

Uncovering the In Vivo Proxisome Using Proximity-Tagging Methods

Yoann G Santin

▶ **To cite this version:**

Yoann G Santin. Uncovering the In Vivo Proxisome Using Proximity-Tagging Methods. BioEssays, 2019, 10.1002/bies.201900131 . hal-02341404

HAL Id: hal-02341404
https://amu.hal.science/hal-02341404

Submitted on 31 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Uncovering the *in vivo* proxisome using proximity-tagging methods

Yoann G. Santin

Laboratoire d'Ingénierie des Systèmes Macromoléculaires, Institut de Microbiologie de la Méditerranée, Aix-Marseille Université – CNRS UMR7255, 31 Chemin Joseph Aiguier, CS70071, 13402 Marseille Cedex 09, France

Email: ysantin@imm.cnrs.fr

The ORCID identification number for the author: 0000-0003-1093-7465

Twitter link: https://twitter.com/santin_yoann

Keywords: Protein-protein interaction, *in vivo* proximity labeling, APEX, BioID, PUP-IT, proxisome, bacteria

Summary

The development of new approaches is critical to gain further insights into biological processes that cannot be obtained by existing methods or technologies. The detection of protein-protein interaction is often challenging, especially for weak and transient interactions or for membrane proteins. Over the last decade, several proximity-tagging methodologies have been developed to explore protein interactions in living cells. Among those, the most efficient are based on protein partner modification, such as biotinylation or pupylation. Such technologies are based on engineered variants of enzymes like peroxidases or ligases that release reactive molecules, in the presence of specific substrates, that bind surrounding proteins. Fusing a protein of interest to these enzymes allows the definition of an unbiased “proxisome”, that is all of the proteins in interaction or in close vicinity of the protein of interest. Here, I describe the different proximity-labeling tools available and provide a comprehensive comparison to discuss advantages and limitations.

1. Introduction

By analogy with a theater play, biological processes require different actors assigned to perform specific tasks in space and time. In living cells, such actors are mostly proteins that "communicate" between each other. Communication is achieved through different types of interactions such as weak, transient, stable or long interactions, resulting in specific biological effects. Protein-protein interaction (PPI) can lead to conformational effects which enable the formation of a structural complex, activate or inactivate a protein, create a new biding site for the interaction with other binding partners or a substrate, serve as regulatory pathway, or allow subcellular relocalization.

Detection of PPIs, in space and time, is therefore critical for deciphering each step of a biological process. A number of methods for assaying PPIs *in vivo* have been developed and are routinely used in laboratories. Co-immunoprecipitation (Co-IP) represents one of the most standard methods of identifying interacting partners *in vivo*[1,2]. Briefly, a protein of interest (POI) is immunoprecipitated from a cellular protein extract by using specific antibodies immobilized on beads (Fig. 1a). While unbound proteins are washed out, binding partners are co-precipitated and can be visualized and identified by Western Blot or mass spectrometry analysis. It is worthy to note that Co-IP is a variant of the pull-down assay, which used a tagged bait protein to capture protein complexes instead of antibodies. Tandem affinity purification (TAP) is also a systematic approach to detect PPIs at near proteome-scale under *in vivo* conditions. TAP consists in two consecutive purifications by using two different tags fused to a bait protein, then considerably reducing the amount of nonspecific contaminants[3,4]. Additional methods such as two-hybrid systems, including bacterial (BACTH) and yeast (Y2H) two-hybrid assays (see[5] for comparative review), are powerful genetic approaches to characterize PPIs in native or near-native context[6–13]. POIs are fused to the two isolated fragments from the *Bordetella pertussis* adenylate cyclase[6,8–10,13] or from the Gal4

transcription factor[7,11,12] for BACTH and Y2H respectively (Fig. 1b). Physical association between tested proteins pairs restores the activity of the adenylate cyclase or Gal4 that can be visualized by a transcriptional-activated reporter. Other interaction-mediated reconstitution-based methods exist, such as the TOXCAT or GALLEX systems for studying transmembrane helix-helix oligomerization in a natural membrane environment[14–16]. With a similar conceptual approach, Förster resonance energy transfer (FRET) is a well-suited biophysical method for the investigation of PPIs that occur between two proteins positioned within 10 nanometers of each other, allowing study of molecular interactions in real time[17,18]. In brief, FRET relies on the energy transfer from a donor fluorescent molecule (i.e., diminution of fluorescence intensity) to an acceptor when these two probes are separated by 10 nanometers or less [19–22] (Fig. 1c). Then, PPIs or conformational changes within a protein (if the two probes are fused to the same protein) can be visualized in real time, by using a fluorometer or a fluorescence microscope. Furthermore, this method can be extend to several practical applications in biology such as the detection of protein cleavage, changes in micro-environment or quantitative analysis of protein interactions (see[23,24] for comprehensive reviews).

Although these methods have improved our knowledge about biological processes by the characterization of number of PPIs, some important limitations are still present. Due to the nature of stringent purifications, co-precipitation based-methods (i.e., Co-IP, pull down and TAP) are not reliable to detect weak or transient PPIs, leading to a lot of false negatives[25]. In two-hybrid systems, proteins are produced from multi-copy plasmids and not at the physiological concentration in the cell, which may cause unbalanced stoichiometry and effects on PPIs, or toxicity within the cell[26]. Proper re-assembly of the reporter is dependent of the spatial association between the two tested proteins, hence steric effects may occur and prevent detection of PPI. Screening only soluble domains of membrane proteins, which need the cell membrane for proper folding, can lead to false negative in PPIs. Further, bait and prey proteins

are artificially put into the same subcellular compartment (i.e., the bacterial cytosol or the yeast nucleus for BACTH and Y2H respectively), which may be different from the native compartment of the POIs. Hence, protein pair could interact in a compartment where they normally may not encounter each other or not interact if the physicochemical conditions are too different from the native compartment(s). Finally, fluorescence based-methods suffer from inherent physical and technical limitations. In FRET, since the distance between the donor and the acceptor must be within the Förster radius (i.e., ~10 nm), poor positioning of fluorescent probes in protein pair might lead to defect in detection of PPI. In addition, spectral leakage, that correspond to the direct excitation of the acceptor and causing false positive detections, donor fluorescence background, photobleaching and trouble for measurements make real PPI detection quite difficult[27].

Deciphering subtle PPIs in living cells is a challenge that cannot be achieved by those conventional methods. Detection of weak and transient interactions, which constitute the major part of PPIs in dynamic processes, interactions with membrane proteins and more importantly, the detection in true *in vivo* context (i.e., both in native cell and at the chromosomal expression level) must be improved to understand the complexity of biological processes. Over the last decade, proximity-tagging approaches have been developed to explore these questions. Based on protein partner modifications, proximity-tagging methodologies allow the definition of an unbiased “proxisome”, that is all of the proteins in direct interaction or in close vicinity of the POI, in native conditions and without loss of elusive PPIs. Currently, three major proximity-tagging approaches are used: the proximity-dependent biotin identification (BioID), the APEX proximity-dependent biotin labeling and the recent pupylation-based interaction tagging (PUP-IT). Here, I detail the principles of these proximity-tagging methodologies and their different applications in biology. Finally, I provide a comprehensive comparison to allow the definition of the most appropriate approach for your studies.

2. Proximity-dependent biotin identification (BioID)

BioID is a useful proximity-dependent labeling method to screen both interacting and neighboring proteins in their *in vivo* context. BioID assay is based on the promiscuous biotinylation generated by a modified variant of BirA[28,29]. BirA is a 35-kDa bacterial biotin ligase that catalyzes the biotinylation of BCCP, a subunit of the acetyl-CoA carboxylase in *Escherichia coli*[30,31]. By using endogenous biotin and ATP, BirA produces biotinoyl-5'-AMP molecules and biotinylates BCCP with an exquisite specificity. To perform proximity-dependent labeling, the R118G mutant BirA protein ($BirA^{R118G}$ or BioID protein) is used. This variant, which normally binds ATP, displays a low affinity for biotinoyl-5'-AMP compared to the WT BirA[32]. Hence, biotinoyl-5'-AMP molecules are released from BioID protein and covalently react with proximate proteins, and specifically on lysine residues, in a range of approximately 10 nm[33,34]. For BioID assay, promiscuous biotinylation is induced in cells producing BioID fusion protein by adding a supraphysiological concentration of biotin in the medium during an optimal labeling period of 16-18h[34–36] (Fig. 2). Biotinylated proteins, which are candidate interactors for the POI, are then enriched on streptavidin beads and identified by mass spectrometry.

Recently, significant efforts have been made to improve BioID assay in living cells. Screen for substitutions in the biotin binding site that modulate BirA activity revealed that R118K substitution reduced biotin affinity and increased the biotinoyl-5'-AMP release[37]. In 2016, Kim and colleagues tested the potential of BirA protein from the thermophilic bacterium *Aquifex aeolicus*[38]. This variant, called BioID2, is 8-kDa smaller than BioID (i.e., 27 kDa) and requires less biotin supplementation for proximity labeling[38]. Interestingly, assaying for dimerization-dependent protein interactions was investigated by split-BioID, where inactive BioID-fragment complementation pair was generated as split-reporter and forms a functional BioID protein that biotinylates substrates and other proximate proteins upon

heterodimerization[39]. Finally, by using yeast display-based directed evolution, the Ting lab recently engineered two enhanced BioID variants: the 35-kDa TurboID and the 28-kDa miniTurbo[40]. These variants display a remarkable labeling power, dropping from 16-18h to only 10 min labeling period with similar size and specificity than standard BioID[40]. Furthermore, proximity labeling activity of TurboID and miniTurbo is conserved in different ranges of pH and at low temperature[40] by contrast to BioID and BioID2 proteins[38]. Taken together, such improvements greatly extended the potentiality of the BioID system for assaying proxisome detection in plethora of contexts.

BioID was initially developed for the study of the *in vivo* nuclear lamin-A proxisome in mammalian cells[34]. Up to now, BioID has been successfully used in diverse contexts. In mammalian cells, BioID was applied to probe architecture of different subcellular locations such as the nuclear pore complex[33,41,42], focal adhesion complexes[43] or the centrosome-cilium interface[44], and also to get insights about molecular mechanisms such as mitochondrial proteostasis[45,46], autophagy[47], mitosis [48–50] or viral infection[51]. BioID was also elegantly adapted to investigate RNA-protein interactions that play important roles in cellular functions and diseases. Named RaPID (for RNA–protein interaction detection), this method enables the detection of proteins which interact with a specific RNA[52]. Briefly, the RNA of interest is flanked by bacteriophage BoxB stem loops which bind by the λN peptide with high affinity[53]. The λN peptide is fused to the N-terminus of BioID protein. Then, BoxB stem loops recruit the λN-BioID fusion protein and thereby biotinylating proteins bound to the RNA of interest in living cells[52]. In addition, authors engineered BASU, a modified biotin ligase from *Bacillus subtilis* that displays faster kinetics (>1000-fold) and increased signal-to-noise ratio (>30-fold) compared to the standard *Escherichia coli* BioID protein. Although less active than miniTurbo and TurboID[40], BASU enabled sufficient labeling of proteins for RaPID study in living cells[52]. Finally, BioID was used for proximity labeling *in planta*[54,55] and directly on animals

such as in mouse[56] and recently in fruit flies (*Drosophila melanogaster*) and worms (*Caenorhabditis elegans*)[40].

3. APEX proximity-dependent biotin labeling

APEX is a spatially resolved proteomic mapping assay to determine *in vivo* proxisome in dynamic systems in living cells. By contrast to BioID, APEX-dependent labeling is based on the activity of the engineered enhanced hemic ascorbate peroxidase APEX, a 27-kDa monomeric derivative of protein from soybean[57,58]. Briefly, APEX catalyzes the H_2O_2-dependent oxidation of phenol derivatives (the biotin-phenol, BP, being the most specific substrate for APEX) into short lived phenoxyl radicals (<1 ms) that covalently react in a small labeling radius (<20 nm) with electron-rich amino acids such as Tyr, Trp, His and Cys[58]. It has been recently established that tyrosine residues are the principal site of biotinylation by APEX with more than 98 % labeling observed[59]. Similar to BioID workflow, APEX labeling starts with the construction of a functional fusion between APEX and the POI. After incubation with BP, cells producing APEX fusion are treated with a 1-min pulse of hydrogen peroxide enabling APEX-dependent biotinylation of proximate proteins in a short labeling radius (Fig. 3)[58,60]. Subsequently, cells are lysed and biotinylated proteins are purified and identified by mass spectrometry analysis.

Due to its low activity when expressed at physiological level, APEX protein was further improved to enhance its sensibility in living cells. By using directed evolution, a single point substitution (A134P), located in a loop at the vicinity of both heme and the aromatic substrate–binding site was found to confer great improvements in stability, kinetics, heme binding and resistance to high H_2O_2 concentrations[61]. Named APEX2, this variant is much more active than APEX in the cell and is consequently better for proteomic mapping[61]. Very recently, a split-APEX (sAPEX) system has been developed[62]. Two fragments called “AP” (a 200-amino-

acid N-terminal fragment) and "EX" (a 50-amino-acid C-terminal fragment) are each inactive but give peroxidase activity upon reconstitution. Authors proposed that the sAPEX system could be useful to address question about interaction-dependent proximity labeling at higher spatiotemporal resolution. Hence, sAPEX reconstitution can be applied for mapping specifically proxisome of target nucleic acids (e.g., mRNA, non-coding RNA or genomic locus) by considerably reducing background caused by the activity of APEX2 that is not bound to the target of interest[63,64], or to capture molecular composition at the interface of two organellar structures such as at mitochondria-ER contact[62].

Over the past 6 years, many applications of APEX labeling have emerged to map proxisome within specific locations and/or in dynamic molecular processes. Extensively used in eukaryotic cells, APEX has been applied for investigating complex proteomes such as in human or in *Drosophila melanogaster* mitochondria[58,60,61,65–67], in mammalian cilia[68] or in yeast Golgi[69]. APEX proximity biotinylation was also used for exploring dynamic macromolecular complexes within specific genomic loci or for RNA subcellular location. In order to define the human mitochondrial nucleoid proteome, fusion between APEX2 and the mitochondrial DNA helicase Twinkle was constructed and revealed new nucleoid-associated proteins[70]. Last year, two independent groups reported the development of GLoPro (for genomic locus proteomics) and C-BERST (for dCas9 biotinylation at genomic elements by restricted spatial tagging), a combination of CRISPR-based genome targeting with APEX2-dependent proximity labeling for defining proteome at specific genomic locus[63,64]. While receptor signaling constitutes a critical process involved in numerous physiological regulations, tracking those processes requires high spatio-temporal resolution and high throughput, missing in current methods. Back-to-back papers from Kruse and Krogan labs elegantly demonstrated the application of APEX assay to G-protein-coupled receptors signaling by tracking interactions at several different time points upon addition of activating ligand and within a native cellular

environment[71,72]. They finally concluded that APEX is a powerful method for spatially resolved quantitative analysis of signaling, applicable to any other signal transduction monitoring[71,72]. Designed at the beginning for proteomic mapping, APEX has been recently used for transcriptomic mapping assay. Spatial organization of RNA within the cell has been at first determined by an indirect approach combining APEX proximity labeling with RNA-protein chemical crosslinking[73]. As this approach was inadequate for spatial specificity in unenclosed cellular locations, authors introduced APEX-seq, a direct APEX-catalyzed labeling of RNA combined with RNA-seq. This direct approach was successfully used in nine distinct subcellular compartments allowing determination of a the RNA atlas in human cell[74] and recently the organization of translation initiation complexes[75]. Finally, Zhou and colleagues recently developed promising biotin-aniline and biotin-naphthylamine probes (Btn-An and Btn-Nap, respectively), a new class of phenolic compounds with significantly higher reactivity towards nucleic acids[76]. Such developments open the field of new broad applications for *in situ* transcriptomic assay to shed light on RNA functions in cell physiology.

Over the last year, APEX biotinylation assay was also performed in different microbial pathogens such as for the mapping of the inclusion membrane proteins in *Chlamydia trachomatis*[77–80], in *Mycobacterium smegmatis* periplam[81], in the diplomonad fish parasite *Spironucleus salmonicida*[82], and in enteroaggregative *Escherichia coli* where stage-blocking mutations were used to define *in vivo* temporal contacts of the TssA protein during type VI secretion system biogenesis[83]. With its broad range of applications in diverse organisms and for labeling of different types of macromolecules, APEX proximity-dependent biotin labeling constitutes a powerful tool for cellular processes that require high spatial specificity and short labeling windows.

4. Pupylation-based interaction tagging (PUP-IT)

As a new proximity tagging method, the PUP-IT system was developed from the ubiquitin-like process present in actinobacteria[84–86]. The pupylation-based modification was extensively studied in *Mycobacterium tuberculosis* for its implication in bacterial proteasome, a cellular system required for causing disease[85,87,88]. In this system, the Pup ligase PafA uses ATP to transfer an activated Pup(E) substrate to a lysine of a target protein[85,89,90]. Thereafter, the Pup(E)-protein complex is targeted to the mycobacterial proteasome for unfolding and degradation[85,90,91]. Taking advantage of this system, Liu and colleagues adapted the Pup system from *Corynebacterium glutamicum* as a PUP-IT tool to define membrane and intercellular PPI[92]. They first determined the potential of PafA to be a suitable proximity-tagging enzyme and revealed that compared to BioID and APEX substrate, activated Pup(E) does not freely diffuse from PafA, suggesting that very close proximity between the bait protein and their partners is required for proper pupylation. They then tested *in vivo* PUP-IT by fusing PafA to the membrane receptor CD28 to identify its cytosolic tail interactome (Fig. 4a). By inducing the production of a biotinylated carboxylase-fused Pup(E) protein (bio-Pup(E)), which was under the control of the locked TET-ON system[93], the authors initiated the pupylation-based labeling and successfully enriched known CD28-interacting proteins such as ITK, p85 or LCK[92]. They finally explored the possibility of extracellular PUP-IT labeling. To avoid long exposure of PafA at the cell surface, authors induced the external heterodimerization of CD28 and PafA by using the FKBP-rapamycin-FRB system[94] (Fig. 4b). Upon reconstitution, the addition of ATP and truncated biotin-fused Pup(E) DE28 peptides (bio-DE28) enabled cell surface labeling suggesting that PUP-IT can be used for extracellular pupylation[92]. While the PUP-IT has been only used in this study, it represents a promising method for assaying *in vivo* PPIs especially involved in cell-to-cell communication, cell-to-cell adhesion, or host-pathogen interactions.

5. What is the best approach (for you)?

Proximity-tagging methods can be used to circumvent inherent technical limitations that conventional approaches for assaying *in vivo* PPIs suffer from. However, each of them display differences which are both important or not depending on the nature of the study. In this section, I will compare the specifications of each method to help defining the most suitable approach for your studies.

5.1. Advantages compared to conventional methods

As mentioned in the introduction, proximity-tagging methods present many benefits compared to conventional methods. First, interactome/proxisome of a POI can be defined under native expression and stoichiometry by fusing tagging enzyme-coding gene at chromosomal locus of the bait POI. Because proximate proteins are covalently labeled before purification, transient, weak or hydrophobic interactions, which are lacking from some conventional methods, can be easily detected. Further, promiscuous labeling of non-interacting-proximate proteins might be useful to understand the environment of the POI and hence, to get new insights about its role and/or its subcellular location in the cell. Interestingly, it has been shown that proxisome can be expanded or reduced by modifying the length of the linker between the bait and the tagging protein[38]. As previously described, labeling molecules (i.e., biotinoyl-5'-AMP for BioID, biotin-phenoxyls for APEX and Pup(E) for PUP-IT) have been shown not to cross plasma membrane[34,58,92], therefore proteome definition into cellular organelle or near to the membrane might be useful to determine topology of transmembrane proteins[58,67,95]. Finally, the use of proximity-tagging methods in several biological contexts, ranging from the bacterial cell to the living animals, makes them powerful versatile tools for investigating endogenous PPIs.

5.2. Common features to keep in mind before starting

As biological applications, proximity-tagging methods can be limited or be not suitable in specific context.

First, characterization of the bait protein is a basic, but important prerequisite to confirm that its location and its function in the cell are conserved in fusion with the tagging enzyme. Hence, the use of proximity labeling method without preliminary results about the POI is strongly discouraged.

While biotin is commonly used for detection and enrichment, it is necessary to examine endogenous biotinylation occurring in your model organism. Overrepresentation of endogenous biotinylated proteins might be detrimental by preventing enrichment of proximity labeled-proteins, especially if the POI is thought to have only few partners. For example, in *Mycobacterium smegmatis*, APEX2-dependent biotinylated proteins were confounded by endogenous biotinylated proteins[81]. To address this problem, authors successfully designed two alternative labeling molecules: tyramide alkyne and tyramide azide, which can be used both for detection and enrichment via a copper-catalyzed alkyne/azide cycloaddition "click" reaction[81]. Thus, the design of new probes could be a good option to circumvent this type of issue in specific organisms.

Proximity labeling is associated with the presence and the accessibility of reactive residues such as lysine in BioID[28,34] and PUP-IT[92] assay, and mostly tyrosine in APEX proximity-labeling[58,59]. Consequently, the efficiency of labeling is dictated by the composition and the folding of proximate proteins. The absence of labeling does not therefore means a lack of interacting or proximity partners, especially for proteins that lack of exposed Tyr/Lys. By contrast, abundance of biotinylated proteins cannot be used to quantitatively assess the presence or the strength of a putative interaction.

Except for PUP-IT system that requires contact for labeling[92], labeling radius in BioID and APEX assay is not a fixed value but instead a "cloud" or a "labeling gradient" where contours depend on the physicochemical micro-conditions such as temperature or pH[28,60]. By defining a "proxisome", it could be necessary to determine whether identified biotinylated proteins are in direct interaction or in close proximity of the POI, especially if the POI is included within a multiprotein complex.

Although weak, labeling of endogenous proteins present at many nanometers away from tagging enzyme might be detected by MS analysis, especially in non-membrane-enclosed cellular regions. Hence, efficient discrimination from proximate biotinylated proteins to the cellular background is essential. To address this question, most studies combined proximity-tagging methods with quantitative proteomics such as stable isotope labeling by amino acids in cell culture[43,60,63,66,95–97] (SILAC), tandem mass tagging[64,70] (TMT), or label-free quantification[92] (LFQ). Schematic pipelines of these approaches applied to proximity-tagging methods are represented in Fig. 5. Although there is no any report for preferentially using one approach compared to another in the proximity-dependent labeling context, technical specifications should be considered before starting your study. For example, SILAC and TMT are much more expensive than LFQ but are more accurate for proteomic quantification. Because LFQ is a non-ratiometric approach, there is no limit in multiplexing (i.e., the number of conditions that can be tested simultaneously) by contrast to SILAC and TMT, limited to 5-plex and 10-plex respectively. SILAC requires metabolically active cells to incorporate labels whereas TMT can be directly performed on extracted peptides. Still, while LFQ presents a better proteome coverage, it requires significantly longer data acquisition time compared to SILAC or TMT. For the choice of the most suitable quantitative approach, see[98] for recent comprehensive comparison.

5.3. Comparative of the three approaches

To discriminate among those proximity-tagging methods, comparison of the main features could be important for specific biological studies. In this context, I decided to distinguish TurboID/miniTurbo assay (hereafter, (mini)TurboID assay) from BioID assay due to striking differences in labeling process. All of the data are summarized in Table 1.

APEX2, BioID2 and miniTurbo are relatively small proteins with a molecular mass of about 27 kDa[38,40,58,61] while Turbo-ID, BioID and PafA proteins are larger, with a molecular mass of about 35 kDa[40], 35 kDa[34] and 54 kDa[92] respectively. Because the tagging-enzyme is fused to the bait protein, its molecular mass can affect the localization and/or the function of the bait protein in the cell. Thus, preliminary results about your POI can be useful to wisely determine where you have to fuse the tagging-enzyme (i.e., in N or C-terminus of the POI), the nature of the linker, in terms of residues and length, and the most suitable enzyme.

In BioID and (mini)TurboID assays, tagging-enzyme uses native ATP and exogenous input of biotin as substrates for proximity labeling[34,40]. Since biotin is essential for life, its incorporation in cells is ensured by specialized transporters[99]. By contrast, APEX2 uses H_2O_2 and non-natural phenol derivatives as substrates[58,76,81]. Although in mammalian cells no specific treatment is used for biotin-phenol incorporation, it has been reported that cell permeability can limit biotin-phenol incorporation in yeast, preventing efficient labeling[69]. In *E. coli*, biotin-phenol uptake is in part mediated by the major γ-proteobacterium biotin transporter YigM (Santin et al., unpublished data). Therefore, biotin-phenol incorporation requires optimization for each organism. Finally, as non-natural compounds, whether phenol derivatives affect cell growth needs to be examined for each case. During labeling, H_2O_2 treatment could affect the general cell oxidative state and cause cellular stress. Hence, APEX assay is perhaps not suited for studies applied to oxidative stress or protein repair. In PUP-IT assay, as Pup(E) derivative-peptides are not commercially available[92], purification of these

374 substrates for extracellular labeling could be a limitation. However, Pup(E) substrates can be
375 genetically expressed in cells, and hence PUP-IT assay might be useful in organisms in which
376 substrate uptake is limiting.
377 To perform efficient *in vivo* labeling, the temperature in which the tagging-enzyme is active
378 has to fit with the growth temperature of the studied organism. In contrast to the BioID protein
379 that has an optimal temperature of 37°C[34,38], BioID2 protein has optimal activity of 50°C,
380 yielding this enzyme well suitable for labeling in thermophilic conditions[38]. Whereas BioID2
381 retains highly efficient biotinylation at 37°C, both BioID and BioID2 proteins exhibit reduced
382 activity below 37ºC[38]. APEX2, miniTurbo and TurboID proteins have an optimal activity at
383 37°C. Because these proteins were evolved from directed evolution in yeast (growing at 30°C),
384 they retain high activity at 30°C[40,61]. In addition, miniTurbo and TurboID have been
385 successfully used in worms and flies which grow at 20°C and 25°C, respectively. While PafA
386 has only been tested in mammalian cells[92] (37°C), the range of temperatures in which it retains
387 its activity has to be determined.
388 An important aspect of the proximity labeling is the temporal resolution of the reaction, both in
389 the control of the labeling initiation and in the labeling time. For example, definition of
390 signaling networks upon ligand-induced activation requires high temporal resolution[71,72] in
391 contrast to studies about proteome definition in specific subcellular locations[33,41,44,58,60,65,67].
392 In BioID and (mini)TurboID assays, the enzyme is active just upon translation and can use
393 endogenous biotin and ATP for proximity labeling, even before correct localization of the
394 fusion protein[34,40]. In the PUP-IT assay, even though Pup(E) expression might be monitored,
395 invariable delay exists between induction and PafA mediated-pupylation[92]. Hence, the
396 initiation of the labeling process may not be precisely controlled in these approaches. Finally,
397 as APEX-proximity labeling is nearly instant[58,60,61], temporal control of the labeling can be
398 achieved by addition of H_2O_2 in the culture at specific time points. APEX assay is therefore

perfectly suited for taking snapshots of PPIs, for example in cellular response upon environmental changes[71,72] or during assembly of macromolecular complexes[83]. With slow tagging kinetics, the BioID and PUP-IT assays require 16-18 hours or 24-36 hours of labeling time for efficient biotinylation/pupylation, respectively[34,92]. These approaches are hence not suited for capturing transient protein interactions but rather provide a history of protein associations over long periods of time, as used in cell cycle[48–50] or in viral infection[51]. TurboID and miniTurbo were designed to improve BioID assay by reducing the labeling time from 16-18 hours to 10 minutes[40]. However, recent studies showed that this time might be adapted depending on the organism studied. While 10 min are sufficient in mammalian cells[40], TurboID-mediated biotinylation requires about 30 minutes to 3 hours in yeast[100], 4 hours in flies[40], 4 hours to several days in worms[40] and 12 hours in plants[101]. Optimization of the TurboID labeling time should therefore be done for each organism, especially to prevent toxicity via chronic endogenous biotinylation or endogenous biotin consumption due to the high activity of these enzymes[40].

Finally, BioID and (mini)TurboID assays have never been tested for extracellular labeling by contrast to APEX[102] and PUP-IT[92]. Nevertheless, as PUP-IT and BioID-derivative assays use a similar combination of ATP and substrates (Pup(E) and biotin, respectively), there is no reason why BioID and (mini)TurboID should not work for extracellular labeling.

6. Conclusions and prospects

In molecular biology, the major question is: who acts in that process? To address this question, many PPI-assaying methodologies have been designed for decades. However, technical limitations in detection of PPIs in living cells, such as weak, transient or hydrophobic interactions, prevent global comprehension. By providing both *in vivo* spatiotemporal information and high-throughput analysis, proximity-tagging methods represent the next

generation of *in vivo* PPIs assay. In addition to the major proximity-tagging methods presented in this review, other technologies exist including for example horseradish peroxidase-dependent labeling approaches (EMARS[103,104] and SPPLAT[105–107]), photoactivated *in vivo* proximity labeling[108], NEDDylator system[109,110] and methylation-based chromatin profiling approach DamID[111,112]. While proximity-tagging methods have been used in a variety of biological contexts, further challenges remain to be overcome. For example, in thermophilic organisms, such as archaea, where tools available for PPIs assay are limited[113], or in protein-lipid interactions, involved in many diseases including cancer, obesity, neurodegenerative disorders or cardiovascular pathologies[114]. However, over the last decade, the rapid development of these proximity-tagging methods, in specificity and spatio-temporal control, has revealed their potential to answer deeply questions in molecular biology. One may assume that other applications will be made in the near future to further unravel the molecular complexity in living cells.

Abbreviations

PPI, protein-protein interaction; POI, protein of interest; Co-IP, Co-Immunoprecipitation; TAP, Tandem Affinity Purification; BACTH, bacterial adenylate cyclase-based two-hybrid; Y2H, yeast two-hybrid assay; FRET, Förster resonance energy transfer; APEX, APEX proximity-dependent biotin labeling BioID, proximity-dependent biotin identification; PUP-IT, pupylation-based interaction tagging; SILAC, stable isotope labeling by amino acids in cell culture; TMT, tandem mass tagging; LFQ, label-free quantitation.

Acknowledgments

I would like to thank Eric Cascales for support, Dukas Jurėnas for language editing, Etienne Vanlioglu for discussions on mass spectrometry analysis, Camille Laberthonnière for helpful advices, the members of the Cascales research group for useful discussions, and Yamamoto Kadèrapé for encouragements. Research in Cascales laboratory is supported by the Centre National de la Recherche Scientifique, the Aix-Marseille Université, and grants from the

Agence Nationale de la Recherche (ANR-14-CE14-0006-02, ANR-17-CE11-0039-01), the Fondation pour la Recherche Médicale (DEQ20180339165), and the Fondation Bettencourt-Schueller. My PhD work is supported by a doctoral fellowship from the French ministry of higher education and research. All the figures were created with BioRender.

References

[1] J.-S. Lin, E.-M. Lai, *Methods Mol. Biol. Clifton NJ* **2017**, *1615*, 211.

[2] C. Lee, *Methods Mol. Biol. Clifton NJ* **2007**, *362*, 401.

[3] O. Puig, F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, B. Séraphin, *Methods San Diego Calif* **2001**, *24*, 218.

[4] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Séraphin, *Nat. Biotechnol.* **1999**, *17*, 1030.

[5] J. Mehla, J. H. Caufield, N. Sakhawalkar, P. Uetz, *Methods Enzymol.* **2017**, *586*, 333.

[6] A. Battesti, E. Bouveret, *Methods San Diego Calif* **2012**, *58*, 325.

[7] S. Fields, O. Song, *Nature* **1989**, *340*, 245.

[8] G. Karimova, J. Pidoux, A. Ullmann, D. Ladant, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 5752.

[9] G. Karimova, N. Dautin, D. Ladant, *J. Bacteriol.* **2005**, *187*, 2233.

[10] G. Karimova, E. Gauliard, M. Davi, S. P. Ouellette, D. Ladant, *Methods Mol. Biol. Clifton NJ* **2017**, *1615*, 159.

[11] J.-S. Lin, E.-M. Lai, *Methods Mol. Biol. Clifton NJ* **2017**, *1615*, 177.

[12] J. Mehla, J. H. Caufield, P. Uetz, *Cold Spring Harb. Protoc.* **2015**, *2015*, 425.

[13] S. P. Ouellette, G. Karimova, M. Davi, D. Ladant, *Curr. Protoc. Mol. Biol.* **2017**, *118*, 20.12.1.

[14] F. Hennecke, A. Müller, R. Meister, A. Strelow, S. Behrens, *Protein Eng. Des. Sel. PEDS* **2005**, *18*, 477.

[15] W. P. Russ, D. M. Engelman, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 863.

[16] D. Schneider, D. M. Engelman, *J. Biol. Chem.* **2003**, *278*, 3105.

[17] F. Festy, S. M. Ameer-Beg, T. Ng, K. Suhling, *Mol. Biosyst.* **2007**, *3*, 381.

[18] R. Roy, S. Hohng, T. Ha, *Nat. Methods* **2008**, *5*, 507.

[19] H. J. Carlson, R. E. Campbell, *Curr. Opin. Biotechnol.* **2009**, *20*, 19.

[20] L. He, T. D. Bradrick, T. S. Karpova, X. Wu, M. H. Fox, R. Fischer, J. G. McNally, J. R. Knutson, A. C. Grammer, P. E. Lipsky, *Cytom. Part J. Int. Soc. Anal. Cytol.* **2003**, *53*, 39.

[21] I. T. Li, E. Pham, K. Truong, *Biotechnol. Lett.* **2006**, *28*, 1971.

[22] B. A. Pollok, R. Heim, *Trends Cell Biol.* **1999**, *9*, 57.

[23] B. Hochreiter, A. P. Garcia, J. A. Schmid, *Sensors* **2015**, *15*, 26281.

[24] K. Okamoto, Y. Sako, *Curr. Opin. Struct. Biol.* **2017**, *46*, 16.

[25] T. Branon, S. Han, A. Ting, *Biochemistry* **2017**, DOI 10.1021/acs.biochem.7b00466.

[26] A. Brückner, C. Polge, N. Lentze, D. Auerbach, U. Schlattner, *Int. J. Mol. Sci.* **2009**, *10*, 2763.

[27] S. J. Leavesley, T. C. Rich, *Cytom. Part J. Int. Soc. Anal. Cytol.* **2016**, *89*, 325.

[28] E. Choi-Rhee, H. Schulman, J. E. Cronan, *Protein Sci. Publ. Protein Soc.* **2004**, *13*, 3043.

[29] J. E. Cronan, *J. Nutr. Biochem.* **2005**, *16*, 416.

[30] A. Chapman-Smith, J. E. Cronan, *Trends Biochem. Sci.* **1999**, *24*, 359.

[31] S. J. Li, J. E. Cronan, *J. Biol. Chem.* **1992**, *267*, 855.

[32] K. Kwon, D. Beckett, *Protein Sci. Publ. Protein Soc.* **2000**, *9*, 1530.

[33] D. I. Kim, K. C. Birendra, W. Zhu, K. Motamedchaboki, V. Doye, K. J. Roux, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E2453.

[34] K. J. Roux, D. I. Kim, M. Raida, B. Burke, *J. Cell Biol.* **2012**, *196*, 801.

[35] D. G. May, K. J. Roux, *Methods Mol. Biol. Clifton NJ* **2019**, *2008*, 83.

[36] K. J. Roux, D. I. Kim, B. Burke, D. G. May, *Curr. Protoc. Protein Sci.* **2018**, *91*, 19.23.1.

[37] L. T. Oostdyk, L. Shank, K. Jividen, N. Dworak, N. E. Sherman, B. M. Paschal, *Methods San Diego Calif* **2018**, DOI 10.1016/j.ymeth.2018.11.003.

[38] D. I. Kim, S. C. Jensen, K. A. Noble, B. Kc, K. H. Roux, K. Motamedchaboki, K. J. Roux, *Mol. Biol. Cell* **2016**, *27*, 1188.

[39] S. De Munter, J. Görnemann, R. Derua, B. Lesage, J. Qian, E. Heroes, E. Waelkens, A. Van Eynde, M. Beullens, M. Bollen, *FEBS Lett.* **2017**, *591*, 415.

[40] T. C. Branon, J. A. Bosch, A. D. Sanchez, N. D. Udeshi, T. Svinkina, S. A. Carr, J. L. Feldman, N. Perrimon, A. Y. Ting, *Nat. Biotechnol.* **2018**, *36*, 880.

[41] W. Xie, A. Chojnowski, T. Boudier, J. S. Y. Lim, S. Ahmed, Z. Ser, C. Stewart, B. Burke, *Curr. Biol. CB* **2016**, *26*, 2651.

[42] C. James, M. Müller, M. W. Goldberg, C. Lenz, H. Urlaub, R. H. Kehlenbach, *J. Biol. Chem.* **2019**, DOI 10.1074/jbc.RA118.007283.

[43] J.-M. Dong, F. P.-L. Tay, H. L.-F. Swa, J. Gunaratne, T. Leung, B. Burke, E. Manser, *Sci. Signal.* **2016**, *9*, rs4.

[44] G. D. Gupta, É. Coyaud, J. Gonçalves, B. A. Mojarad, Y. Liu, Q. Wu, L. Gheiratmand, D. Comartin, J. M. Tkach, S. W. T. Cheung, M. Bashkurov, M. Hasegan, J. D. Knight, Z.-Y. Lin, M. Schueler, F. Hildebrandt, J. Moffat, A.-C. Gingras, B. Raught, L. Pelletier, *Cell* **2015**, *163*, 1484.

[45] A. Cole, Z. Wang, E. Coyaud, V. Voisin, M. Gronda, Y. Jitkova, R. Mattson, R. Hurren, S. Babovic, N. Maclean, I. Restall, X. Wang, D. V. Jeyaraju, M. A. Sukhai, S. Prabha, S. Bashir, A. Ramakrishnan, E. Leung, Y. H. Qia, N. Zhang, K. R. Combes, T. Ketela, F. Lin, W. A. Houry, A. Aman, R. Al-Awar, W. Zheng, E. Wienholds, C. J. Xu, J. Dick, J. C. Y. Wang, J. Moffat, M. D. Minden, C. J. Eaves, G. D. Bader, Z. Hao, S. M. Kornblau, B. Raught, A. D. Schimmer, *Cancer Cell* **2015**, *27*, 864.

[46] S. Konovalova, X. Liu, P. Manjunath, S. Baral, N. Neupane, T. Hilander, Y. Yang, D. Balboa, M. Terzioglu, L. Euro, M. Varjosalo, H. Tyynismaa, *Redox Biol.* **2018**, *19*, 37.

[47] C. A. Lamb, S. Nühlen, D. Judith, D. Frith, A. P. Snijders, C. Behrends, S. A. Tooze, *EMBO J.* **2016**, *35*, 281.

[48] A. Caballe, D. M. Wenzel, M. Agromayor, S. L. Alam, J. J. Skalicky, M. Kloc, J. G. Carlton, L. Labrador, W. I. Sundquist, J. Martin-Serrano, *eLife* **2015**, *4*, e06547.

[49] K. Jacquet, S. L. Banerjee, F. J. M. Chartier, S. Elowe, N. Bisson, *Mol. Cell. Proteomics MCP* **2018**, *17*, 1979.

[50] C. Yeh, É. Coyaud, M. Bashkurov, P. van der Lelij, S. W. T. Cheung, J. M. Peters, B. Raught, L. Pelletier, *Curr. Biol. CB* **2015**, *25*, 2290.

[51] P. V'kovski, M. Gerber, J. Kelly, S. Pfaender, N. Ebert, S. Braga Lagache, C. Simillion, J. Portmann, H. Stalder, V. Gaschen, R. Bruggmann, M. H. Stoffel, M. Heller, R. Dijkman, V. Thiel, *eLife* **2019**, *8*, DOI 10.7554/eLife.42037.

[52] M. Ramanathan, K. Majzoub, D. S. Rao, P. H. Neela, B. J. Zarnegar, S. Mondal, J. G. Roth, H. Gai, J. R. Kovalski, Z. Siprashvili, T. D. Palmer, J. E. Carette, P. A. Khavari, *Nat. Methods* **2018**, *15*, 207.

[53] R. J. Austin, T. Xia, J. Ren, T. T. Takahashi, R. W. Roberts, *J. Am. Chem. Soc.* **2002**, *124*, 10966.

[54] B. Conlan, T. Stoll, J. J. Gorman, I. Saur, J. P. Rathjen, *Front. Plant Sci.* **2018**, *9*, 1882.

[55] A. Mair, S.-L. Xu, T. C. Branon, A. Y. Ting, D. C. Bergmann, *eLife* **2019**, *8*, DOI 10.7554/eLife.47864.

[56] A. Uezu, D. J. Kanak, T. W. A. Bradshaw, E. J. Soderblom, C. M. Catavero, A. C. Burette, R. J. Weinberg, S. H. Soderling, *Science* **2016**, *353*, 1123.

[57] J. D. Martell, T. J. Deerinck, Y. Sancak, T. L. Poulos, V. K. Mootha, G. E. Sosinsky, M. H. Ellisman, A. Y. Ting, *Nat. Biotechnol.* **2012**, *30*, 1143.

[58] H.-W. Rhee, P. Zou, N. D. Udeshi, J. D. Martell, V. K. Mootha, S. A. Carr, A. Y. Ting, *Science* **2013**, *339*, 1328.

[59] N. D. Udeshi, K. Pedram, T. Svinkina, S. Fereshetian, S. A. Myers, O. Aygun, K.

554 Krug, K. Clauser, D. Ryan, T. Ast, V. K. Mootha, A. Y. Ting, S. A. Carr, *Nat. Methods* **2017**,
555 DOI 10.1038/nmeth.4465.

556 [60] V. Hung, N. D. Udeshi, S. S. Lam, K. H. Loh, K. J. Cox, K. Pedram, S. A. Carr, A. Y.
557 Ting, *Nat. Protoc.* **2016**, *11*, 456.

558 [61] S. S. Lam, J. D. Martell, K. J. Kamer, T. J. Deerinck, M. H. Ellisman, V. K. Mootha,
559 A. Y. Ting, *Nat. Methods* **2014**, *12*, 51.

560 [62] Y. Han, T. C. Branon, J. D. Martell, D. Boassa, D. Shechner, M. H. Ellisman, A. Ting,
561 *ACS Chem. Biol.* **2019**, DOI 10.1021/acschembio.8b00919.

562 [63] X. D. Gao, L.-C. Tu, A. Mir, T. Rodriguez, Y. Ding, J. Leszyk, J. Dekker, S. A.
563 Shaffer, L. J. Zhu, S. A. Wolfe, E. J. Sontheimer, *Nat. Methods* **2018**, DOI 10.1038/s41592-
564 018-0006-2.

565 [64] S. A. Myers, J. Wright, R. Peckner, B. T. Kalish, F. Zhang, S. A. Carr, *Nat. Methods*
566 **2018**, *15*, 437.

567 [65] C.-L. Chen, Y. Hu, N. D. Udeshi, T. Y. Lau, F. Wirtz-Peitz, L. He, A. Y. Ting, S. A.
568 Carr, N. Perrimon, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12093.

569 [66] V. Hung, P. Zou, H.-W. Rhee, N. D. Udeshi, V. Cracan, T. Svinkina, S. A. Carr, V. K.
570 Mootha, A. Y. Ting, *Mol. Cell* **2014**, *55*, 332.

571 [67] S.-Y. Lee, M.-G. Kang, J.-S. Park, G. Lee, A. Y. Ting, H.-W. Rhee, *Cell Rep.* **2016**,
572 *15*, 1837.

573 [68] D. U. Mick, R. B. Rodrigues, R. D. Leib, C. M. Adams, A. S. Chien, S. P. Gygi, M. V.
574 Nachury, *Dev. Cell* **2015**, *35*, 497.

575 [69] J. Hwang, P. J. Espenshade, *Biochem. J.* **2016**, *473*, 2463.

576 [70] S. Han, N. D. Udeshi, T. J. Deerinck, T. Svinkina, M. H. Ellisman, S. A. Carr, A. Y.
577 Ting, *Cell Chem. Biol.* **2017**, *24*, 404.

578 [71] B. T. Lobingier, R. Hüttenhain, K. Eichel, K. B. Miller, A. Y. Ting, M. von Zastrow,
579 N. J. Krogan, *Cell* **2017**, *169*, 350.

580 [72] J. Paek, M. Kalocsay, D. P. Staus, L. Wingler, R. Pascolutti, J. A. Paulo, S. P. Gygi,
581 A. C. Kruse, *Cell* **2017**, *169*, 338.

582 [73] P. Kaewsapsak, D. M. Shechner, W. Mallard, J. L. Rinn, A. Y. Ting, *eLife* **2017**, *6*,
583 DOI 10.7554/eLife.29224.

584 [74] F. M. Fazal, S. Han, K. R. Parker, P. Kaewsapsak, J. Xu, A. N. Boettiger, H. Y.
585 Chang, A. Y. Ting, *Cell* **2019**, DOI 10.1016/j.cell.2019.05.027.

586 [75] A. Padrón, S. Iwasaki, N. T. Ingolia, *Mol. Cell* **2019**, *75*, 875.

587 [76] Y. Zhou, G. Wang, P. Wang, Z. Li, T. Yue, J. Wang, P. Zou, *Angew. Chem. Int. Ed*
588 *Engl.* **2019**, DOI 10.1002/anie.201905949.

589 [77] M. S. Dickinson, L. N. Anderson, B.-J. M. Webb-Robertson, J. R. Hansen, R. D.
590 Smith, A. T. Wright, K. Hybiske, *PLoS Pathog.* **2019**, *15*, e1007698.

[78] E. A. Rucks, M. G. Olson, L. M. Jorgenson, R. R. Srinivasan, S. P. Ouellette, *Front. Cell. Infect. Microbiol.* **2017**, *7*, 40.

[79] M. G. Olson, L. M. Jorgenson, R. E. Widner, E. A. Rucks, *Methods Mol. Biol. Clifton NJ* **2019**, *2042*, 245.

[80] M. G. Olson, R. E. Widner, L. M. Jorgenson, A. Lawrence, D. Lagundzin, N. T. Woods, S. P. Ouellette, E. A. Rucks, *Infect. Immun.* **2019**, IAI.00537.

[81] U. S. Ganapathy, L. Bai, L. Wei, K. A. Eckartt, C. M. Lett, M. L. Previti, I. S. Carrico, J. C. Seeliger, *ACS Infect. Dis.* **2018**, *4*, 918.

[82] Á. Ástvaldsson, K. Hultenby, S. G. Svärd, J. Jerlström-Hultqvist, *mSphere* **2019**, *4*, DOI 10.1128/mSphereDirect.00153-19.

[83] Y. G. Santin, T. Doan, R. Lebrun, L. Espinosa, L. Journet, E. Cascales, *Nat. Microbiol.* **2018**, *3*, 1304.

[84] Y. Akhter, S. Thakur, *Microbiol. Res.* **2017**, *204*, 9.

[85] M. J. Pearce, J. Mintseris, J. Ferreyra, S. P. Gygi, K. H. Darwin, *Science* **2008**, *322*, 1104.

[86] F. Striebel, F. Imkamp, D. Özcelik, E. Weber-Ban, *Biochim. Biophys. Acta* **2014**, *1843*, 103.

[87] J. Barandun, C. L. Delley, E. Weber-Ban, *BMC Biol.* **2012**, *10*, 95.

[88] M. I. Samanovic, K. H. Darwin, *Trends Microbiol.* **2016**, *24*, 26.

[89] E. Guth, M. Thommen, E. Weber-Ban, *J. Biol. Chem.* **2011**, *286*, 4412.

[90] F. Striebel, F. Imkamp, M. Sutter, M. Steiner, A. Mamedov, E. Weber-Ban, *Nat. Struct. Mol. Biol.* **2009**, *16*, 647.

[91] T. Wang, K. H. Darwin, H. Li, *Nat. Struct. Mol. Biol.* **2010**, *17*, 1352.

[92] Q. Liu, J. Zheng, W. Sun, Y. Huo, L. Zhang, P. Hao, H. Wang, M. Zhuang, *Nat. Methods* **2018**, *15*, 715.

[93] S. Urlinger, U. Baron, M. Thellmann, M. T. Hasan, H. Bujard, W. Hillen, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 7963.

[94] L. A. Banaszynski, C. W. Liu, T. J. Wandless, *J. Am. Chem. Soc.* **2005**, *127*, 4715.

[95] D. I. Kim, J. A. Cutler, C. H. Na, S. Reckel, S. Renuse, A. K. Madugundu, R. Tahir, H. L. Goldschmidt, K. L. Reddy, R. L. Huganir, X. Wu, N. E. Zachara, O. Hantschel, A. Pandey, *J. Proteome Res.* **2017**, DOI 10.1021/acs.jproteome.7b00775.

[96] J. A. Groves, A. O. Maduka, R. N. O'Meally, R. N. Cole, N. E. Zachara, *J. Biol. Chem.* **2017**, *292*, 6493.

[97] N. Opitz, K. Schmitt, V. Hofer-Pretz, B. Neumann, H. Krebber, G. H. Braus, O. Valerius, *Mol. Cell. Proteomics MCP* **2017**, *16*, 2199.

[98] J. A. Ankney, A. Muneer, X. Chen, *Annu. Rev. Anal. Chem. Palo Alto Calif* **2018**, *11*,

49.

[99] A. Azhar, G. W. Booker, S. W. Polyak, *Biochem. Anal. Biochem.* **2015**, *4*, 210.

[100] M. Larochelle, D. Bergeron, B. Arcand, F. Bachand, *J. Cell Sci.* **2019**, *132*, DOI 10.1242/jcs.232249.

[101] Y. Zhang, G. Song, N. K. Lal, U. Nagalakshmi, Y. Li, W. Zheng, P.-J. Huang, T. C. Branon, A. Y. Ting, J. W. Walley, S. P. Dinesh-Kumar, *Nat. Commun.* **2019**, *10*, 3252.

[102] K. H. Loh, P. S. Stawski, A. S. Draycott, N. D. Udeshi, E. K. Lehrman, D. K. Wilton, T. Svinkina, T. J. Deerinck, M. H. Ellisman, B. Stevens, S. A. Carr, A. Y. Ting, *Cell* **2016**, *166*, 1295.

[103] N. Kotani, J. Gu, T. Isaji, K. Udaka, N. Taniguchi, K. Honke, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 7405.

[104] K. Honke, A. Miyagawa-Yamaguchi, N. Kotani, *Methods Mol. Biol. Clifton NJ* **2019**, *2008*, 1.

[105] J. S. Rees, *Methods Mol. Biol. Clifton NJ* **2019**, *2008*, 13.

[106] J. S. Rees, X.-W. Li, S. Perrett, K. S. Lilley, A. P. Jackson, *Curr. Protoc. Protein Sci.* **2015**, *80*, 19.27.1.

[107] X.-W. Li, J. S. Rees, P. Xue, H. Zhang, S. W. Hamaia, B. Sanderson, P. E. Funk, R. W. Farndale, K. S. Lilley, S. Perrett, A. P. Jackson, *J. Biol. Chem.* **2014**, *289*, 14434.

[108] D. B. Beck, R. Bonasio, *Curr. Protoc. Chem. Biol.* **2017**, *9*, 128.

[109] Z. B. Hill, S. B. Pollock, M. Zhuang, J. A. Wells, *J. Am. Chem. Soc.* **2016**, *138*, 13123.

[110] M. Zhuang, S. Guan, H. Wang, A. L. Burlingame, J. A. Wells, *Mol. Cell* **2013**, *49*, 273.

[111] B. van Steensel, S. Henikoff, *Nat. Biotechnol.* **2000**, *18*, 424.

[112] F. Cléard, F. Karch, R. K. Maeda, *Methods Mol. Biol. Clifton NJ* **2014**, *1196*, 279.

[113] V. Visone, W. Han, G. Perugino, G. Del Monaco, Q. She, M. Rossi, A. Valenti, M. Ciaramella, *PloS One* **2017**, *12*, e0185791.

[114] P. V. Escribá, X. Busquets, J. Inokuchi, G. Balogh, Z. Török, I. Horváth, J. L. Harwood, L. Vígh, *Prog. Lipid Res.* **2015**, *59*, 38.

Legend to Figures

Figure 1. Schematic representation of conventional methods used for detection of PPIs. **a)** Co-immunoprecipitation workflow. From left to right: cellular proteins are extracted from cells and incubated with specific antibody-coated beads directed against a protein of interest (green). The

660 protein of interest and its direct (blue) and indirect (yellow) partners are then co-
661 immunoprecipitated. Detection of PPIs is achieved by Western Blot analysis (Total, total
662 proteins in cellular extract; Co-IP, partners co-immunoprecipitated with the protein of interest)
663 **b)** Two-hybrid assay. Two-hybrid assay is based on the reconstitution of a specific protein
664 (adenylate cyclase for BACTH and Gal4 for Y2H) that consists in two domains (light and dark
665 grey). Bait (in green) and prey (in blue or red) proteins are translationally fused to isolated
666 domains. Spatial association between bait and prey proteins restores the proximity of the two
667 domains and then the protein activity, leading *in fine* to the transcription of a reporter gene. **c)**
668 FRET assay. A bait (in green) and a prey (in blue or red) protein are translationally fused in
669 tandem to fluorescent proteins. An example is showed with the cyan fluorescent protein (CFP,
670 cyan) and the yellow fluorescent protein (YFP, yellow). Energy transfer (i.e., FRET) between
671 fluorescent reporters can occur if the bait and the prey proteins are within 10 nm of distance
672 (left panel). The loss of fluorescence intensity from the donor (CFP) and the increase of
673 fluorescence intensity from the acceptor (YFP) can be visualized by monitoring fluorescence
674 spectra (right panel).
675
676 **Figure 2.** BioID workflow. Fusion between BioID protein (blue) and a bait protein (black) is
677 expressed in native context within cells. *In vivo* labeling (blue radius) is initiated by the addition
678 of a supra-physiological concentration of biotin during many hours. Cells are then lysed, and
679 biotinylated proteins are purified by enrichment on streptavidin beads. Trypsin digestion
680 provides the generation of peptides that are then identified by mass spectrometry.
681
682 **Figure 3.** APEX proximity-dependent biotin labeling workflow. Fusion between APEX2
683 protein (blue) and a bait protein (black) is expressed in native context within cells. After
684 incubation in presence of biotin-phenol, a H_2O_2 pulse allows biotin labeling (red radius) of both

direct partners (yellow and green) and proteins in close environment of the bait protein (brown). While unbiotinylated proteins are washed, biotinylated proteins are purified by enrichment on streptavidin beads after cell lysis. Trypsin digestion provides the generation of peptides that are then identified by mass spectrometry.

Figure 4. PUP-IT workflow for intra and extracellular labeling. **a)** For intracellular labeling, the cytosolic domain of the membrane receptor CD28 (blue) is fused to the PUP ligase PafA (green). The expression of the bio-Pup(E) substrate is under the control of TET-ON system. Addition of doxycycline (Dox) induces bio-Pup(E) production and initiates the PafA-mediated labeling (green radius). **b)** For cell surface labeling, the extracellular domain of the membrane receptor CD28 (blue) and PafA (green) are fused to specific adaptors, FKBP and FRB, respectively. Addition of rapamycin (purple pill) allows the link between the two adaptors and hence, the formation of a functional PUP-IT complex. PafA-mediated labeling (green radius) is initiated upon the addition of ATP and truncated bio-Pup(E) peptides (bio-DE28).

Figure 5. Schematic pipelines for major quantitative approaches used for proximity-tagging quantification. For each approach, 1 corresponds to the negative control (without tagging enzyme), 2 corresponds to the specificity control (where the tagging enzyme is independently expressed in the same cellular compartment that the bait protein), and 3 corresponds to the test (the bait protein is fused to the tagging enzyme). **a)** Three-state SILAC experiment. Cells are grown in presence of light (L, yellow, negative control), medium (M, blue, specificity control) or heavy (H, red, test) stable isotope–enriched amino acids. Upon proximity labeling, L, M and H protein extracts are combined at equal ratio before further processing. Biotinylated proteins are then enriched, digested and quantified by LC-MS/MS analysis. In this case, H/L intensity ratio represents the extent of biotinylation by the tagging enzyme and H/M intensity ratio

represents the extent of specific biotinylation by the tagging enzyme versus non-specific endogenous proteins. **b)** Triplex TMT experiment. Cells are grown in standard conditions. Upon proximity labeling, biotinylated proteins are enriched, digested and then chemically tagged with MS-differentiable TMT labels (126, red, negative control; 127, blue, specificity control; 128, green, test). Tagged-peptides are combined at equal ratio and quantified by LC-MS/MS analysis. 128/126 intensity ratio is used to distinguish the extent of biotinylation by the tagging enzyme and 128/127 intensity ratio represents the extent of specific biotinylation by the tagging enzyme versus non-specific endogenous proteins. **c)** LFQ assay. Cells are grown in standard conditions. Upon proximity labeling, biotinylated proteins are enriched and digested. Samples are then independently quantified and compared by using ion intensity or spectral counting (as represented).

Figure 1.

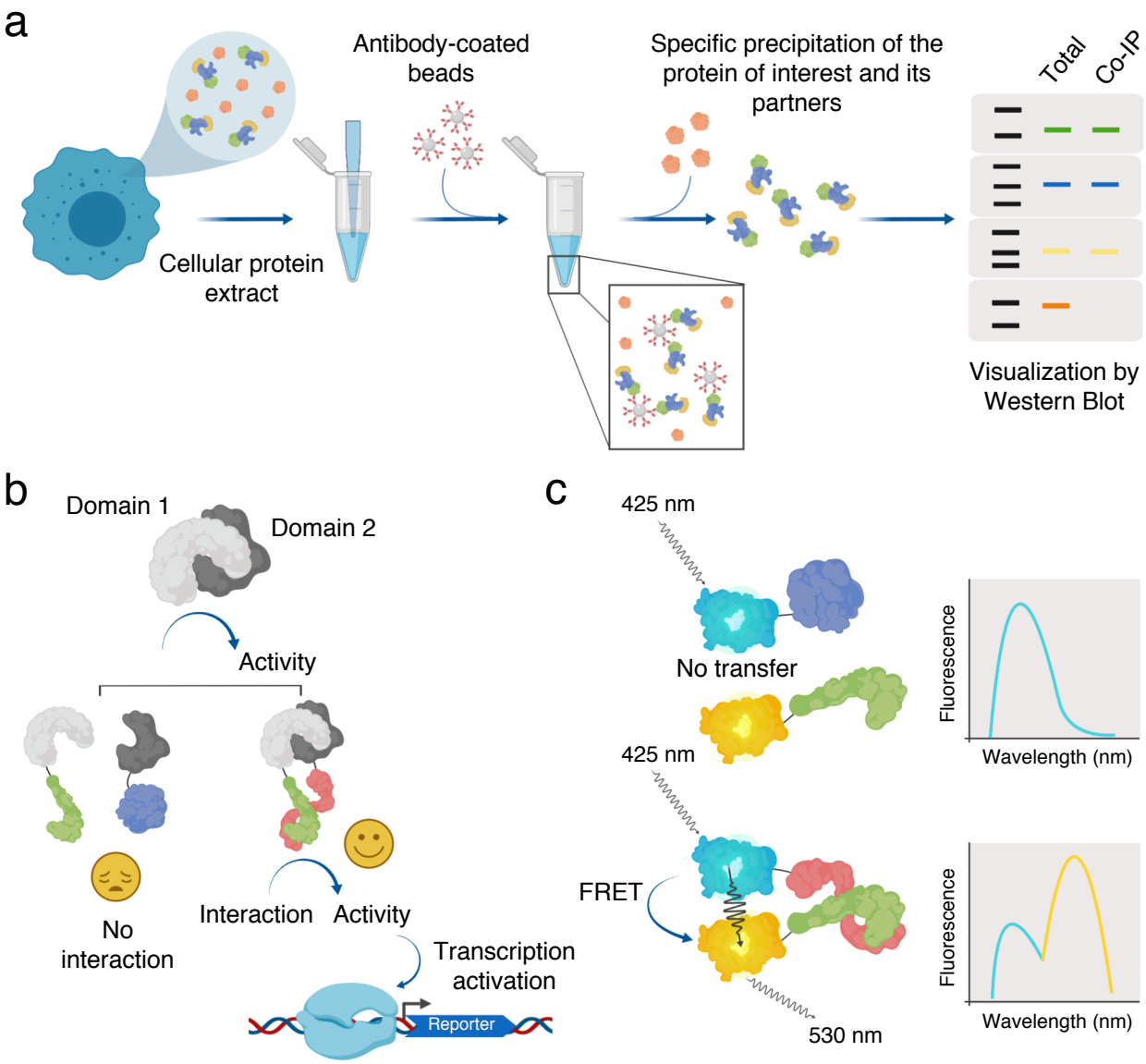

a
Antibody-coated beads
Specific precipitation of the protein of interest and its partners
Total
Co-IP
Cellular protein extract
Visualization by Western Blot
b
Domain 1
Domain 2
Activity
No interaction
Interaction
Activity
Transcription activation
Reporter
c
425 nm
No transfer
Fluorescence
Wavelength (nm)
425 nm
FRET
530 nm
Fluorescence
Wavelength (nm)

Figure 2.

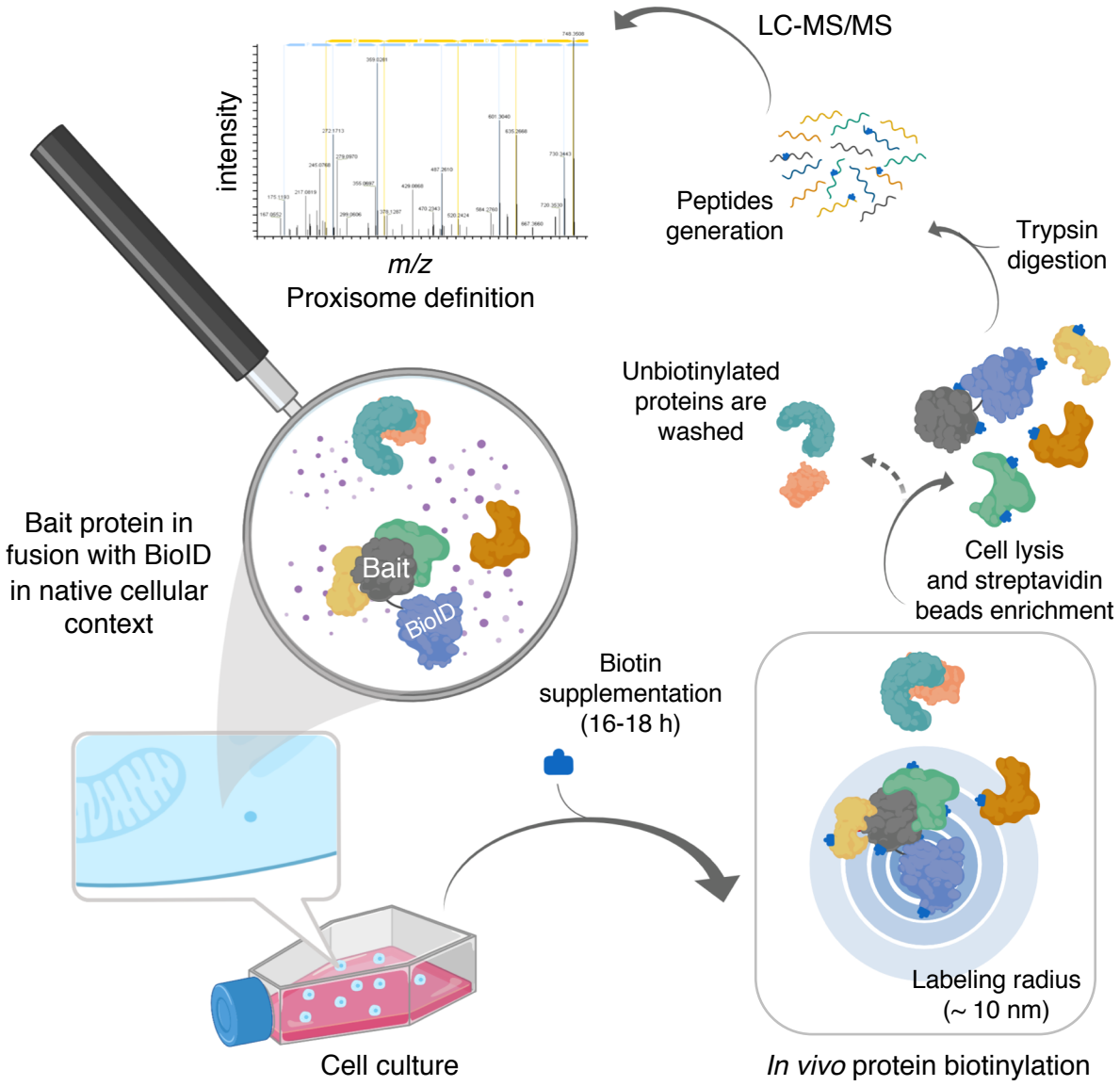

LC-MS/MS
intensity
m/z
Proxisome definition
Peptides generation
Trypsin digestion
Unbiotinylated proteins are washed
Cell lysis and streptavidin beads enrichment
Bait protein in fusion with BioID in native cellular context
Bait
BioID
Biotin supplementation (16-18 h)
Labeling radius (~ 10 nm)
Cell culture
In vivo protein biotinylation

Figure 3.

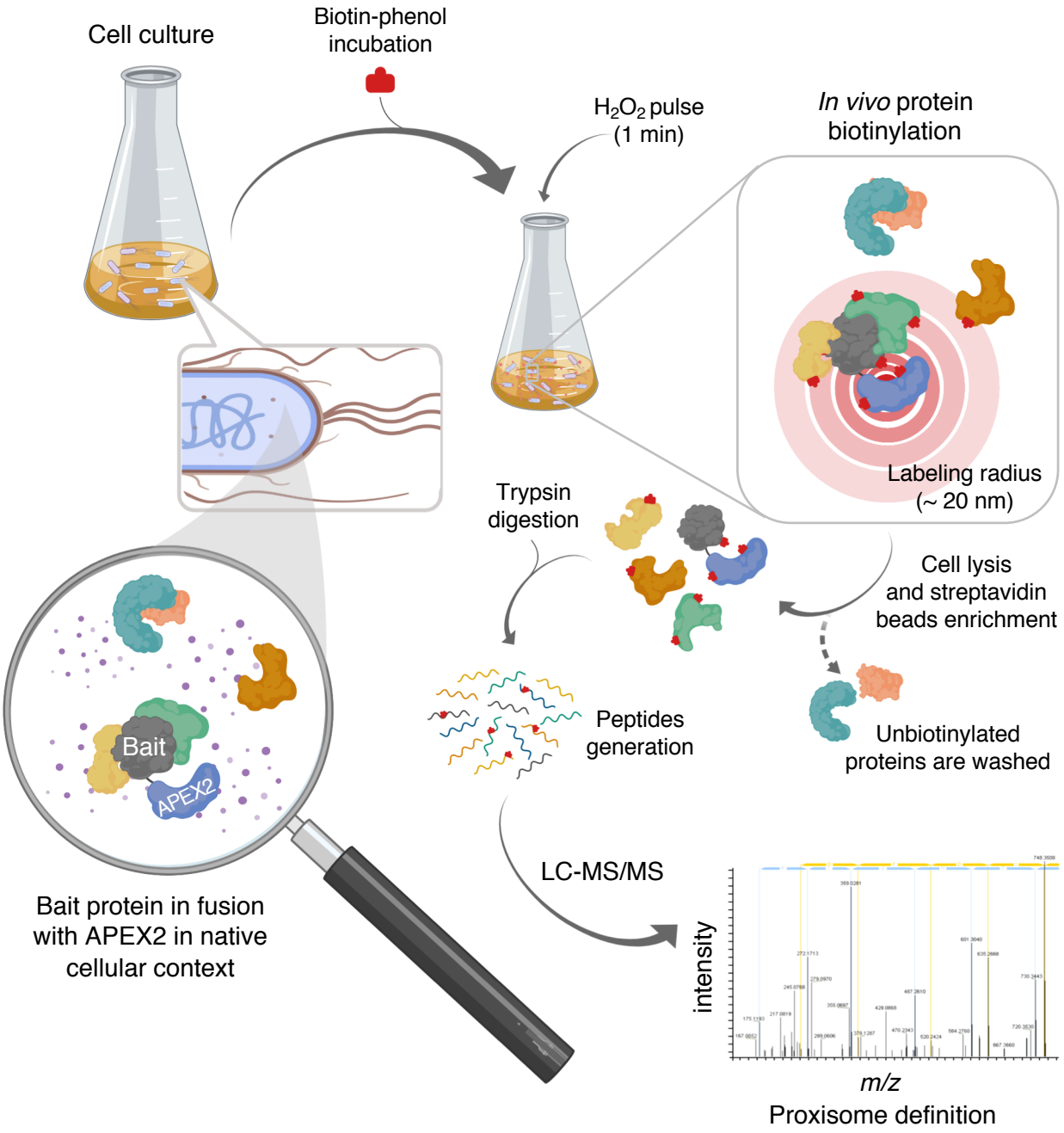

Cell culture
Biotin-phenol incubation
H_2O_2 pulse (1 min)
In vivo protein biotinylation
Labeling radius (~ 20 nm)
Trypsin digestion
Cell lysis and streptavidin beads enrichment
Peptides generation
Unbiotinylated proteins are washed
LC-MS/MS
Bait
APEX2
Bait protein in fusion with APEX2 in native cellular context
intensity
m/z
Proxisome definition

Figure 4.

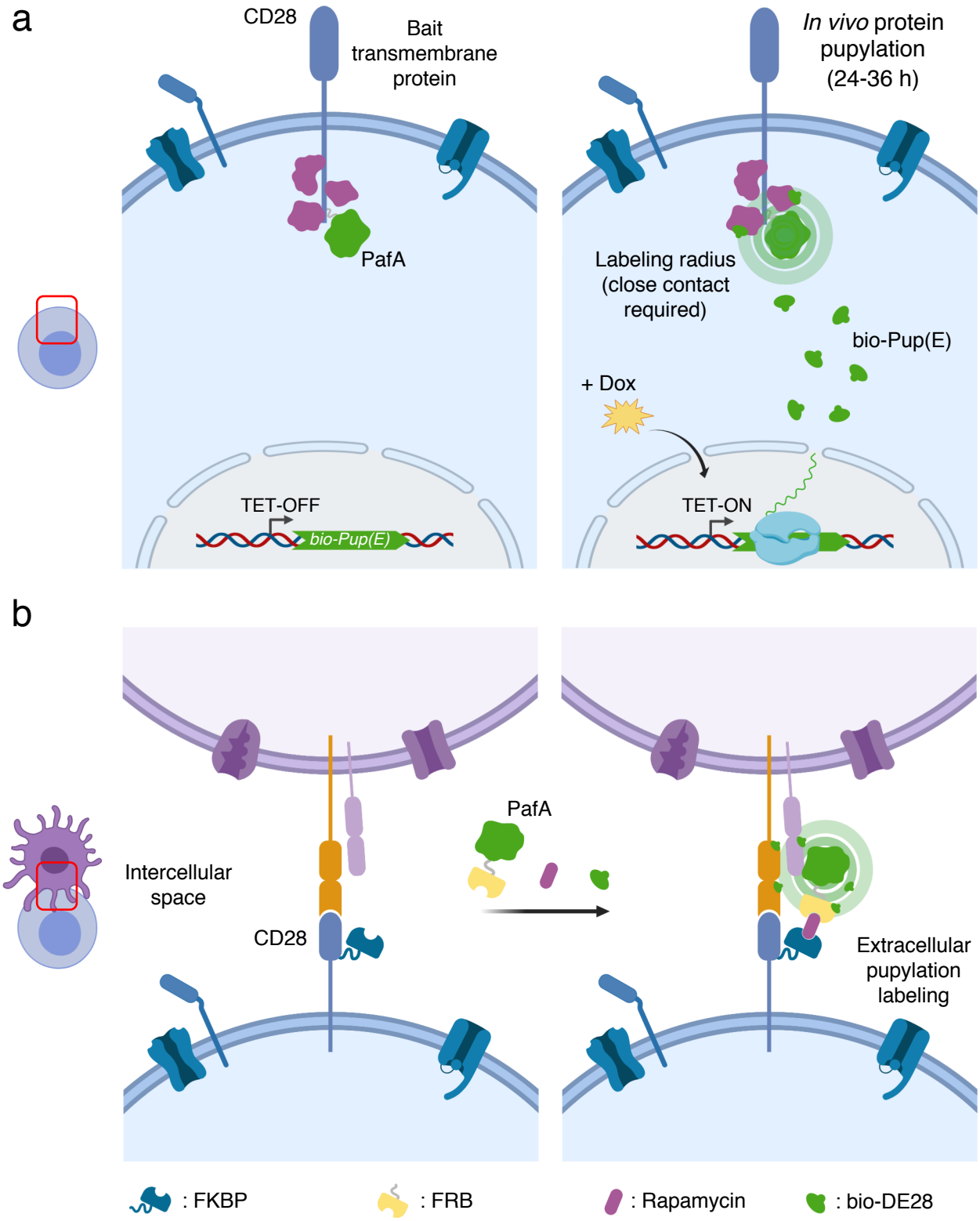

a
CD28
Bait transmembrane protein
PafA
TET-OFF
bio-Pup(E)
In vivo protein pupylation (24-36 h)
Labeling radius (close contact required)
bio-Pup(E)
+ Dox
TET-ON
b
Intercellular space
CD28
PafA
Extracellular pupylation labeling
: FKBP
: FRB
: Rapamycin
: bio-DE28

Figure 5.

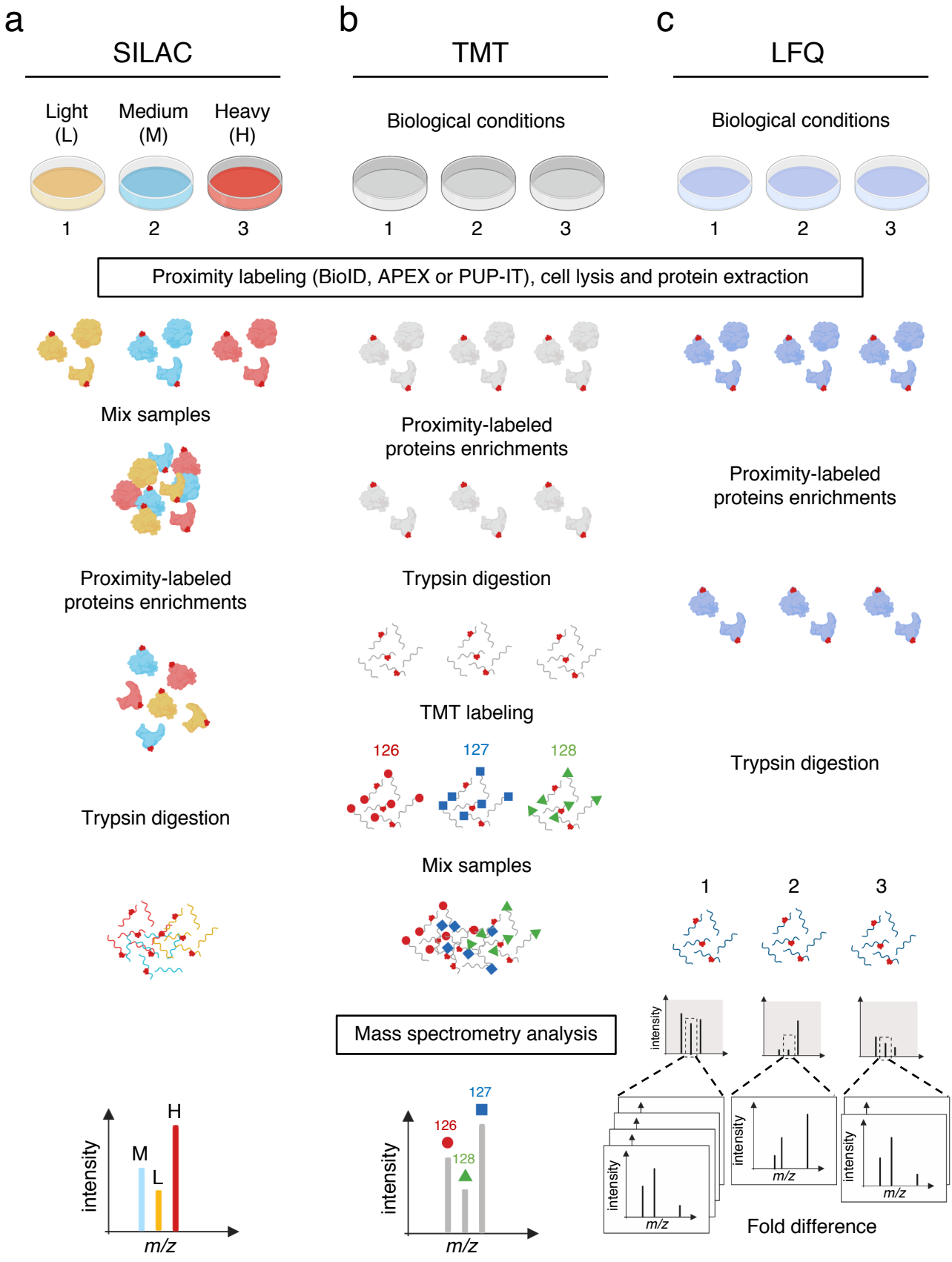

a
SILAC
Light (L)
Medium (M)
Heavy (H)
1
2
3
b
TMT
Biological conditions
1
2
3
c
LFQ
Biological conditions
1
2
3
Proximity labeling (BioID, APEX or PUP-IT), cell lysis and protein extraction
Mix samples
Proximity-labeled proteins enrichments
Trypsin digestion
Mix samples
Proximity-labeled proteins enrichments
Trypsin digestion
TMT labeling
126
127
128
Mix samples
Proximity-labeled proteins enrichments
Trypsin digestion
1
2
3
Mass spectrometry analysis
intensity
H
M
L
m/z
intensity
127
126
128
m/z
intensity
m/z
Fold difference

Table 1. Comparative table of the main features of the different proximity-tagging methods

	BioID	(mini)TurboID	APEX	PUP-IT
Origin	Bacteria (BirA from *Escherichia coli*)	Bacteria (BirA from *Escherichia coli*)	Plantae (APX from *Glycine max*)	Bacteria (PafA from *Corynebacterium glutamicum*)
Molecular mass	35 kDa (for BioID) 27 kDa (for BioID2)	28 kDa (for miniTurbo) 35 kDa (for TurboID)	27 kDa	54 kDa
Substrates	Biotin and ATP	Biotin and ATP	H_2O_2 and phenol derivative-molecules	Pup(E) derivative-peptides and ATP
Range of temperature[a]	37°C for BioID and 37°C to 50°C for BioID2	20°C to 37°C	30°C to 37°C	37°C
Enzyme activity[b] (at 37°C)	+	+++	++++	++
Time of labeling	16-18 hours	10 min to several hours (depending on the organism)	1 min	24-36 hours
Labeling radius	10-15 nm	10-15 nm	20 nm	Very close contact
Labeled macromolecules	Proteins	Proteins	Proteins and nucleic acids	Proteins
Labeled residues	Lys	Lys	Tyr (98%) and Trp, His, Cys (2%) for proteins; mostly guanosine for nucleic acids	Lys
Temporal control (labeling initiation)	No	No	Yes	No
Information about the POI	Long history	Intermediate	Snapshot	Long history

[a] Enzyme works at these temperatures but temperatures between this range have not necessary been tested

[b] ++++, strong activity; +++, very good activity; ++, good activity; +, weak activity