Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Protein engineering for selective proteomics

Katarzyna Radziwon and Amy M. Weeks

Summary

Post-translational modifications, complex formation, subcellular localization, and cell-type-specific expression create functionally distinct protein subpopulations that enable living systems to execute rapid and precise responses to changing conditions. Systems-level analysis of these subproteomes remains challenging, requiring preservation of spatial information or enrichment of species that are transient and present at low abundance. Engineered proteins have emerged as important tools for selective proteomics based on their capacity for highly specific molecular recognition and their genetic targetability. Here, we focus on new developments in protein engineering for selective proteomics of post-translational modifications, protein complexes, subcellular compartments, and cell types. We also address remaining challenges and future opportunities to integrate engineered protein tools across different subproteome scales to map the proteome with unprecedented depth and detail.

Addresses

Department of Biochemistry, University of Wisconsin–Madison, Madison, WI, 53706, USA

Corresponding author: Weeks, Amy M (amweeks@wisc.edu)

Current Opinion in Chemical Biology 2021, **60**:10–19

This reviews comes from a themed issue on **Omics**

Edited by **Nichollas Scott** and **Laura Edgington-Mitchell**

For a complete overview see the Issue and the Editorial

https://doi.org/10.1016/j.cbpa.2020.07.003

Keywords

Protein engineering, Proteomics, Proximity tagging, Post-translational modifications, Enzymatic capture, Bioorthogonal chemistry.

Introduction

Functionally distinct subpopulations of proteins arise as a result of post-translational modifications (PTMs), formation of complexes, subcellular localization, and cell-type-specific expression. The presence of a protein within a specific subpopulation may determine whether it is active or inhibited, whether it is colocalized with substrates or partners that it acts on, and what the fate of its product or output will be. Defining the modifications of individual proteins and the molecular composition of protein complexes, subcellular compartments, and specific cell types is therefore a central challenge in assigning protein function. Beyond advancing our fundamental understanding of biological regulatory mechanisms, this information has led to the development of new drugs and therapeutic hypotheses for the treatment of human disease. These include antibody–drug conjugates [1] and chimeric antigen receptor T cells that target cells expressing specific disease-associated proteins on their surfaces [2]; chimeric small molecules that redirect the activity of ubiquitin ligases toward unnatural substrates by induced protein complex formation for targeted protein degradation [3]; small molecule correctors that restore trafficking of receptors to the cell surface [4]; and drugs that target PTM enzymes, including kinases [5] and proteases [6].

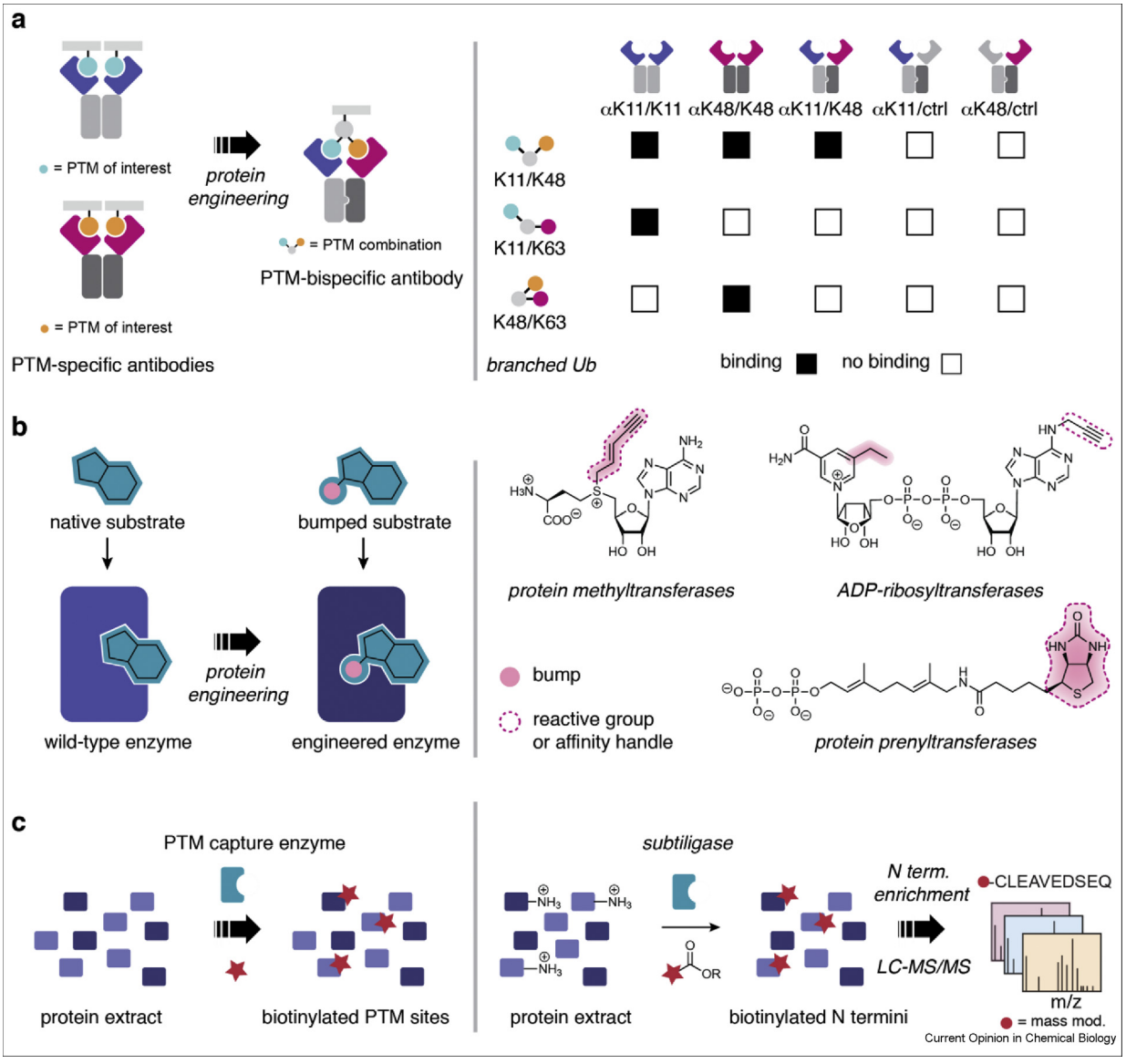
Over the past two decades, advances in mass spectrometry (MS) have enabled rapid and deep profiling of the proteome. However, selective analysis of specific subproteomes remains significantly more challenging than measurement of bulk protein abundance. PTMs are often present at low stoichiometry [7] and therefore must be enriched to remove high background from unmodified proteins, while protein complexes, subcellular compartments, and specific cell types must be isolated before analysis to preserve spatial information [8,9]. These protocols are made especially challenging by the dynamic nature of PTMs, the potential for protein complexes to be transient, and the inability to isolate certain cellular compartments, including major organelles such as the endoplasmic reticulum [8], membraneless organelles formed via phase separation, and biochemical compartments [10], by standard differential centrifugation-based subcellular fractionation approaches. Engineered proteins recently have emerged as important tools to address these challenges in selective proteomics. Proteins have at least three key advantages as tools for selective proteomics, including their capacity to perform highly selective molecular recognition of specific functional groups; their ability to be genetically targeted to specific cells, compartments, or protein complexes; and their evolution to operate on fast timescales in the cellular environment for capture of transient protein subpopulations. This review focuses on recent advances in protein engineering that enable capture of specific subproteomes for analysis by MS-based proteomics. We discuss engineered proteins that enable enrichment of

proteins bearing specific chemical modifications, proteins that reside in specific complexes or compartments, and proteins that are expressed in specific cell types. We also present a perspective on future applications and challenges in applying engineered proteins across different subproteome types to gain detailed insights into protein function.

Capturing PTMs with engineered proteins

PTMs to proteins provide a mechanism for proteomic diversification through introduction of chemical functional groups not found in the genetically encoded amino acids and by alteration of peptide bond connectivity. PTMs identified to date include phosphorylation, glycosylation, acetylation, and proteolysis, among

Figure 1



Engineered proteins for selective enrichment of post-translationally modified proteins and peptides. (a) Phage-display-derived antibodies for PTM binding and enrichment. Left: Antibody phage display has been used for the selection of PTM-specific antibodies. These PTM-specific antibodies can be incorporated into two arms of an IgG to generate PTM-bispecific antibodies that require two different PTMs for binding. Right: Bispecific antibodies that recognize branched ubiquitin chains that contain both K11 and K48 linkages. **(b)** Left: A schematic of the bump-hole strategy for generating PTM enzymes to install a bioorthogonally reactive or biotin handle. Right: Examples of bump-hole substrate–enzyme pairs for selective proteomics. **(c)** Left: PTM capture enzymes recognize and covalently modify specific PTM sites. Right: Subtiligase is an engineered enzyme that selectively biotinylates protein N termini, including those generated by proteolysis, enabling their enrichment and sequencing using LC-MS/MS.

hundreds of others [11]. These modifications play important roles in regulating protein structure, enzymatic activity, lifetime, and localization. A critical step toward assigning the biological functions of PTMs is to understand their residue localization and spatiotemporal dynamics. Although MS has emerged as a routine tool for proteomic analysis, profiling PTMs remains challenging because many modifications are present only at low stoichiometry [7] and are dynamically regulated by multiple enzymes [12]. In recent years, engineered proteins have emerged as robust tools for PTM-selective proteomics, enabling enrichment of modified proteins to reduce sample complexity and increase sampling depth. Engineered proteins function in PTM enrichment based on two primary mechanisms: molecular recognition of specific modifications and/or covalent modification of specific functional groups based on their unique chemical properties.

Engineered proteins for affinity capture of PTMs

Protein scaffolds can perform high affinity, high specificity molecular recognition of a wide variety of structures and functional groups and therefore represent the most universal strategy for PTM enrichment (Figure 1a). Although alternative strategies have proven very useful for certain PTMs, such as immobilized metal affinity chromatography for phosphopeptide enrichment [13], such techniques typically cannot enrich for specific modified residues and linkage types. For example, although immobilized metal affinity chromatography enables enrichment of pSer, pThr, and pTyr peptides, pTyr comprises <1% of all cellular phosphorylation sites and is therefore challenging to sample with appropriate depth [14]. Antibody-based molecular recognition of pTyr enables isolation of this small but biologically important subset of phosphopeptides for MS analysis [15,16]. Although the mostly commonly used antibodies for PTM enrichment were developed by animal immunization, modern protein engineering methods provide high-throughput platforms for identifying and improving PTM-specific antibodies and other PTM-binding protein scaffolds. Display approaches have recently been used for the development of nature-inspired phosphospecific antibodies [17,18] and Src homology 2 domain-based pTyr binders [19] for proteomics. Antibody phage display has also enabled the development of antibodies for linkage-specific enrichment of ubiquitin (Ub)-modified proteins. The versatility of Ub modification in directing diverse signaling outcomes in eukaryotic cells depends on the ability of the cellular ubiquitylation machinery to construct Ub chains that differ by their linkages and topologies (linear vs branched) to specify different signaling outcomes. Previous work used antibody phage display to develop linkage-specific antibodies recognizing Ub chains at K11 [20] or K48 [21]. More recently, a knobs-into-holes heterodimerization strategy was used to combine these linkage-specific antibodies to generate a bispecific antibody that recognizes branched Ub chains containing both K11 and K48 linkages [22] (Figure 1a). This bispecific antibody was used to identify cellular conditions that lead to the accumulation of K11/K48-linked Ub chains, such as proteotoxic stress. The antibody was further used to enrich proteins modified by this specific branched-chain Ub topology, enabling their identification by quantitative MS and leading to the identification of UBR4 and UBR5 as the E3 Ub ligases primarily responsible for modifying proteins with this Ub topology. The multispecific antibody strategy has potential as a generalizable approach for enriching proteoforms with specific combinations of PTMs.

Engineered enzymes for covalent PTM capture

Engineered enzymes represent attractive tools for selective proteomics based on their potential to combine protein-based molecular recognition with the formation of a covalent bond to a chemically reactive group or affinity handle such as biotin. This enables enrichment of proteins bearing specific PTMs and their subsequent identification using liquid chromatography (LC)-MS/MS. Nature has evolved a diverse repertoire of enzymes capable of installing, modifying, and recognizing specific PTMs. Protein engineering approaches have enabled many of these enzymes to be repurposed as tools for selective proteomics, facilitating systems-level study of biological signaling pathways. Enzymes have the advantages that they can be genetically targeted to specific subcellular compartments, providing an opportunity for subcellular spatial resolution of PTMs, and that they have typically evolved to function on fast timescales in the cellular environment, providing an opportunity for temporal resolution of PTMs.

Early methods for covalent labeling of PTM sites took advantage of the ability of cellular enzymes to accept synthetic, bioorthogonally reactive precursors to PTMs, termed metabolic chemical reporters (MCRs), that are incorporated into PTMs after feeding to cells or animals [23,24]. After incorporation, the bioorthogonal tag can be conjugated with biotin, enabling enrichment of proteins bearing the PTM of interest. However, a limitation of the MCR approach is that some PTMs, such as various glycan structures, use the same building blocks. The development of enzymatic tools that recognize and modify specific PTMs has helped to overcome this challenge. PTM-modifying enzymes have been developed for the capture of O-GlcNAc [25], fucose-α(1,2)-galactose [26], N-acetyllactosamine, N-acetylneuraminic acid-α(2−3)-galactose [27], and cell surface N-linked glycans [28]. More recently, a new enzymatic tool for glycoproteomic mapping of mucin-domain glycoproteins, which are densely O-glycosylated, has been

developed [29]. Rather than labeling O-glycosylated mucins with a reactive reporter, this strategy relies on recognition of the proteins of interest by StcE, an *Escherichia coli* protease that specifically cleaves mucin domains by recognizing a peptide- and glycan-based motif. StcE has been deployed as a tool for proteomics to improve sequence coverage, glycosite mapping, and glycoform analysis of purified mucin samples, and has also been used as an enrichment reagent for O-glycosylated mucins by conjugating it to beads and enriching in the presence of EDTA, which inhibits StcE activity. Future characterization of additional bacterial mucinases and engineering of StcE holds promise for the development of new tools for selective proteomics of proteins modified with specific glycan structures.

An alternative strategy for chemoenzymatic PTM capture relies on the use of substrate analogs that have been modified with PTM reporters that can only be accepted as substrates by PTM enzymes that have been engineered to recognize them (Figure 1b). This 'bump-hole' approach has the advantage that it enables the identification of individual protein substrates of closely related PTM enzymes because only the enzyme engineered to have a hole to accommodate the bumped substrate can use it to install PTMs. The bump-hole strategy has been successfully implemented to enable selective proteomics of substrates of kinases [30], methyltransferases [31], prenyltransferases [32], acetyltransferases [33], and ADP-ribosyltransferases (ARTs) [34–36]. Recent work has combined a proximity labeling approach with a bump-hole approach to identify substrates and interactors of PARP14, a human ART that plays an important role in immune function [37]. A mutant PARP14 was designed to accommodate the NAD^+ analog 5-benzyl-6-alkyne-NAD^+, enabling biotinylation of the resulting ADP-ribose modification using click chemistry for enrichment and identification using LC-MS/MS. In parallel, PARP14 interactors were selectively biotinylated by BioID proximity labeling using a PARP14-BioID fusion protein and were identified using LC-MS/MS. Combining these two data sets resulted in high confidence identification of 114 substrates of PARP14. This combined proximity labeling/PTM capture pipeline represents an approach that could be generalized to many other PTMs for accurate identification of PTM-substrate pairs. A remaining limitation of the bump-hole approach is that bumped substrate analogs are often cell impermeable, limiting the method to application in cell lysates. In the future, the MCR approach could be combined with the bump-hole protein engineering strategy to map PTM enzyme–substrate relationships in living cells.

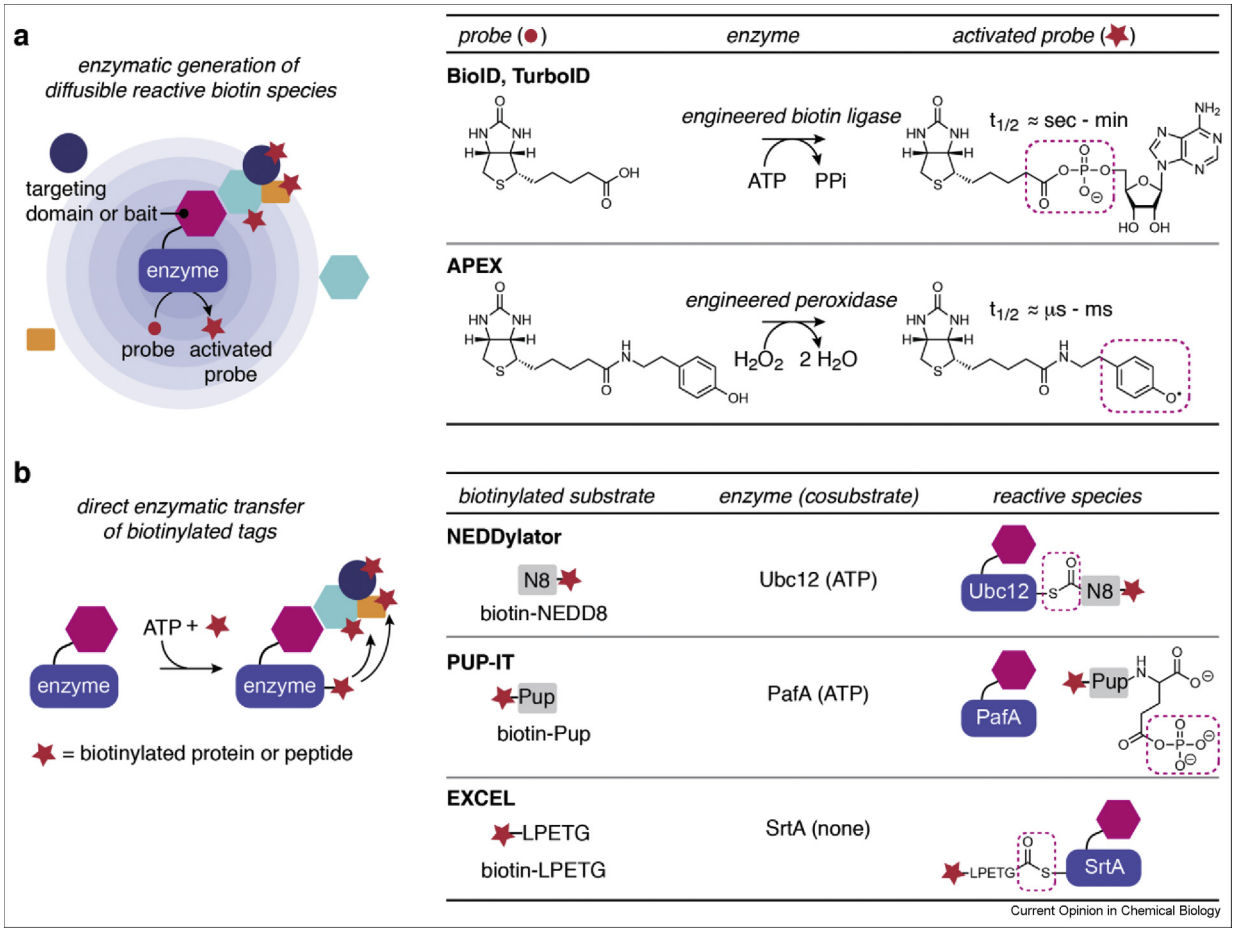
Proteolysis is a unique PTM in that it does not add a new functional group to proteins, but rather alters peptide bond connectivity by generating neo-N and neo-C termini. Although protein N and C termini have unique structures compared to other biological amines and carboxylates, they are nonetheless challenging to modify selectively because their reactivity is similar to that of abundant amino acid side chains such as Lys (N termini) and Asp/Glu (C termini). Proteases that have been engineered to favor peptide ligation have emerged as useful tools for terminal modification of proteins to enable selective proteomics of N and C termini, or 'terminomics' [38,39] (Figure 1c). The designed peptide ligase subtiligase was the first enzymatic tool to be used in N terminomics, enabling the study of proteolytic pathways involved in apoptosis, inflammation, viral and bacterial pathogenesis, and protein trafficking [40]. Subtiligase catalyzes a ligation reaction between C-terminal peptide esters and the N-terminal α-amines of peptides or proteins [40]. In N terminomics studies, subtiligase is applied to globally modify N termini in cell lysates with a biotinylated peptide ester substrate to enable enrichment of N-terminal peptides and their identification and quantification using LC-MS/MS. Recent work using the proteomic identification of ligation sites approach, which uses highly diverse, proteome-derived peptide libraries as pools of substrates for subtiligase, identified a family of subtiligase mutants that modify N-terminal sequences that cannot be efficiently modified by the wild-type enzyme [41]. This enabled application of subtiligase to study proteolytic cleavage catalyzed by signal peptidase, a broad specificity protease that cleaves many sequences not efficiently captured by wild-type subtiligase. Further protein engineering work led to the development of subtiligase-TM, a subtiligase variant that is targeted to the plasma membrane for efficient and specific capture of cell surface N termini in live cells [42]. Subtiligase-TM was used to sequence hundreds of cell surface N termini and to quantify changes in their abundance in response to proteolysis-inducing biological stimuli. Application of subtiligase to spatially resolved mapping of proteolysis inside living cells is currently limited by the need to use a cell impermeable peptide ester substrate, a problem that could be addressed by future substrate design and protein engineering efforts. Although efficient N-terminal modifying enzymes have been developed, no enzymatic tools have yet been reported that modify free C termini, likely due to the low intrinsic reactivity of the carboxylate group. This challenge may be overcome in the future through further genome mining or protein engineering efforts that enable development of enzymatic tools for selective C terminomics.

Selective proteomics of subcellular compartments and protein complexes

Localization of proteins to macromolecular complexes and subcellular compartments organizes cellular processes, allowing them to occur within specialized

environments and in the presence of an appropriate repertoire of biomolecules. Over the past decade, the development of genetically targetable, proximity-dependent labeling enzymes has transformed our understanding of the composition of the subproteomes of compartments and complexes [43–45] (Figure 2). Proximity labeling enzymes have been engineered from several different protein scaffolds but have in common that they generate a reactive species, such as a radical or an electrophile, that is covalently tethered to biotin. The reactive biotin species can nonspecifically tag proteins within a few nanometers of the proximity tagging enzyme. The small tagging radius is enforced by the short half-life of the activated biotin species (Figure 2a) or by the need for direct enzymatic transfer to a partner protein (Figure 2b). In proximity labeling experiments, the proximity tagging enzyme is targeted by genetic fusion to a protein or compartment of interest. Addition of a biotinylated substrate initiates enzymatic generation of the reactive biotin species and subsequent tagging of proximal proteins. Biotinylated proteins are then enriched on immobilized avidin and identified using LC-MS/MS. Proximity tagging enzymes include BioID [43,45], which generates 5′-biotinoyl-AMP that reacts with surface lysines; APEX, which generates a biotin phenoxyl radical that reacts primarily with surface tyrosines [44]; and the NEDDylator [46,47], PUP-IT [48], and EXCELL [49], which generate reactive electrophiles on proteins or peptides that are enzymatically transferred to surface lysine residues. Since their introduction, proximity tagging approaches have been widely applied to address an array of questions in cell biology. In a recent study, the BioID system was applied to generate a proximity biotinylation map of a human

Figure 2

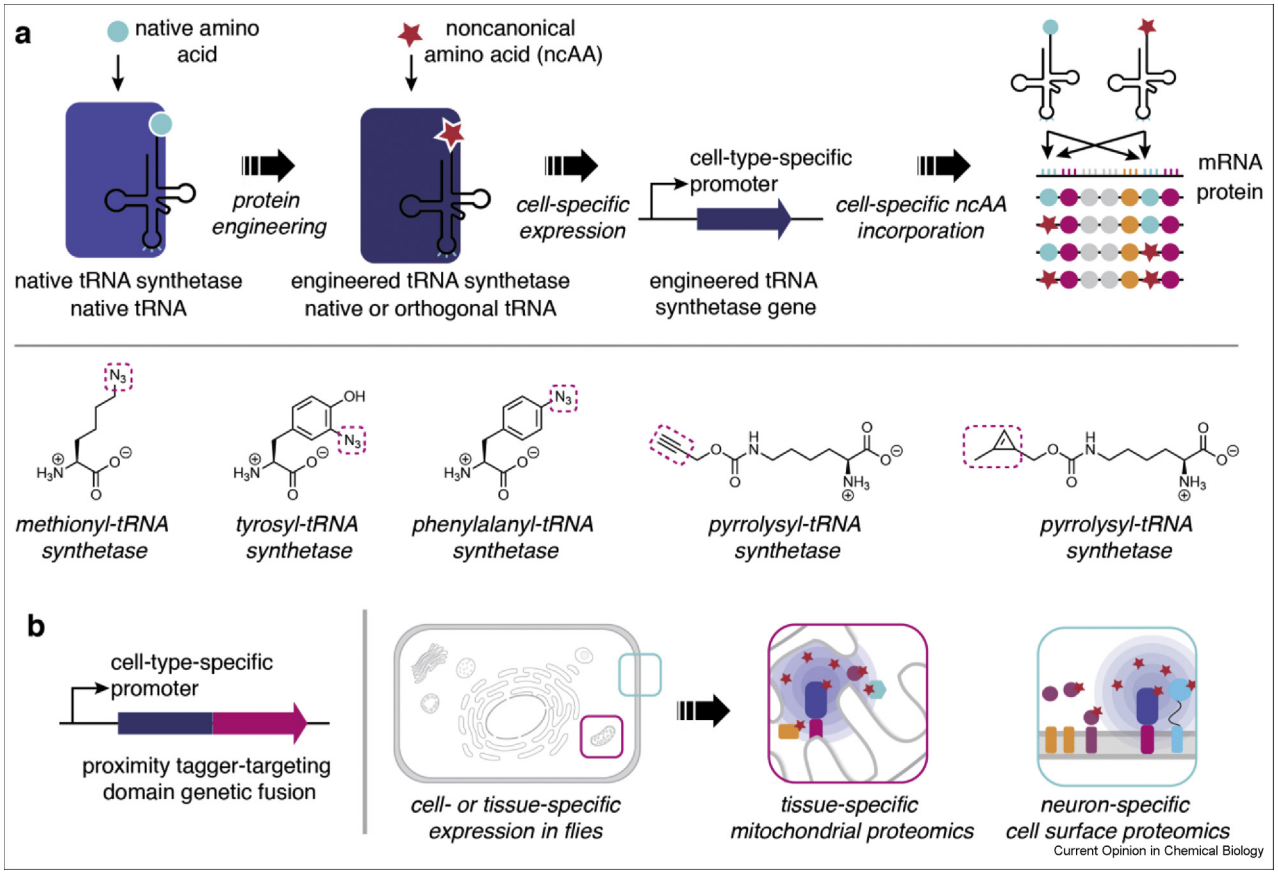


Engineered proximity tagging enzymes for selective proteomics of protein complexes and subcellular compartments. (a) Enzymes for generating diffusible reactive species to modify proximal proteins. Left: This group of proximity tagging enzymes accepts biotinylated probes as substrates and converts them to short half-life reactive species. Right: Examples of proximity taggers that generate diffusible reactive species. **(b)** Enzymes that directly transfer biotinylated tags to interacting proteins. Left: These enzymes accept biotinylated peptides or proteins as their substrates and use ATP to convert them to reactive electrophiles such as thioesters or phosphate esters that can be enzymatically transferred to lysine residues on proximal proteins. Right: Examples of direct-transfer proximity taggers and their substrates.

cell, with BioID fused to 192 proteins from 32 subcellular compartments [50]. This resulted in the assignment of localization for 4145 proteins and in the identification of 35,902 high confidence protein–protein interactions.

The two most widely adopted proximity tagging systems, BioID and APEX, have distinct advantages and limitations. BioID is advantageous because it requires only the nontoxic molecules biotin and ATP as substrates. However, it is limited by slow kinetics, generally requiring 18 h of labeling time before sufficient signal is obtained. In contrast, APEX has fast kinetics, tagging proteins on a second to minute timescale, but requires the toxic substrate H_2O_2. To overcome the slow kinetics of BioID, a yeast display screening approach was applied to identify a BioID variant, TurboID, with 15 mutations that increase its activity such that only a 10-min labeling time is required [51]. This improvement in activity made it possible to apply TurboID for proximity tagging in flies, worms, and plants [52], contexts in which BioID gave insufficient signal. Another enzyme, human arylamine N-acetyltransferase (NAT), has recently been developed for subcellular proximity labeling that avoids the use of toxic substrates and functions on fast timescales [53]. Arylamine NAT catalyzes the conversion of N-acetyl-N-hydroxyarylamines to N-acetoxyarylamines, which undergo rapid heterolytic cleavage of the N–O bond to generate nitrenium ions. These electrophilic ions are stable on a microsecond timescale and are highly reactive toward nucleophilic amino acid side chains. This strategy produced fast timescale labeling with

Figure 3



Engineered enzymes for cell-type-selective proteomics. (a) Incorporation of noncanonical amino acids (ncAAs) during protein synthesis. Top: Engineered tRNA synthetase enzymes generate tRNAs charged with bioorthogonally reactive ncAAs that can compete with native tRNAs during protein synthesis. Expression of these enzymes under cell-type-specific promoters tags newly synthesized proteins in only the targeted cell type. Bottom: Examples of bioorthogonally reactive ncAAs used in cell-type-specific proteomics. Modification of the reactive group with biotin after incorporation enables enrichment of the cell-type-specific proteome. **(b)** Proximity tagging enzymes for cell-type-specific proteomics. Left: Expression of a proximity tagging enzyme fused to a targeting domain or bait protein restricts tagging activity to a specific cell type and a specific subcellular compartment or protein complex. Right: Engineered peroxidases have been applied for tissue-specific mitochondrial proteomics and for neuron-specific cell surface proteomics.

subcellular resolution and has high potential for optimization by protein engineering and future application in selective proteomics studies.

Cell-type-specific selective proteomics

Proteomic characterization of tissues and organisms has provided significant biological insights but does not capture the diversity of biological function in different cell types. Efforts to purify specific cell types based on cell surface markers often compromise cellular structures and networks and may therefore lead to an inaccurate picture of the biological response. Methods to encode from which subcellular population of a protein is derived have been developed to overcome these challenges and to enable cell-selective proteomics (Figure 3). These approaches generally rely on cell-specific expression of enzymes that tag cellular proteins, enabling enrichment of cell-type-specific subproteomes from mixed cultures or whole organisms. Because these techniques rely on enzyme expression, protein engineering has had a strong impact on the field of cell-type-specific proteomics. Screening a library of mutant *E. coli* methionyl-tRNA synthetases led to the identification of NLL-MetRS, a variant that efficiently charges cognate tRNAs with azidonorleucine (ANL) [54] (Figure 3a). Incorporation of ANL into the proteome is restricted to cells expressing NLL-MetRS, enabling its application to cell-specific proteomics. After incorporation of ANL, cells are subjected to bioorthogonal noncanonical amino acid tagging (BONCAT) for biotinylation of proteins from the cell type of interest, enabling their enrichment and identification using LC-MS/MS. This method has been expanded for application in worms [55], flies [56], mammalian cell lines [57], and mice [58,59]. Another method, stochastic orthogonal recoding of translation (SORT) is conceptually similar but relies on a pyrrolysyl-tRNA synthetase/tRNA pair to incorporate bioorthogonally reactive noncanonical amino acids for biotinylation [60,61] (Figure 3a). SORT has been adapted for application in *E. coli*, mammalian cell lines, flies, and mice [62]. Recent studies have reported the application of noncanonical amino acid labeling to isolate mouse neuronal and glial proteomes [62] and to study changes in hippocampal neuron protein synthesis during long-term memory formation in mice [63].

Proximity tagging enzymes have been developed as tools for cell-type-specific proteomics that also provide subcellular spatial resolution of the proteome (Figure 3b). This is achieved by expressing a proximity tagging enzyme, such as APEX or HRP, that is targeted to a specific subcellular location, under a cell-type-specific promoter. This strategy has been applied to map the mitochondrial proteome across different *Drosophila* tissue types [64] and, more recently, to profile the cell surface proteome of the fly brain to identify new regulators of neuronal wiring [65].

Conclusions and future perspectives

Engineered protein tools have been deployed to map PTM sites, protein–protein interactions, subcellular protein localization, and tissue-specific and cell-specific proteomes with unprecedented depth and detail. Despite these successes, many protein tools remain limited by low activity in the cellular environment, precluding their use in live cells or whole animals. Developments in protein engineering, such as improved display techniques [66,67], microfluidics-based [68] and chip-based [69] screening, and phage-assisted continuous evolution [70], are poised to address these challenges and have already made inroads in broadening the applicability of some tools [51]. A second remaining limitation is the reliance of many engineered enzymes on cell-impermeable substrates, another obstacle to their use in live cells and animals. Design of new MCRs [71] and their combination with engineered protein tools are likely to address this challenge in the future. Moving forward, there are many opportunities to integrate different engineered protein tools to obtain tissue-specific, subcellular, and PTM-scale resolution in a single proteomics experiment. Future efforts toward protein engineering for selective proteomics will therefore propel the field of proteomics toward meeting the current challenge of moving beyond cataloging proteins toward assigning their biological functions. The results of these experiments are likely to provide a detailed and dynamic picture of cellular signaling that will advance our fundamental understanding of biological regulatory mechanisms and fuel translational science in the coming years.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank C Frazier, A Amiridis, S Coyle, and D Sashital for helpful discussions and comments on the manuscript. AMW acknowledges financial support from a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (1017065) and a Steenbock Career Award from the University of Wisconsin.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Beck A, Goetsch L, Dumontet C, Corvaïa N: **Strategies and challenges for the next generation of antibody-drug conjugates**. *Nat Rev Drug Discov* 2017, **16**:315–337.

2. Guedan S, Ruella M, June CH: **Emerging cellular therapies for cancer**. *Annu Rev Immunol* 2019, **37**:145–171.

3. Burslem GM, Crews CM: **Proteolysis-targeting chimeras as therapeutics and tools for biological discovery**. *Cell* 2020, **181**:102–114.

4. Dechecchi MC, Tamanini A, Cabrini G: **Molecular basis of cystic fibrosis: from bench to bedside**. *Ann Transl Med* 2018, **6**. 334–334.

5. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, *et al.*: **Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia**. *N Engl J Med* 2001, **344**:1031–1037.

6. Koren MJ, Scott R, Kim JB, Knusel B, Liu T, Lei L, Bolognese M, Wasserman SM: **Efficacy, safety, and tolerability of a monoclonal antibody to proprotein convertase subtilisin/kexin type 9 as monotherapy in patients with hypercholesterolaemia (MENDEL): a randomised, double-blind, placebo-controlled, phase 2 study**. *Lancet* 2012, **380**:1995–2006.

7. Tsai C-F, Wang Y-T, Yen H-Y, Tsou C-C, Ku W-C, Lin P-Y, Chen H-Y, Nesvizhskii AI, Ishihama Y, Chen Y-J: **Large-scale determination of absolute phosphorylation stoichiometries in human cells by motif-targeting quantitative proteomics**. *Nat Commun* 2015, **6**:6622–6628.

8. Chen X, Karnovsky A, Sans MD, Andrews PC, Williams JA: **Molecular characterization of the endoplasmic reticulum: insights from proteomic studies**. *Proteomics* 2010, **10**: 4040–4052.

9. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong S-E, Walford GA, Sugiana C, Boneh A, Chen WK, *et al.*: **A mitochondrial protein compendium elucidates complex I disease biology**. *Cell* 2008, **134**:112–123.

10. Youn J-Y, Dyakov BJA, Zhang J, Knight JDR, Vernon RM, Forman-Kay JD, Gingras A-C: **Properties of stress granule and P-body proteomes**. *Mol Cell* 2019, **76**:286–294.

11. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, Costello CE, Cravatt BF, Fenselau C, Garcia BA, *et al.*: **How many human proteoforms are there?** *Nat Chem Biol* 2018, **14**:206–214.

12. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M: **Global, in vivo, and site-specific phosphorylation dynamics in signaling networks**. *Cell* 2006, **127**:635–648.

13. Riley NM, Coon JJ: **Phosphoproteomics in the age of rapid and deep proteome profiling**. *Anal Chem* 2016, **88**:74–94.

14. Sharma K, D'Souza RCJ, Tyanova S, Schaab C, Wiśniewski JR, Cox J, Mann M: **Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling**. *Cell Rep* 2014, **8**:1583–1594.

15. Di Palma S, Zoumaro-Djayoon A, Peng M, Post H, Preisinger C, Munoz J, Heck AJR: **Finding the same needles in the haystack? A comparison of phosphotyrosine peptides enriched by immuno-affinity precipitation and metal-based affinity chromatography**. *J Proteomics* 2013, **91**:331–337.

16. Kettenbach AN, Gerber SA: **Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: application to general and phosphotyrosine-specific phosphoproteomics experiments**. *Anal Chem* 2011, **83**: 7635–7644.

17. Koerber JT, Thomsen ND, Hannigan BT, Degrado WF, Wells JA: **Nature-inspired design of motif-specific antibody scaffolds**. *Nat Biotechnol* 2013, **31**:916–921.

18. Mou Y, Zhou XX, Leung K, Martinko AJ, Yu J-Y, Chen W, Wells JA: **Engineering improved antiphosphotyrosine antibodies based on an immunoconvergent binding motif**. *J Am Chem Soc* 2018, **140**:16615–16624.

19. Bian Y, Li L, Dong M, Liu X, Kaneko T, Cheng K, Liu H, Voss C, Cao X, Wang Y, *et al.*: **Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder**. *Nat Chem Biol* 2016, **12**:959–966.

20. Matsumoto ML, Wickliffe KE, Dong KC, Yu C, Bosanac I, Bustos D, Phu L, Kirkpatrick DS, Hymowitz SG, Rape M, *et al.*: **K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody**. *Mol Cell* 2010, **39**:477–484.

21. Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, Dugger D, Gordon N, Sidhu SS, Fellouse FA, *et al.*: **Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies**. *Cell* 2008, **134**:668–678.

22. ** Yau RG, Doerner K, Castellanos ER, Haakonsen DL, Werner A, Wang N, Yang XW, Martinez-Martin N, Matsumoto ML, Dixit VM, *et al.*: **Assembly and function of heterotypic ubiquitin chains in cell-cycle and protein quality control**. *Cell* 2017, **171**: 918–933.e20.

This article describes a nice example of the development of a bispecific antibody to recognize proteoforms bearing two specific PTMs. The antibody developed is applied to uncover new facets of ubiquitin signaling.

23. Saxon E, Bertozzi CR: **Cell surface engineering by a modified Staudinger reaction**. *Science* 2000, **287**:2007–2010.

24. Darabedian N, Gao J, Chuh KN, Woo CM, Pratt MR: **The metabolic chemical reporter 6-azido-6-deoxy-glucose further reveals the substrate promiscuity of O-GlcNAc transferase and catalyzes the discovery of intracellular protein modification by O-glucose**. *J Am Chem Soc* 2018, **140**:7092–7100.

25. Khidekel N, Ficarro SB, Clark PM, Bryan MC, Swaney DL, Rexach JE, Sun YE, Coon JJ, Peters EC, Hsieh-Wilson LC: **Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics**. *Nat Chem Biol* 2007, **3**: 339–348.

26. Chaubard J-L, Krishnamurthy C, Yi W, Smith DF, Hsieh-Wilson LC: **Chemoenzymatic probes for detecting and imaging fucose-α(1-2)-galactose glycan biomarkers**. *J Am Chem Soc* 2012, **134**:4489–4492.

27. Wen L, Zheng Y, Jiang K, Zhang M, Kondengaden SM, Li S, Huang K, Li J, Song J, Wang PG: **Two-step chemoenzymatic detection of N-acetylneuraminic acid-α(2-3)-galactose glycans**. *J Am Chem Soc* 2016, **138**:11473–11476.

28. Sun T, Yu S-H, Zhao P, Meng L, Moremen KW, Wells L, Steet R, Boons G-J: **One-step selective exoenzymatic labeling (SEEL) strategy for the biotinylation and identification of glycoproteins of living cells**. *J Am Chem Soc* 2016, **138**:11575–11582.

29. ** Malaker SA, Pedram K, Ferracane MJ, Bensing BA, Krishnan V, Pett C, Yu J, Woods EC, Kramer JR, Westerlind U, *et al.*: **The mucin-selective protease StcE enables molecular and functional analysis of human cancer-associated mucins**. *Proc. Natl. Acad. Sci. U.S.A.* 2019, **116**:7278–7287.

The authors characterize the specificity of a bacterial mucinase, StcE, and demonstrate the utility of the enzyme in various proteomics applications, including digestion, enrichment, and glycosylation mapping.

30. Hertz NT, Berthet A, Sos ML, Thorn KS, Burlingame AL, Nakamura K, Shokat KM: **A neo-substrate that amplifies catalytic activity of Parkinson's-disease-related kinase PINK1**. *Cell* 2013, **154**:737–747.

31. Islam K, Chen Y, Wu H, Bothwell IR, Blum GJ, Zeng H, Dong A, Zheng W, Min J, Deng H, *et al.*: **Defining efficient enzyme-cofactor pairs for bioorthogonal profiling of protein methylation**. *Proc. Natl. Acad. Sci. U.S.A.* 2013, **110**:16778–16783.

32. Nguyen UTT, Guo Z, Delon C, Wu Y, Deraeve C, Fränzel B, Bon RS, Blankenfeldt W, Goody RS, Waldmann H, *et al.*: **Analysis of the eukaryotic prenylome by isoprenoid affinity tagging**. *Nat Chem Biol* 2009, **5**:227–235.

33. Yang C, Mi J, Feng Y, Ngo L, Gao T, Yan L, Zheng YG: **Labeling lysine acetyltransferase substrates with engineered enzymes and functionalized cofactor surrogates**. *J Am Chem Soc* 2013, **135**:7791–7794.

34. Gibson BA, Zhang Y, Jiang H, Hussey KM, Shrimp JH, Lin H, Schwede F, Yu Y, Kraus WL: **Chemical genetic discovery of PARP targets reveals a role for PARP-1 in transcription elongation**. *Science* 2016, **353**:45–50.

35. Carter-O'Connell I, Jin H, Morgan RK, Zaja R, David LL, Ahel I, Cohen MS: **Identifying family-member-specific targets of mono-ARTDs by using a chemical genetics approach**. *Cell Rep* 2016, **14**:621–631.

36. Carter-O'Connell I, Jin H, Morgan RK, David LL, Cohen MS: **Engineering the substrate specificity of ADP-ribosyltransferases for identifying direct protein targets**. *J Am Chem Soc* 2014, **136**:5201–5204.

37. ** Carter-O'Connell I, Vermehren-Schmaedick A, Jin H, Morgan RK, David LL, Cohen MS: **Combining chemical genetics with proximity-dependent labeling reveals cellular targets of poly(ADP-ribose) polymerase 14 (PARP14)**. *ACS Chem Biol* 2018, **13**:2841–2848.
This study combines a bump-hole chemoenzymatic capture strategy for ADP-ribosylated proteins with a BioID proximity tagging approach to identify interactors of the ADP-ribosyltransferase PARP14, providing high-confidence substrate identification for this enzyme.

38. Mahrus S, Trinidad JC, Barkan DT, Sali A, Burlingame AL, Wells JA: **Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini**. *Cell* 2008, **134**:866–876.

39. Xu G, Shin SBY, Jaffrey SR: **Chemoenzymatic labeling of protein C-termini for positive selection of C-terminal peptides**. *ACS Chem Biol* 2011, **6**:1015–1020.

40. Weeks AM, Wells JA: **Subtiligase-catalyzed peptide ligation**. *Chem Rev* 2020, **120**:3127–3160.

41. ** Weeks AM, Wells JA: **Engineering peptide ligase specificity by proteomic identification of ligation sites**. *Nat Chem Biol* 2018, **14**:50–57.
This paper reports a set of subtiligase variants with expanded N-terminal specificity, as well as a protein engineering platform based on proteome-derived peptide libraries that is generalizable to other peptide ligases.

42. ** Weeks AM, Byrnes JR, Lui I, Wells JA: **Mapping proteolytic neo-N termini at the surface of living cells**. *bioRxiv* 2020, https://doi.org/10.1101/2020.02.23.961904.
This study deploys subtiligase N terminomics in the context of live cells for the first time to identify substrates of proteolytic cleavage at the plasma membrane.

43. 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* 2012, **196**:801–810.

44. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, Ting AY: **Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging**. *Science* 2013, **339**:1328–1331.

45. * Samavarchi-Tehrani P, Samson R, Gingras A-C: **Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches**. *Mol Cell Proteomics* 2020, **19**:757–773.
This manuscript provides and in-depth review of proximity tagging enzymes.

46. Hill ZB, Pollock SB, Zhuang M, Wells JA: **Direct proximity tagging of small molecule protein targets using an engineered NEDD8 ligase**. *J Am Chem Soc* 2016, **138**:13123–13126.

47. Zhuang M, Guan S, Wang H, Burlingame AL, Wells JA: **Substrates of IAP ubiquitin ligases identified with a designed orthogonal E3 ligase, the NEDDylator**. *Mol Cell* 2013, **49**:273–282.

48. * Liu Q, Zheng J, Sun W, Huo Y, Zhang L, Hao P, Wang H, Zhuang M: **A proximity-tagging system to identify membrane protein-protein interactions**. *Nat Methods* 2018, **15**:715–722.
The authors report a new system based on the bacterial Pupylation machinery that is deployed for capture of protein–protein interactions at the cell surface.

49. * Ge Y, Chen L, Liu S, Zhao J, Zhang H, Chen PR: **Enzyme-mediated intercellular proximity labeling for detecting cell-cell interactions**. *J Am Chem Soc* 2019, **141**:1833–1837.
The authors report a sortase-based proximity tagging enzyme that is able to capture interactions between cells.

50. ** Go CD, Knight JDR, Rajasekharan A, Rathod B, Hesketh GG, Abe KT, Youn J-Y, Samavarchi-Tehrani P, Zhang H, Zhu LY, *et al.*: **A proximity biotinylation map of a human cell**. *bioRxiv* 2019, https://doi.org/10.1101/796391.
This is the most comprehensive proximity tagging study of protein interactions and localization in human cells, with nearly 200 human proteins used as bait in a BioID fusion construct.

51. ** Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY: **Efficient proximity labeling in living cells and organisms with TurboID**. *Nat Biotechnol* 2018, **36**:880–887.
The authors report a yeast display-based strategy for directed evolution of the BioID enzyme, resulting in a 100-fold improvement of biotinylation kinetics and increased activity that expands the applicability of BioID-like proximity tagging to yeast, worms, and flies.

52. ** Mair A, Xu S-L, Branon TC, Ting AY, Bergmann DC: **Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID**. *Elife* 2019, **8**:451.
This paper reports the first study of plant organellar proteomes using TurboID proximity tagging.

53. ** Kleinpenning F, Eising S, Berkenbosch T, Garzero V, Schaart JM, Bonger KM: **Subcellular protein labeling by a spatially restricted arylamine N-acetyltransferase**. *ACS Chem Biol* 2018, **13**:1932–1937.
The authors report a new strategy for generation of a diffusible reactive probe (a nitrenium ion) for proximity tagging in live cells.

54. Ngo JT, Champion JA, Mahdavi A, Tanrikulu IC, Beatty KE, Connor RE, Yoo TH, Dieterich DC, Schuman EM, Tirrell DA: **Cell-selective metabolic labeling of proteins**. *Nat Chem Biol* 2009, **5**:715–717.

55. Yuet KP, Doma MK, Ngo JT, Sweredoski MJ, Graham RLJ, Moradian A, Hess S, Schuman EM, Sternberg PW, Tirrell DA: **Cell-specific proteomic analysis in *Caenorhabditis elegans***. *Proc. Natl. Acad. Sci. U.S.A.* 2015, **112**:2705–2710.

56. Erdmann I, Marter K, Kobler O, Niehues S, Abele J, Müller A, Bussmann J, Storkebaum E, Ziv T, Thomas U, *et al.*: **Cell-selective labelling of proteomes in *Drosophila melanogaster***. *Nat Commun* 2015, **6**:7521.

57. Ngo JT, Schuman EM, Tirrell DA: **Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells**. *Proc. Natl. Acad. Sci. U.S.A.* 2013, **110**:4992–4997.

58. Alvarez-Castelao B, Schanzenbächer CT, Hanus C, Glock C, Tom Dieck S, Dörrbaum AR, Bartnik I, Nassim-Assir B, Ciirdaeva E, Mueller A, *et al.*: **Cell-type-specific metabolic labeling of nascent proteomes in vivo**. *Nat Biotechnol* 2017, **35**:1196–1201.

59. Alvarez-Castelao B, Schanzenbächer CT, Langer JD, Schuman EM: **Cell-type-specific metabolic labeling, detection and identification of nascent proteomes in vivo**. *Nat Protoc* 2019, **14**:556–575.

60. Elliott TS, Townsley FM, Bianco A, Ernst RJ, Sachdeva A, Elsässer SJ, Davis L, Lang K, Pisa R, Greiss S, *et al.*: **Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal**. *Nat Biotechnol* 2014, **32**:465–472.

61. Elliott TS, Bianco A, Townsley FM, Fried SD, Chin JW: **Tagging and enriching proteins enables cell-specific proteomics**. *Cell Chem. Biol.* 2016, **23**:805–815.

62. ** Krogager TP, Ernst RJ, Elliott TS, Calo L, Beránek V, Ciabatti E, Spillantini MG, Tripodi M, Hastings MH, Chin JW: **Labeling and identifying cell-specific proteomes in the mouse brain**. *Nat Biotechnol* 2018, **36**:156–159.
The authors report improvements in stochastic orthogonal recoding of translation (SORT) to expand its applicability to transgenic mice.

63. ** Evans HT, Bodea L-G, Götz J: **Cell-specific non-canonical amino acid labelling identifies changes in the de novo proteome during memory formation**. *Elife* 2020, **9**:16432.
NLL-MetRS is expressed in mouse hippocampal neurons to capture their proteomes via BONCAT, providing inside into long-term memory formation.

64. Chen C-L, Hu Y, Udeshi ND, Lau TY, Wirtz-Peitz F, He L, Ting AY, Carr SA, Perrimon N: **Proteomic mapping in live**

***Drosophila* tissues using an engineered ascorbate peroxidase**. *Proc. Natl. Acad. Sci. U.S.A.* 2015, **112**:12093–12098.

65. ** Li J, Han S, Li H, Udeshi ND, Svinkina T, Mani DR, Xu C, Guajardo R, Xie Q, Li T, *et al.*: **Cell-surface proteomic profiling in the fly brain uncovers wiring regulators**. *Cell* 2020, **180**: 373–386. e15.
This study combines proximity labeling based on horseradish peroxidase with cell-type-specific expression of the peroxidase enzyme to map wiring regulators in the fly brain.

66. Chen I, Dorr BM, Liu DR: **A general strategy for the evolution of bond-forming enzymes using yeast display**. *Proc. Natl. Acad. Sci. U.S.A.* 2011, **108**:11399–11404.

67. Atwell S, Wells JA: **Selection for improved subtiligases by phage display**. *Proc. Natl. Acad. Sci. U. S. A.* 1999, **96**: 9497–9502.

68. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret J-C, Marquez M, Klibanov AM, Griffiths AD, Weitz DA: **Ultrahigh-throughput screening in drop-based microfluidics for directed evolution**. *Proc. Natl. Acad. Sci. U.S.A.* 2010, **107**: 4004–4009.

69. Chen B, Lim S, Kannan A, Alford SC, Sunden F, Herschlag D, Dimov IK, Baer TM, Cochran JR: **High-throughput analysis and protein engineering using microcapillary arrays**. *Nat Chem Biol* 2016, **12**:76–81.

70. Esvelt KM, Carlson JC, Liu DR: **A system for the continuous directed evolution of biomolecules**. *Nature* 2011, **472**: 499–503.

71. Westcott NP, Fernandez JP, Molina H, Hang HC: **Chemical proteomics reveals ADP-ribosylation of small GTPases during oxidative stress**. *Nat Chem Biol* 2017, **13**:302–308.