaches is the large size of Cas13, which may sterically interfere with RBP binding; alternative strategies to target PL enzymes to specific RNAs may further improve RNA-centric discovery.

Proximity labeling for profiling protein-DNA interactions

Protein-DNA interactions play vital roles in the regulation of gene expression, genome integrity, and chromatin organization. Chromatin Immunoprecipitation-Sequencing (ChIP-seq) is widely used to capture and sequence DNA regions associated with a POI[6]. The PL adaptation of this approach occurs in living cells and uses the peroxidase APEX to biotinylate proteins proximal to a bait, which are in turn crosslinked by formaldehyde to neighboring DNA regions. Subsequently, biotinylated protein-DNA fragment complexes are enriched by streptavidin and analyzed by next-generation sequencing. ALaP (for APEX-mediated chromatin labeling and purification) is conceptually analogous to APEX-RIP for RNA identification[114], and offers improved sensitivity but decreased specificity in comparison to traditional ChIP-seq. ALaP has also been further adapted for mapping the genomic contact sites of promyelocytic leukemia (PML) bodies, phase-separated nuclear structures that physically interact with chromatin.

For DNA-centered mapping, wherein proteins proximal to a genomic locus or chromatin complex of interest are identified in an unbiased manner, several methods have been

developed. Three groups independently combined PL with CRISPR-based genome targeting[115-117]. Fusing APEX2 with catalytically inactive dCas9 to target specific genomic loci (e.g. telomeres and centromeres) allowed associated proteins to be biotinylated, enriched, and analyzed by mass spectrometry[115,116] (Fig. 2g). For discovery of proteins associated with specific chromatin complexes, RIME[118] and ChIP-MS[119] were reported. More recently, ChromID was used to interrogate protein interactomes at specific chromatin marks by fusing BASU to engineered readers specific to chromatin modifications[120] (Fig. 2h). ChromID identified novel promoter regions modified by H3K4me3 and H3K27me3[120]. Although the presence of targeting enzymes may affect the interactors that bind to chromatin, these studies provide a unique tool to investigate the regulatory mechanisms of chromatin functions. Of note, APEX-based PL has also been applied for mapping proteins associated with mitochondrial DNA, uncovering seven previously unknown mitochondrial nucleoid-associated proteins[26].

Limitations of proximity labeling

Molecular interaction mapping with PL-based approaches requires direct fusion of a PL enzyme to the protein of interest, requiring either transfection of the fusion construct or an alternative induction method such as viral infection. The fusion can potentially affect the function, localization, or even interactome of the target. Thus, it is crucial that functional and localization assays are performed to confirm that the PL enzyme fusion construct remains physiologically relevant and behaves similarly to the endogenous protein of interest. Furthermore, the selection of PL enzymes depends highly on the specific application, as each enzyme has its own advantages and disadvantages (Table 1). For example, the requirement of hydrogen peroxide in APEX labeling may compromise redox-sensitive proteins or pathways and hinder *in vivo* applications, whereas biotin ligases, such as BioID or TurboID, are less toxic and more suitable in these scenarios.

Because some published PL datasets do not utilize quantitative approaches for data collection and analyses, these datasets may be considered candidate lists that may contain considerable false positives. However, PL experiments can produce highly specific datasets if quantitative mass spectrometry is used while including proper controls for ratiometric or statistical analyses. For example, we have previously used APEX2 to generate a highly specific proteome of the outer mitochondrial membrane (OMM) by comparing the extent of biotinylation of proteins by APEX2 targeted to the OMM versus APEX2 expressed in the cytosol[23]. However, PL-based technology may exhibit decreased sensitivity due to various reasons. For instance, in the example described above, the ratiometric analysis filters out dual-localized proteins - proteins that reside in both the cytosol and on the OMM. Furthermore, proteins that lack surface exposed tyrosines (in the case of APEX) or lysines (in the case of BioID/TurboID) may not be detected, and different PL enzymes may exhibit biases towards labeling certain protein substrates[121]. Additional details regarding setting up, optimizing, analyzing, and troubleshooting PL experiments may be found in two protocols publications from our laboratory[18,122].

Conclusions and Outlook

The technological advances in molecular interaction mapping using proximity labeling have enabled biological investigations previously difficult to access. However, additional tool development and engineering may allow more comprehensive interactome maps and improve spatiotemporal specificity in a greater diversity of model systems. While biotin ligases such as BioID and now TurboID have been successfully utilized in a number of organisms for *in vivo* proteomic mapping, further optimization such as use of non-biotin probes to avoid background from endogenously biotinylated proteins, may improve compatibility for PL *in vivo*. For protein-nucleic acid mapping, improving the efficiency of RNA/DNA labeling by PL enzymes will boost sensitivity and analysis of transcriptomes and genomes in distinct cell populations. Furthermore, improvements in CRISPR-based nucleic acid targeting and binding stability should improve PL approaches that use this mechanism and enable application to endogenous transcripts expressed at low levels. Multiplexing PL enzymes and enrichment strategies could allow simultaneous molecular interactome mapping for multiple complexes at a time. While PL has enabled molecular interaction mapping in many previously intractable biological systems (e.g. transient interactions, insoluble baits, *in vivo* interactions, etc.), continuing development of increasingly sophisticated PL technology may vastly expand the range of PL-based discoveries and address more challenging questions, such as determining the affinity, stoichiometry, and contact sites of molecular interactions.

Acknowledgements

This work was supported by the NIH R01-DK121409 (to A.Y.T.) and Stanford Wu Tsai Neurosciences Institute Big Ideas Initiative (to A.Y.T.). K.F.C. was supported by NIH Training Grant 2T32CA009302-41 and the Blavatnik Graduate Fellowship. P.E.C. was supported by the NSF Graduate Research Fellowship. A.Y.T. is an investigator of the Chan Zuckerberg Biohub.

References

1. Keskin O, Tuncbag N & Gursoy A Predicting Protein-Protein Interactions from the Molecular to the Proteome Level. Chemical Reviews 116, 4884–4909 (2016). [PubMed: 27074302]
2. Hentze MW, Castello A, Schwarzl T & Preiss T A brave new world of RNA-binding proteins. Nature Reviews Molecular Cell Biology 19, 327–341 (2018). [PubMed: 29339797]
3. Brückner A, Polge C, Lentze N, Auerbach D & Schlattner U Yeast two-hybrid, a powerful tool for systems biology. International Journal of Molecular Sciences 10, 2763–2788 (2009). [PubMed: 19582228]
4. Dunham WH, Mullin M & Gingras AC Affinity-purification coupled to mass spectrometry: Basic principles and strategies. Proteomics 12, 1576–1590 (2012). [PubMed: 22611051]
5. Visa N & Jordán-Pla A ChIP and ChIP-related techniques: Expanding the fields of application and improving ChIP performance. in Methods in Molecular Biology 1689, 1–7 (2018). [PubMed: 29027160]
6. Park PJ ChIP-seq: Advantages and challenges of a maturing technology. Nature Reviews Genetics 10, 669–680 (2009).
7. Liu F, Rijkers DTS, Post H & Heck AJR Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nat. Methods 12, 1179–1184 (2015). [PubMed: 26414014]
8. Trigg SA et al. CrY2H-seq: A massively multiplexed assay for deep-coverage interactome mapping. Nat. Methods 14, 819–825 (2017). [PubMed: 28650476]

9. Lam SS et al. Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat. Methods 12, 51–54 (2014). [PubMed: 25419960] This study used directed evolution to develop APEX2
10. Kotani N et al. Biochemical visualization of cell surface molecular clustering in living cells. Proc. Natl. Acad. Sci. U. S. A 105, 7405–7409 (2008). [PubMed: 18495923]
11. Roux KJ, Kim DI, Raida M & Burke B A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. 196, 801–810 (2012). [PubMed: 22412018] This study introduced the first variant of BioID for promiscuous proximity labeling.
12. Choi-Rhee E, Schulman H & Cronan JE Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. Protein Sci. 13, 3043–3050 (2008).
13. Kim DI et al. An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell 27, 1188–1196 (2016). [PubMed: 26912792]
14. Ramanathan M. et al. RNA-protein interaction detection in living cells. Nat. Methods 15, 207–212 (2018). [PubMed: 29400715] This was the first PL-based study to identify RBPs associated with an RNA of interest.
15. Branon TC et al. Efficient proximity labeling in living cells and organisms with TurboID. Nature Biotechnology 36, 880–898 (2018). This study used directed evolution to engineer TurboID from BioID.
16. Martell JD et al. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. Nat. Biotechnol. 30, 1143–1148 (2012). [PubMed: 23086203]
17. Kim DI et al. Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proc. Natl. Acad. Sci. U. S. A 111, E2453–E2461 (2014). [PubMed: 24927568]
18. Hung V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat. Protoc 11, 456–475 (2016). [PubMed: 26866790]
19. Rhee HW et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science (80-.). 339, 1328–1331 (2013). [PubMed: 23371551]
20. Martell JD et al. A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. Nat. Biotechnol 34, 774–780 (2016). [PubMed: 27240195]
21. Markmiller S. et al. Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell 172, 590–604.e13 (2018). [PubMed: 29373831]
22. Loh KH et al. Proteomic Analysis of Unbounded Cellular Compartments: Synaptic Clefts. Cell 166, 1295–1307.e21 (2016). [PubMed: 27565350]
23. Hung V. et al. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. Elife 6, (2017).
24. Youn JY et al. High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol. Cell 69, 517–532.e11 (2018). [PubMed: 29395067]
25. Kehrer J, Frischknecht F & Mair GR Proteomic analysis of the plasmodium berghei gametocyte egressome and vesicular bioid of osmiophilic body proteins identifies merozoite trap-like protein (MTRAP) as an essential factor for parasite transmission. Mol. Cell. Proteomics 15, 2852–2862 (2016). [PubMed: 27371728]
26. Han S. et al. Proximity Biotinylation as a Method for Mapping Proteins Associated with mtDNA in Living Cells. Cell Chem. Biol 24, 404–414 (2017). [PubMed: 28238724]
27. Miyagawa-Yamaguchi A, Kotani N & Honke K Expressed glycosylphosphatidylinositol-anchored horseradish peroxidase identifies co-clustering molecules in individual lipid raft domains. PLoS One 9, (2014).
28. Gingras AC, Abe KT & Raught B Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. Curr. Opin. Chem. Biol 48, 44–54 (2019). [PubMed: 30458335]
29. Gullberg M. et al. Cytokine detection by antibody-based proximity ligation. Proc. Natl. Acad. Sci. U. S. A 101, 8420–8424 (2004). [PubMed: 15155907]
30. Fredriksson S. et al. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol 20, 473–477 (2002). [PubMed: 11981560]

31. Yang Y. et al. Genetically encoded protein photocrosslinker with a transferable mass spectrometry-identifiable label. Nat. Commun 7, (2016).
32. Liu Q. et al. A proximity-tagging system to identify membrane protein–protein interactions. Nat. Methods 15, 715–722 (2018). [PubMed: 30104635]
33. Geri JB et al. Microenvironment mapping via Dexter energy transfer on immune cells. Science (80-.). 367, 1091–1097 (2020). [PubMed: 32139536]
34. Dumont AA, Dumont L, Berthiaume J & Auger-Messier M p38α MAPK proximity assay reveals a regulatory mechanism of alternative splicing in cardiomyocytes. Biochim. Biophys. Acta - Mol. Cell Res 1866, (2019).
35. Prikas E, Poljak A & Ittner A Mapping p38α mitogen-activated protein kinase signaling by proximity-dependent labeling. Protein Sci. 29, 1196–1210 (2020). [PubMed: 32189389]
36. Couzens AL et al. Protein interaction network of the mammalian hippo pathway reveals mechanisms of kinase-phosphatase interactions. Sci. Signal 6, rs15–rs15 (2013). [PubMed: 24255178] This study fused BioID to 19 members of the Hippo pathway for interactome mapping
37. Liu G. et al. Mechanism of adrenergic CaV1.2 stimulation revealed by proximity proteomics. Nature 577, 695–700 (2020). [PubMed: 31969708]
38. Paek J. et al. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. Cell 169, 338–349.e11 (2017). [PubMed: 28388415] This study utilized APEX to map the dynamic interactomes of the GPCRs, AT1R and β2AR, in response to agonist activation.
39. Lobingier BT et al. An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. Cell 169, 350–360.e12 (2017). [PubMed: 28388416] This study utilized APEX to map the dynamic interactomes of the GPCR, δ-opioid receptor, in response to agonist activation.
40. Coyaud E. et al. BioID-based identification of skp cullin F-box (SCF)β-TrCP1/2 E3 ligase substrates. Mol. Cell. Proteomics 14, 1781–1795 (2015). [PubMed: 25900982]
41. Mannix KM, Starble RM, Kaufman RS & Cooley L Proximity labeling reveals novel interactomes in live Drosophila tissue. Development 146, (2019).
42. Cutler JA et al. Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. Leukemia 31, 1513–1524 (2017). [PubMed: 28210003]
43. Rodríguez-Fraticelli AE et al. Developmental regulation of apical endocytosis controls epithelial patterning in vertebrate tubular organs. Nat. Cell Biol 17, 241–250 (2015). [PubMed: 25706235]
44. Holthenrich A, Drexler HCA, Chehab T, Naß J & Gerke V Proximity proteomics of endothelial Weibel-Palade bodies identifies novel regulator of von Willebrand factor secretion. Blood 134, 979–982 (2019). [PubMed: 31262780]
45. Liao YC et al. RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. Cell 179, 147–164.e20 (2019). [PubMed: 31539493]
46. Chung CY et al. In Situ Peroxidase Labeling and Mass-Spectrometry Connects Alpha-Synuclein Directly to Endocytic Trafficking and mRNA Metabolism in Neurons. Cell Syst. 4, 242–250.e4 (2017). [PubMed: 28131823]
47. Santin YG et al. In vivo TssA proximity labelling during type VI secretion biogenesis reveals TagA as a protein that stops and holds the sheath. Nat. Microbiol 3, 1304–1313 (2018). [PubMed: 30275513]
48. Opitz N. et al. Capturing the Asc1p/Receptor for Activated C Kinase 1 (RACK1) Microenvironment at the head region of the 40s ribosome with quantitative BioID in Yeast. Mol. Cell. Proteomics 16, 2199–2218 (2017). [PubMed: 28982715]
49. Larochelle M, Bergeron D, Arcand B & Bachand F Proximity-dependent biotinylation mediated by TurboID to identify protein-protein interaction networks in yeast. J. Cell Sci 132, (2019).
50. Li J. et al. Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. Cell 180, 373–386.e15 (2020). [PubMed: 31955847]
51. Domsch K. et al. The hox transcription factor ubx stabilizes lineage commitment by suppressing cellular plasticity in drosophila. Elife 8, (2019).
52. Khan M, Youn JY, Gingras AC, Subramaniam R & Desveaux D In planta proximity dependent biotin identification (BioID). Sci. Rep 8, (2018).

53. Mair A, Xu SL, Branon TC, Ting AY & Bergmann DC Proximity labeling of protein complexes and cell type specific organellar proteomes in Arabidopsis enabled by TurboID. Elife 8, (2019).
54. Feng W. et al. Identifying the cardiac dyad proteome in vivo by a BioID2 knock-in strategy. Circulation 141, 940–942 (2020). [PubMed: 32176542]
55. Uezu A. et al. Identification of an elaborate complex mediating postsynaptic inhibition. Science (80-.). 353, 1123–1129 (2016). [PubMed: 27609886]
56. Dingar D. et al. BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. J. Proteomics 118, 95–111 (2015). [PubMed: 25452129] This was the first study to use BioID in vivo in mice.
57. Bar DZ et al. Biotinylation by antibody recognition—a method for proximity labeling. Nat. Methods 15, 127–133 (2018). [PubMed: 29256494]
58. Birendra KC et al. VRK2A is an A-type lamin-dependent nuclear envelope kinase that phosphorylates BAF. Mol. Biol. Cell 28, 2241–2250 (2017). [PubMed: 28637768]
59. Fu Y. et al. MacroH2A1 associates with nuclear lamina and maintains chromatin architecture in mouse liver cells. Sci. Rep 5, (2015).
60. Mackmull M. et al. Landscape of nuclear transport receptor cargo specificity. Mol. Syst. Biol 13, 962 (2017). [PubMed: 29254951]
61. Chou CC et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. Nat. Neurosci 21, 228–239 (2018). [PubMed: 29311743]
62. Phelan JD et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature 560, 387–391 (2018). [PubMed: 29925955]
63. Cui Y. et al. The NF2 tumor suppressor merlin interacts with Ras and RasGAP, which may modulate Ras signaling. Oncogene 38, 6370–6381 (2019). [PubMed: 31312020]
64. Che Y. et al. KRAS regulation by small non-coding RNAs and SNARE proteins. Nat. Commun 10, (2019).
65. Mirza AN et al. LAP2 Proteins Chaperone GLI1 Movement between the Lamina and Chromatin to Regulate Transcription. Cell 176, 198–212.e15 (2019). [PubMed: 30503211]
66. Park S, Ponce-balbuena D, Kuick R & Guerrero-serna G Kir2 . 1 interactome mapping uncovers PKP4 as a modulator of the Kir2 . 1-regulated inward rectifier potassium currents . Mol. Cell. Proteomics (2020).
67. Chang L. et al. Identification of siglec ligands using a proximity labeling method. J. Proteome Res 16, 3929–3941 (2017). [PubMed: 28899088]
68. Wu G, Nagala M & Crocker PR Identification of lectin counter-receptors on cell membranes by proximity labeling. Glycobiology 27, 800–805 (2017). [PubMed: 28810661]
69. Zheng N & Shabek N Ubiquitin Ligases: Structure, Function, and Regulation. Annu. Rev. Biochem 86, 129–157 (2017). [PubMed: 28375744]
70. Grainger S. et al. EGFR is required for Wnt9a–Fzd9b signalling specificity in haematopoietic stem cells. Nat. Cell Biol 21, 721–730 (2019). [PubMed: 31110287]
71. James C. et al. Proteomic mapping by rapamycin-dependent targeting of APEX2 identifies binding partners of VAPB at the inner nuclear membrane. J. Biol. Chem 294, 16241–16254 (2019). [PubMed: 31519755]
72. De Munter S. et al. Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. FEBS Letters 591, 415–424 (2017). [PubMed: 28032891]
73. Schopp IM et al. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. Nat. Commun 8, (2017).
74. Kwak C. et al. Contact-ID, a new tool for profiling organelle contact site, reveals proteins of mitochondrial-associated membrane formation. Proc. Natl. Acad. Sci (2020). doi:10.1073/pnas.1916584117
75. Han Y. et al. Directed Evolution of Split APEX2 Peroxidase. ACS Chem. Biol 14, 619–635 (2019). [PubMed: 30848125]
76. Cho KF et al. Split-TurboID enables contact-dependent proximity labeling in cells. Proc. Natl. Acad. Sci. U. S. A 117, 12143–12154 (2020). [PubMed: 32424107]

77. Conlan B, Stoll T, Gorman JJ, Saur I & Rathjen JP Development of a rapid in planta bioid system as a probe for plasma membrane-Associated immunity proteins. Front. Plant Sci. 871, (2018).

78. Zhang Y. et al. TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. Nat. Commun 10, (2019).

79. Batsios P, Meyer I & Gräf R Proximity-Dependent Biotin Identification (BioID) in Dictyostelium Amoebae. in Methods in Enzymology 569, 23–42 (2016). [PubMed: 26778551]

80. Pitzen V, Askarzada S, Gräf R & Meyer I CDK5RAP2 Is an Essential Scaffolding Protein of the Corona of the Dictyostelium Centrosome. Cells 7, 32 (2018). [PubMed: 29690637]

81. Morriswood B. et al. Novel bilobe components in Trypanosoma brucei identified using proximity-dependent biotinylation. Eukaryot. Cell 12, 356–367 (2013). [PubMed: 23264645]

82. Chen AL et al. Novel components of the toxoplasma inner membrane complex revealed by BioID. MBio 6, (2015).

83. McAllaster MR et al. Proteomic identification of novel cytoskeletal proteins associated with TbPLK, an essential regulator of cell morphogenesis in Trypanosoma brucei. Mol. Biol. Cell 26, 3013–3029 (2015). [PubMed: 26133384]

84. Khosh-Naucke M. et al. Identification of novel parasitophorous vacuole proteins in P. falciparum parasites using BioID. Int. J. Med. Microbiol 308, 13–24 (2018). [PubMed: 28784333]

85. Tu V. et al. The Toxoplasma gondii cyst wall interactome. MBio 11, (2020).

86. Shinoda N, Hanawa N, Chihara T, Koto A & Miura M Dronc-independent basal executioner caspase activity sustains Drosophila imaginal tissue growth. Proc. Natl. Acad. Sci. U. S. A 116, 20539–20544 (2019). [PubMed: 31548372]

87. Carnesecchi J. et al. Multi-level and lineage-specific interactomes of the Hox transcription factor Ubx contribute to its functional specificity. Nat. Commun. 11, (2020).

88. Ramanathan M, Porter DF & Khavari PA Methods to study RNA–protein interactions. Nature Methods 16, 225–234 (2019). [PubMed: 30804549]

89. Licatalosi DD et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456, 464–469 (2008). [PubMed: 18978773]

90. Hafner M. et al. Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. Cell 141, 129–141 (2010). [PubMed: 20371350]

91. Kim B & Kim VN fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA. Methods 152, 3–11 (2019). [PubMed: 29902563]

92. Van Nostrand EL et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods 13, 508–514 (2016). [PubMed: 27018577]

93. Konig J. et al. ICLIP - transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution. J. Vis. Exp (2011). doi:10.3791/2638

94. Zarnegar BJ et al. IrCLIP platform for efficient characterization of protein-RNA interactions. Nat. Methods 13, 489–492 (2016). [PubMed: 27111506]

95. Trendel J. et al. The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. Cell 176, 391–403.e19 (2019). [PubMed: 30528433]

96. Kaewsapsak P, Shechner DM, Mallard W, Rinn JL & Ting AY Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. Elife 6, (2017).

97. Benhalevy D, Anastasakis DG & Hafner M Proximity-CLIP provides a snapshot of protein-occupied RNA elements in subcellular compartments. Nat. Methods 15, 1074–1082 (2018). [PubMed: 30478324]

98. Fazal FM et al. Atlas of Subcellular RNA Localization Revealed by APEX-Seq. Cell 178, 473–490.e26 (2019). [PubMed: 31230715] This study used APEX to systematically investigate subcellular localization of RNA.

99. Padrón A, Iwasaki S & Ingolia NT Proximity RNA Labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules. Mol. Cell 75, 875–887.e5 (2019). [PubMed: 31442426]

100. Zhou Y. et al. Expanding APEX2 Substrates for Proximity-Dependent Labeling of Nucleic Acids and Proteins in Living Cells. Angew. Chemie - Int. Ed 58, 11763–11767 (2019).

101. Wang P. et al. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. Nat. Chem. Biol 15, 1110–1119 (2019). [PubMed: 31591565]
102. McHugh CA & Guttman M RAP-MS: A method to identify proteins that interact directly with a specific RNA molecule in cells. in Methods in Molecular Biology 1649, 473–488 (2018). [PubMed: 29130217]
103. Zeng F. et al. A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. Nat. Protoc 1, 920–927 (2006). [PubMed: 17406325]
104. Matia-González AM, Iadevaia V & Gerber AP A versatile tandem RNA isolation procedure to capture in vivo formed mRNA-protein complexes. Methods 118–119, 93–100 (2017). [PubMed: 27746303]
105. Simon MD et al. The genomic binding sites of a noncoding RNA. Proc. Natl. Acad. Sci. U. S. A 108, 20497–20502 (2011). [PubMed: 22143764]
106. Chu C, Qu K, Zhong FL, Artandi SE & Chang HY Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. Mol. Cell 44, 667–678 (2011). [PubMed: 21963238]
107. Tsai BP, Wang X, Huang L & Waterman ML Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach. Mol. Cell. Proteomics 10, (2011).
108. Mukherjee J. et al. β-Actin mRNA interactome mapping by proximity biotinylation. Proc. Natl. Acad. Sci. U. S. A 116, 12863–12872 (2019). [PubMed: 31189591]
109. Han S. et al. RNA-protein interaction mapping via MS2 or Cas13-based APEX targeting. bioRxiv 2020.02.27.968297 (2020). doi:10.1101/2020.02.27.968297
110. Li Y. et al. CBRPP a new RNA-centric method to study RNA-protein interactions. bioRxiv 2020.04.09.033290 (2020). doi:10.1101/2020.04.09.033290
111. Lin X & Lawrenson K In Vivo Analysis of RNA Proximity Proteomes Using RiboPro. bioRxiv 2020.02.28.970442 (2020). doi:10.1101/2020.02.28.970442
112. Zhang Z. et al. Capturing RNA-protein interaction via CRUIS. Nucleic Acids Res. 48, e52 (2020). [PubMed: 32140725]
113. Yi W. et al. CRISPR-assisted detection of RNA – protein interactions in living cells. Nat. Methods (2020). doi:10.1038/s41592-020-0866-0
114. Kurihara M. et al. Genomic Profiling by ALaP-Seq Reveals Transcriptional Regulation by PML Bodies through DNMT3A Exclusion. Mol. Cell 78, 493–505.e8 (2020). [PubMed: 32353257]
115. Myers SA et al. Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. Nat. Methods 15, 437–439 (2018). [PubMed: 29735997]
116. Gao XD et al. C-BERST: Defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. Nat. Methods 15, 433–436 (2018). [PubMed: 29735996]
117. Qiu W. et al. Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. Nucleic Acids Res. 47, (2019).
118. Mohammed H. et al. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nature Protocols 11, 316–326 (2016). [PubMed: 26797456]
119. Wang CI et al. Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. Nat. Struct. Mol. Biol 20, 202–209 (2013). [PubMed: 23295261]
120. Villaseñor R. et al. ChromID identifies the protein interactome at chromatin marks. Nat. Biotechnol (2020). doi:10.1038/s41587-020-0434-2 This study used biotin ligase to identify proteins associated with specific chromatin marks.
121. Minde DP, Ramakrishna M & Lilley KS Biotin proximity tagging favours unfolded proteins and enables the study of intrinsically disordered regions. Commun. Biol. (2020). doi:10.1038/s42003-020-0758-y
122. Cho KF et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. Nat. Protoc (accepted) (2020).
123. Cole A. et al. Inhibition of the Mitochondrial Protease ClpP as a Therapeutic Strategy for Human Acute Myeloid Leukemia. Cancer Cell 27, 864–876 (2015). [PubMed: 26058080]

124. Shin JJH, Gillingham AK, Begum F, Chadwick J & Munro S TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi. Nat. Cell Biol 19, 1424–1432 (2017). [PubMed: 29084197]

125. Liu L, Doray B & Kornfeld S Recycling of Golgi glycosyltransferases requires direct binding to coatomer. Proc. Natl. Acad. Sci. U. S. A 115, 8984–8989 (2018). [PubMed: 30126980]

126. Jia J. et al. Galectins Control mTOR in Response to Endomembrane Damage. Mol. Cell 70, 120–135.e8 (2018). [PubMed: 29625033]

127. Tenenbaum SA, Carson CC, Lager PJ & Keene JD Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. Proc. Natl. Acad. Sci. U. S. A 97, 14085–14090 (2000). [PubMed: 11121017]

128. McMahon AC et al. TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. Cell 165, 742–753 (2016). [PubMed: 27040499]

129. Lapointe CP, Wilinski D, Saunders HAJ & Wickens M Protein-RNA networks revealed through covalent RNA marks. Nat. Methods 12, 1163–1170 (2015). [PubMed: 26524240]

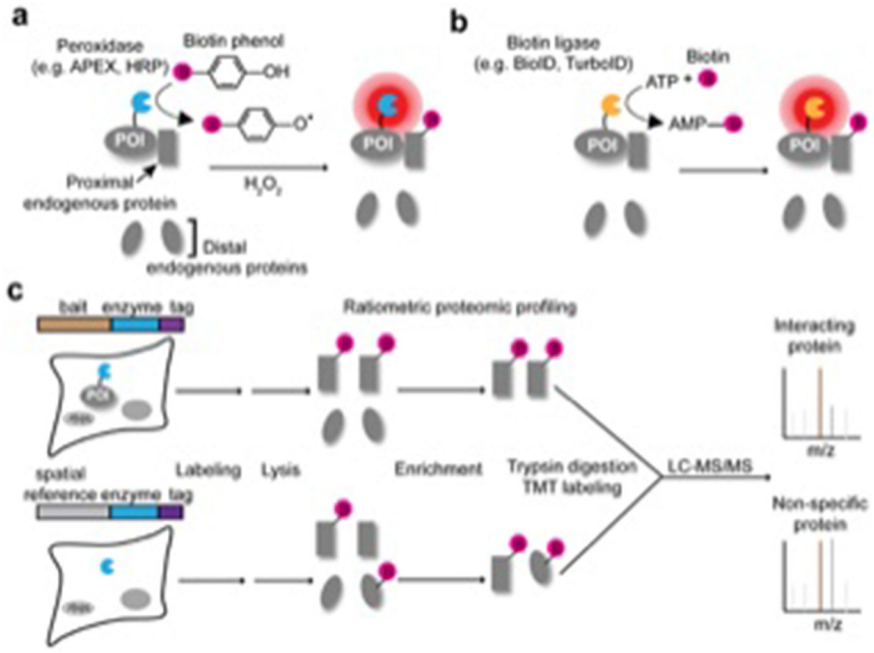

Figure 1.

Peroxidase- and biotin ligase-based proximity labeling methods for PPI mapping. (**a**) Peroxidase-based approaches, such as APEX or HRP, oxidize biotin phenol into reactive phenoxyl radicals using hydrogen peroxide, which preferentially labels proximal over distal endogenous proteins. (**b**) Biotin ligase-based approaches, such as BioID or TurboID, utilize ATP and biotin to catalyze the formation of reactive biotin-5’-AMP intermediates, which diffuse and label proximal proteins. (**c**) Schematic of example proteomic workflow for mapping PPI. PL enzymes fused to the bait of interest and a spatial reference control are expressed in separate samples. Biotinylated proteins from each sample are enriched and analyzed via quantitative mass spectrometry. Proteins that preferentially interact with the bait of interest can be identified by ratiometric analysis.

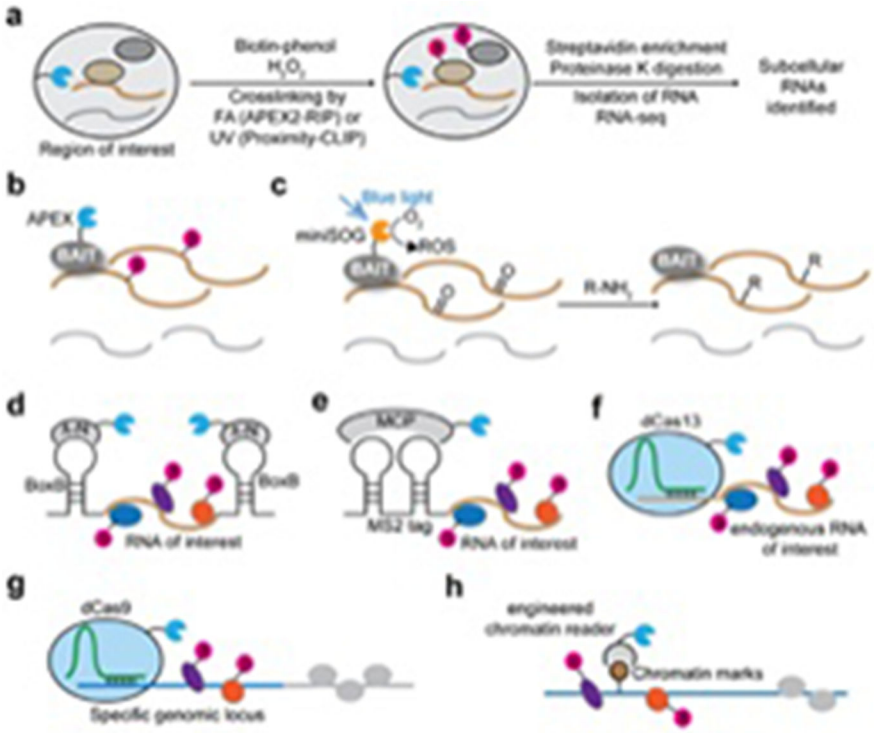

Figure 2.

PL-based methods to investigate protein-nucleic acid interactions. (**a**) Schematic of APEX-RIP and Proximity-CLIP. APEX targeted to a specific subcellular location catalyzes the biotinylation of proximal proteins, and the RNA-protein interactions are subsequently crosslinked by either UV or formaldehyde (FA). The subcellular RBP-occupied RNA can be captured via streptavidin-based enrichment of the biotinylated RBPs. (**b**) Schematic of APEX-seq. APEX directly biotinylates proximal RNA (yellow), but not distal RNA (grey), of a protein bait. (**c**) Schematic of Cap-seq. Upon blue light illumination, miniSOG generates ROS that react with guanine nucleobases in RNA. The photo-oxidation intermediates are intercepted by amine probes (R-NH2) to form covalent adducts. (**d**) Schematic of RaPID. An RNA of interest is tagged with a BoxB aptamer to recruit a fusion protein of λ-N and a promiscuous biotin ligase, which can biotinylate associated RBPs. (**e**) PL strategies based on MS2 tags and MCP to capture RBPs associated with an RNA of interest. (**f**) dCas13-based PL strategies to biotinylate RBPs associated with an endogenous RNA of interest. (**g**) dCas9-based PL strategies to biotinylate DNA-binding proteins at specific genomic locus. (**h**) Schematic of ChromID. BASU is fused to engineered chromatin readers that can specifically recognize particular chromatin marks, leading to the biotinylation of chromatin-binding proteins.

Table 1.

Overview of PL enzymes.

Enzyme	Type	Size (kDa)	Labeling time	Approx. Labeling radius (nm)	Modification sites	Advantages	Limitations
APEX	Peroxidase	28	1 min	20	Tyr; Trp; Cys; His	High temporal resolution; versatility for both protein and RNA labeling	Limited application *in vivo* due to the toxicity of H_2O_2
APEX2	Peroxidase	28	1 min	20	Tyr; Trp; Cys; His	High temporal resolution; versatility for both protein and RNA labeling	Limited application *in vivo* due to the toxicity of H_2O_2
HRP	Peroxidase	44	1 min	20	Tyr; Trp; Cys; His	High temporal resolution; versatility for both protein and RNA labeling	Limited application *in vivo* due to the toxicity of H_2O_2; limited to secretory pathway and extracellular applications
BioID	Biotin ligase	35	18 h	10	Lys	Non-toxic for *in vivo* applications	Poor temporal resolution due to the low catalytic activity
BioID2	Biotin ligase	27	18 h	10	Lys	Non-toxic for *in vivo* applications	Poor temporal resolution due to the low catalytic activity
BASU	Biotin ligase	29	18 h	10	Lys	Non-toxic for *in vivo* applications	Poor temporal resolution due to the low catalytic activity
TurboID	Biotin ligase	35	10 min	10	Lys	Highest activity biotin ligase; non-toxic for *in vivo* applications	Potentially less control of labeling window due to high biotin affinity
miniTurbo	Biotin ligase	28	10 min	10	Lys	High activity; non-toxic for in *vivo* applications; smaller than TurboID	Lower catalytic activity and stability compared to TurboID

Table 2.

Sample of proximity labeling studies for mapping PPIs.

PPI Category	Note(s)	Enzyme	Bait(s)	Reference
Protein aggregates	Insoluble complexes by definition	BioID	TDP43 aggregates	74
Nuclear membrane and nuclear structures	Low solubility complexes due to membrane function and/or complex size	BioID	Lamin A	21
		BioID	Lamin B1	72
		BioID	Various nuclear transport receptors	73
		BioID2	Lamin A, Sun2	71
Enzyme-substrate interactions	Low-affinity/transient interactions due to enzyme turnover	BioID	Hippo pathway (including Mst1/Mst2 kinases)	50
		BioID	p190/p210 BCR-ABL kinases	55
		APEX2	p38 MAPK	48
		BioID2	p38 MAPK	49
		BioID	SCF E3 ligases	53
		APEX2	KREP, Kelch E3 ligase adaptors	54
		BioID	ClpP protease	142
Other signaling pathways	Low-affinity/transient interactions	BioID2	TLR9, MYD88 (NFκB pathway)	76
		BioID2	KRas4B	77
		APEX2	$Ca_v1.2$ GPCR (adrenergic pathway)	51
Intracellular sorting	Transient interactions, low-affinity interactors for trafficking machinery	BioID	Dynein machinery	143
		BioID	Golgin-97, Golgin-245	144
		APEX2	LAMP1	58
		BioID2	Golgi glycosyltransferases	145
Dynamic processes	Utilized APEX for minute-scale interactome capture	APEX2	DOR (GPCR)	41
		APEX2	AT1R, β2AR (GPCRs)	52
		APEX2	MOR (GPCR)	87
		APEX2	Fzd9b (GPCR)	81
		APEX2	Gal8, Gal3, Gal9	146
		APEX2	TssA (bacteria)	60
In vivo PL in plants	Proximity labeling in plant systems	BioID	OsFD2	147
		BioID	HopF2	65
		BioID	AvrPto	95
		TurboID	N NLR	96
		TurboID	FAMA	33
In vivo PL in other organisms	Biotin-ligase based *in vivo* PL	BioID	Sun1 (Dictyostelium)	97
		BioID	CDK5RAP2 (Dictyostelium)	98
		BioID	ISP3 (*Toxoplasma gondii*)	100

PPI Category	Note(s)	Enzyme	Bait(s)	Reference
		BioID	Cyst wall proteins (*Toxoplasma gondii*)	103
		BioID	TbMORN1 (*Trypanosoma brucei*)	99
		BioID	TbPLK (*Trypanosoma brucei*)	101
		BioID	Parasitophorous vacuole (*Plasmodium falciparum*)	102
		BioID	c-MYC (mouse xenograft)	68
		BioID	Gephyrin (mouse)	67
		TurboID	Rmt3 (*Schizosaccharomyces pombe*)	62
		TurboID	Dcp-1, Drice, Dronc (*Drosophila melanogaster*)	104

Table 3.

Comparison of protein-nucleic acid approaches.

RNA	**RNA centric – (identifying the proteins interacting with an RNA of interest)**				
	Method	**Brief Description**	**Pros**	**Cons**	**Ref**
	RAP-MS, PAIR, TRIP, CHART, ChIRP	UV or formaldehyde cross-linking followed by RNA pulldown using biotinylated nucleic acid probes.	No genetic engineering or exogenous expression of components. UV crosslinking is highly specific for direct RNA-protein interactions.	Cross-linking, UV in particular, has low efficiency, requiring 10^8-10^9 cells. FA is less specific and results in protein-protein crosslinks that increase background. DNA probes must be optimized and can lead to nonspecific capture of RNAs or contribute to lower efficiency.	119-123
	MS2-Biotrap	UV cross-linking followed by RNA pulldown via MS2-coat protein interaction.	Improves the pulldown workflow by avoiding ASO capture.	Cross-linking is low efficiency, requiring 10^8-10^9 cells. Exogenous expression of MS2-tagged RNA may not recapitulate physiological concentrations or conditions	124
	RaPID	MS2-modified endogenous RNA recruits a coat protein-PL enzyme fusion to biotinylate proteins interacting with the RNA of interest.	Avoids crosslinking and associated problems. Enables direct biotinylation of interacting proteins. Can be applied *in vivo*.	Biotinylation of the general location necessitates spatial references (e.g. scrambled RNA control) to eliminate false positives. PL captures indirect interactors. Exogenous expression of MS2-target RNA may not recapitulate physiological concentrations or conditions. Biotinylated proteins may be proximal to the MS2 site and not the RNA in general, making this method better for shorter RNAs.	24,125,126
	CRUIS, CBRPP, CARPID, dCas13d-dsRBD-APEX2	Proximity labeling enzyme (PafA / BioID/ BASU/ APEX2) fusion to catalytically inactive dCas13 to biotinylate proteins interacting with an endogenous transcript.	Enables direct biotinylation of proteins interacting with endogenous RNA transcripts. *In vivo* compatible and can be easily engineered for different targets. Avoids crosslinking.	Incomplete localization of Cas 13 can produce high background. May require guide optimization, as well as spatial references (non-targeting guide) to account for non-specific labeling. PL captures indirect interactors. Biotinylated proteins are proximal to the guide RNA site, and not the entire target RNA in general.	126,129
	Protein centric - (identifying the RNAs interacting with a protein of interest)				
	Method	**Brief Description**	**Pros**	**Cons**	**Ref**
	CLIP-Seq, eCLIP, iCLIP, irCLIP PAR-CLIP, fCLIP	Cross-linking Immuno-Precipitation. There are many variations of the CLIP-seq protocol, but generally, crosslinking of proteins to RNA is carried out by UV (CLIP-seq), by UV using incorporated thiouridine (PAR-CLIP), or using FA (fCLIP). A protein of interest is isolated by antibody pulldown, and the covalently bound RNA is sequenced.	UV crosslinking is highly specific. Does not require genetic engineering or exogenous expression of components.	Can be difficult to obtain enough cross-linked RNA due to low efficiency of crosslinking, poor antibody pull-down, or low abundance of the RNA-RBP complex. Requires IP-grade antibodies.	106-111
	RIP-seq	Antibody pulldown of a protein of interest under non-denaturing conditions to recover the associated RNAs.	Higher RNA yield than CLIP. Simple protocol without genetic engineering or exogenous expression.	Lower signal to noise than CLIP, may capture indirect interactors, and has a higher chance of false positives due to FA crosslinking.	148
	RNA Tagging, TRIBE	RNA Tagging uses a poly-U-polymerase fused with the POI to	Does not require antibody purification.	Exogenous expression of RBPs can lead to false positives/	149,150

		extend poly uracil at the 3' end of proximal RNAs, which can be subsequently enriched using poly-A ASO capture. TRIBE uses ADAR fused with the POI and mediates A to I editing of interacting RNAs, which can then be identified by sequencing.	Does not require crosslinking.	negatives. RNA tagging may be biased towards 3' interactors.	
	APEX-RIP, Proximity - Clip	Proteins are biotinylated by APEX2 labeling, and RNA and Proteins are crosslinked by UV and 4SU (proximity CLIP) or FA (APEX-RIP). Streptavidin pulldown enables the enrichment of RNA of a specific subcellular location.	Does not rely on antibody purification. Can recover organelle or location- specific RNAs. UV crosslinking captures direct interactors.	Formaldehyde crosslinking results in poor specificity, which can be overcome by UV crosslinking at the expense of efficiency. Adapting this method to RBP-specific capture necessitates IP-grade antibodies or genetic-tagging of RBP of interest.	113,114
	APEX-seq, CAP-seq	Proximity labeling of RNAs directly by a proximity labeling enzyme enables the enrichment of RNA that interacts with a POI or located in specific subcellular locations.	Direct labeling of RNA improves workflow, specificity, and efficiency. Can be performed *in vivo*.	Proximity labeling can capture indirect interactors. Adapting these techniques to studying specific RBPs requires exogenous expression of the RBP-PL fusion protein.	115-118
DNA	**Protein centric – (identifying the DNAs associated with a protein of interest)**				
	Method	**Brief Description**	**Pros**	**Cons**	**Ref**
	ChIP-Seq	Chromatin Immuno-Precipitation. Antibody pulldown of a POI under non-denaturing conditions allows the identification of associated DNA fragments.	Widely adopted and straight-forward protocol, relatively unbiased, and does not require exogenous expression.	Requires IP-grade antibodies.	131
	ALaP	APEX2 is fused to a protein of interest to detect associated DNA.	Does not require antibody pulldown.	Requires a spatial reference to improve SNR. Exogenous expression of fusion protein may not reflect physiological conditions.	132
	Chromatin modification centric – (identifying proteins associated with a specific chromatin modification)				
	ChromID	Fusion of BASU promiscuous biotin ligase to 'reader domains' that specifically bind to chromatin modifications (e.g. H3K4me3), which enables the identification of proteins associated with specific chromatin modifications.	Direct labeling of proteins associated with a specific chromatin modification.	Overexpression of the reader domains may perturb the normal occupancy of chromatin modifications. Proximity labeling may require a spatial reference to improve SNR.	139
	DNA centric – (identifying proteins associated with a specific DNA sequence)				
	RIME, ChIP-MS	DNA-protein crosslinking followed by immunoprecipitation.	Enable the assessment of chromatin-bound protein complexes.	Crosslinking has low efficiency and may result in false positives.	137,138
	(APEX-DNA Binding Protein fusion)	Fusion of a PL enzyme to a DNA binding protein (DBP) enables the labeling of proteins associated with the DNA-binding site of the DBP.	Does not require crosslinking or antibody pulldown. Can be performed *in vivo*.	May require a spatial reference to improve SNR. Exogenous expression of a DNA binding protein can perturb the studied system.	39
	CASPEX, C-BERST	APEX2-dCas9 fusion proteins are expressed in a cell along with targeting guides to enable labeling of proteins associated with a specific DNA sequence.	Easily reprogrammed and simple protocol. Proteins can be directly enriched and avoids crosslinking or IP.	May requires a spatial reference (e.g. non-targeting guide) to improve SNR. Exogenous expression of Cas9 can perturb the studied system.	134,135

Abbreviations: PL – proximity labeling; FA – formaldehyde; RBP – RNA binding protein; SNR – signal to noise ratio; POI – protein of interest.