Check for updates

ChemBioChem

Perspective
doi.org/10.1002/cbic.202400005

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Special Collection

Climbing into their Skin to Understand Contextual Protein–Protein Associations and Localizations: Functional Investigations in Transgenic Live Model Organisms

Marcus J. C. Long[a] and Yimon Aye*[b]

Borrowing some quotes from Harper Lee's novel "To Kill A Mockingbird" to help frame our manuscript, we discuss methods to profile local proteomes. We initially focus on chemical biology regimens that function in live organisms and use reactive biotin species for this purpose. We then consider ways to add new dimensions to these experimental regimens, principally by releasing less reactive (i.e., more selective) (preter)natural electrophiles. Although electrophile release methods may have lower resolution and label fewer proteins than biotinylation methods, their ability to probe simultaneously protein function and locale raises new and interesting possibilities for the field.

One of the most beloved quotes in literature is the sage advice given by Atticus Finch to his daughter Scout, that you need "to climb into his skin and walk around in it" to really understand a person. In the post genomic era, one of our equivalents has become using engineered proteins to literally climb beneath the skin/cuticle/scales/exoskeleton of specific organisms into specific cells, and investigate the (sub)proteomes they contain as a function of specific locales and stimuli with nanometer resolution. This information can inform on context-specific changes in locale, and in select few cases, also function, that occur during specific signaling events, or at specific signalosomes. These tools that were often developed through protein engineering, and pioneered in cell culture, are starting to push the envelope to investigate protein localization and trafficking changes in whole organisms. As these tools combine genetics and chemical biology, they can be performed in a tissue- and locale–specific manner, offering a huge range of biological contexts to be explored. Here we will start by briefly discussing specific methods of mapping local proteomes, and advantages and current limitations with their use. Subsequently, we will discuss specific examples of where and how these tools have been deployed in model organisms to understand localized interactomes. We will finally discuss similar methods that may add new functional dimensions to locale-directed proximity profiling experiments.

[a] Dr. M. J. C. Long
University of Lausanne (UNIL), Epalinges, 1066 Switzerland

[b] Dr. Y. Aye
Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, 1015 Switzerland
E-mail: yimon.aye@epfl.ch

EFMC Collection: Advances in chemical probing concepts for chemical biology applications

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

"Boo was our Neighbor. He gave us two Soap Dolls, a Broken Watch and Chain, a Pair of Good–Luck Pennies, and our Lives": Methods to Identify Local Interactomes

The cellular environment is particularly complex. There are potentially millions of possible proteoforms (including splice variants, truncated versions, and the like), and each may have peculiar interactomes, often extending way beyond binary associations. Thus, the number of potential interactions in cells and organisms is effectively infinite. Nonetheless, just as is the case in our daily lives, specific associations between select individuals form preferentially. Indeed, the nature of the cellular milieu can favor associations that are not typically present under standard *in vitro* conditions, such as lysates or purified proteins. This difference in stabilities can be due to localized build–up of specific proteins boosting concentrations and aiding association, reduction of diffusive capacities due to crowding favoring low entropy situations, and changes in pH, among a plethora of other variables.[1] Indeed, several known "quinary" protein–protein interactions with association constants close to 100 μM *in vitro* are now established as important for cellular function.[2–4] Traditional methods to identify protein–protein interactions, in which cells are lysed and then an association between a bait protein is tested by co-immunoprecipitation (co-IP),[5] are likely insufficient to detect such interactions (Figure 1A). These traditional methods may also lack sensitivity to detect complexes of low fractional occupancy as can occur in many signaling pathways.[6] It has thus become clear that new methods to identify associations that can be carried out under more native situations, i.e., where the association is assayed in a native cellular environment, are in need.

Several different experimental strategies have now been disclosed to meet this purpose. Basic chemical biology methods, such as chemical cross–linking (Figure 1B), are also commonly used. However, these methods offer limited spatial or temporal control. Imaging methods, such as proximity

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH

doi.org/10.1002/cbic.202400005

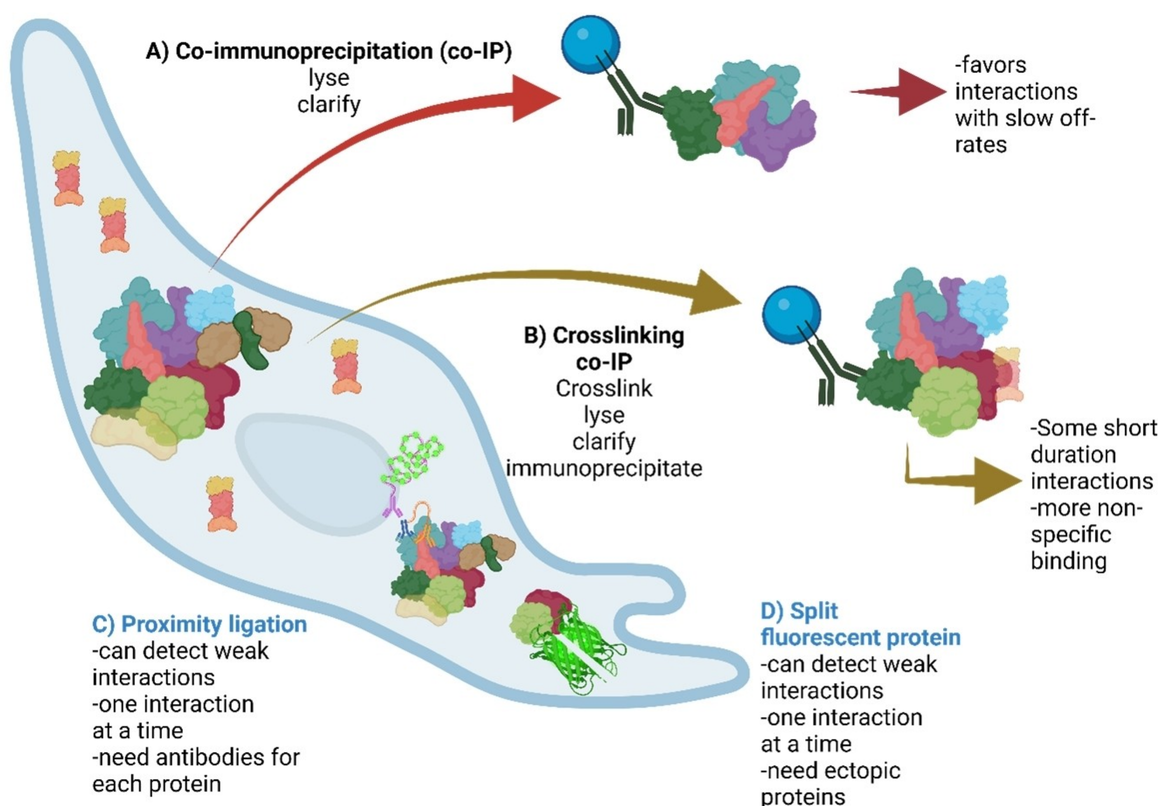

Figure 1. A selection of different methods to assess protein–protein interactions in cells. Method titled in **dark** require cell lysis; those in **blue** are performed in cells. A) Co-immunoprecipitation (co-IP) involves cell lysis, clarification, and IP. Associating proteins can be analyzed by LC–MS/MS or western blot, for example. B) is similar to A) but cells or lysates are treated with a chemical cross–linker. C) Proximity ligation assay (PLA) uses specifically–modified secondary antibodies that can template rolling circle amplification (RCA)-based polymerase chain reaction (PCR), if the antibodies are in proximity to each other. D) Split fluorescent proteins are specifically–designed fluorescent protein fragments that when in contact with each other can form an active fluorophore.

ligation assay (PLA) (Figure 1C), are also available.[7,8] PLA works on pairs of endogenous proteins, and can be used for assessing interactions as a function of locale. However, PLA does not have a huge multiplexing capability and requires well–validated, primary antibodies of high specificity for each protein. Two hybrid methods, including split ubiquitin[9] and split GFP[10] (Figure 1D), beyond split transcription factors, are also commonly used, and can be performed in an unbiased manner on "global proteomes". However, several of these methods are really only applicable to pairwise interactions and the setup necessitates overexpressed proteins. Thus, methods that can screen a large number of interactions in an unbiased manner, and are applicable to endogenous proteins, particularly remain at a premium.

The most recent years have witnessed the emergence of proteomics–based methods that can profile subcellular or sub–tissue localization of a large number of proteins. Methods centered around fractionation of cells[11] or tissues[12] are commonly used. However, among all approaches available to date, biotinylation–based proximity–labeling methods have arguably become the methods of choice to define subinteractomes and (subpopulation)cell–type–specific expression patterns (Figure 2A). The reason is closely linked to biotinylation's general utility, suitability to a wide range of cells and model organisms, high sensitivity, and broad applicability. In biotinylation methods, an enzyme that generates a reactive biotin–derived species is expressed in a desired locale, and locale–specific proteins are non–specifically biotinylated, and then enriched and identified by mass spectrometry (MS), relative to specific controls. There are also manifold technical reasons why on the whole the field has gravitated to biotinylation methods: (1) they do not require the use of specialized antibodies or other complex protein-specific detection methods; (2) the workflow is readily compatible with MS-based approaches allowing potentially a window into the whole proteome in one experiment; (3) the high affinity of the biotin streptavidin interaction allows high sensitivity; and (4) the high reactivity of the biotin-based species released in locale–specific contexts (and short diffusion distance, and ideally relatively–brief duration needed for labeling) renders the spatial and temporal resolution of biotinylation processes highly favorable.

Many early biotinylation methods used engineered enzymes called APEX that needed high concentrations of exogenous peroxide to become activated. APEX enzymes use biotin–phenol as substrates creating ephemeral reactive phenol radicals that covalently label proteins non–discriminately and with a very short diffusion distance (with labeling radii in nanometers, although precise range remains equivocal, ranging from 1–10 nm[13] to 200–300 nm,[14] Figure 2A). APEX enzymes are particularly rapid and give sufficient labeling for proteomics

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH
14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

doi.org/10.1002/cbic.202400005
Chemistry Europe
European Chemical Societies Publishing

14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

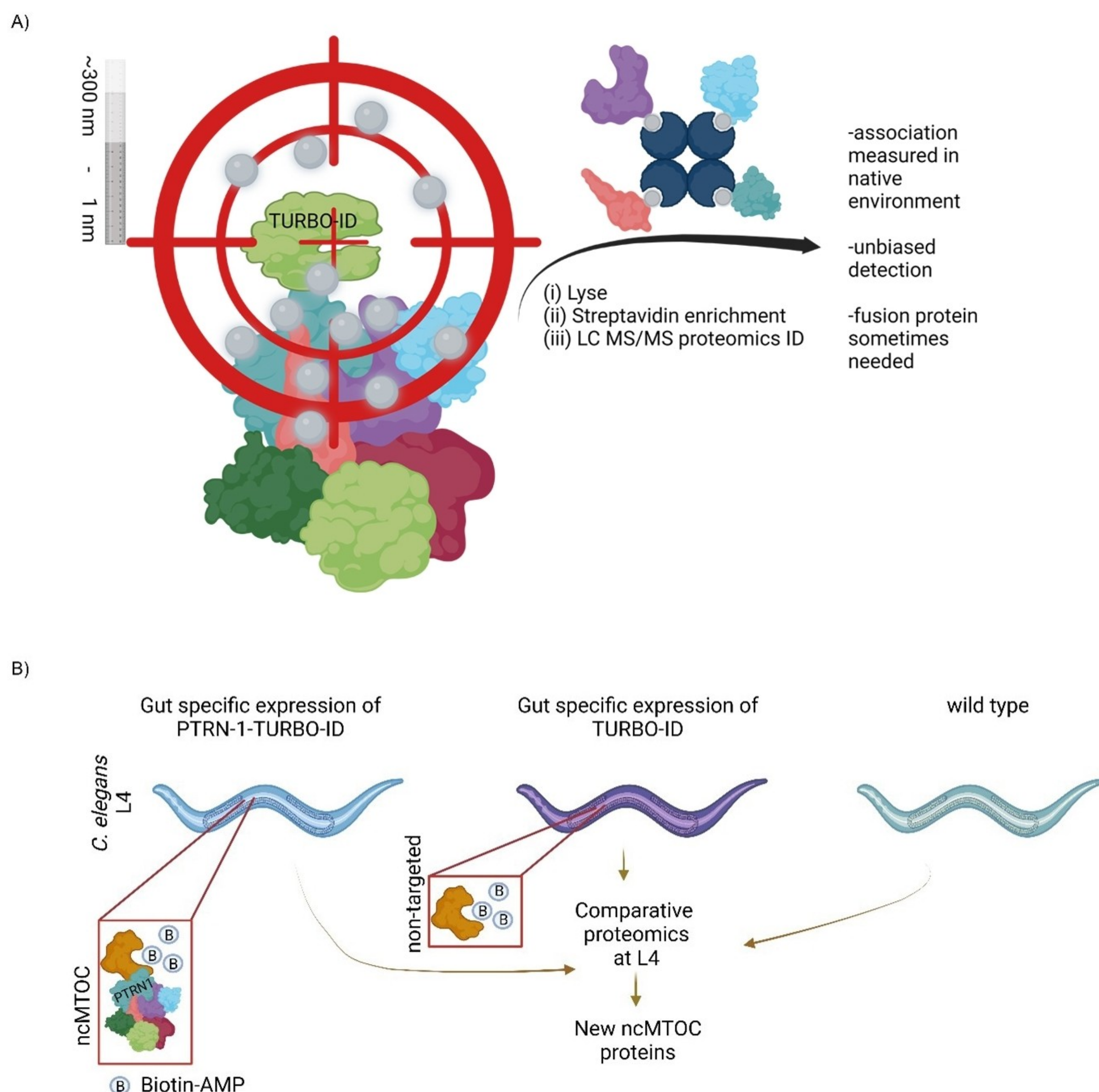

Figure 2. A) TURBO-ID liberates reactive biotin from a point source that can be localized to a specific protein or locale, in cells or in specific animal tissues. This setup can detect proteins resident in specific locales or tissues, or proteins in proximity to a specific protein as a function of different tissues. B) Comparison of proteins biotinylated in nematode worms with gut–specific expression of PTRN-1 – TURBO-ID versus worms with gut–specific expression of (TURBO-ID) TURBO-ID alone, and wild–type (N2) worms, allow discovery of new ncMTOC proteins in *C. elegans*.[27]

or imaging–based experiments in relatively short times (on the order of minutes post exposure to peroxide). This setup allowed some early work on model organisms, mostly isolated organs and tissues.[15] However, the necessity to treat with peroxide was severely limiting as peroxide can rapidly alter signaling pathways and interactomes itself.[16,17] An alternative method harnessing a series of biotinylating–enzyme variants, BioID–that creates biotin–AMP as a reactive species and does not need pretreatment with peroxide to be active–was for many years rarely applied to model organisms, likely due to its sluggish turnover rates. However, recent incarnations, of Bio-ID, such as TURBO- and ULTRA-ID, have relatively fast kinetics (labeling occurs in tens of minutes to hours, with labeling radii of ~40 nm).[18] Moreover, as there is no need for peroxide treatment, longer biotinylation reaction times are possible, potentially allowing capture of weaker protein–protein associations, or proteins that associate in a specific locale/with a specific protein only under specific circumstances. The same favorable properties have also allowed application to model organisms, as discussed below.

However, such improvements bring some negative aspects that also need to be considered. TURBO- and ULTRA-ID[19] are effectively constitutively active, providing biotin is present, as it is in most culture media, and many food sources for model organisms. Thus, temporal control, that may be particularly desired in biotinylation experiments, is inherently less tight for BioID-based methods than with APEX-based methods. Several systems have been deployed to overcome these issues,

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH

Chemistry Europe
European Chemical Societies Publishing

including inducible expression, use of biotin–free/depleted media or food, and examining only specific developmental stages of the animals.[20] Other approaches that are, perhaps, less clearly applicable to the gamut of model organisms are photoinducible TURBO-ID, phototurbo,[21] and LOV-TURBO.[22] Furthermore, excessive biotinylation has been proposed to introduce toxicity issues, either through depletion of endogenous biotin, or through biotinylation events altering specific protein's functions. The high constitutive activity of TURBO-ID may also increase the practical labeling radius compared to BioID.[18,23] Regardless of these aspects, it is important to bear in mind that very carefully considered controls need to be deployed when designing biotinylation–based profiling experiments; these may be intrinsic to the specific experimental goal.

Crawling into the Subject's Skin: Biotinylation–Based Labeling Methods

We now go through examples of how biotinylation methods have been deployed in model organisms, particularly mice and worms. We focus on methods that have been used in whole organisms in conjunction with native systems to localize the reactive biotinylating enzyme. We will thus avoid systems where biotinylation–based proximity labeling was performed on extracts or more artificial systems, such as tissue homogenates, or isolated organs.[24,25] It should be noted that deployment of TURBO-ID and ULTRA-ID in model organisms represents a particular challenge, in part because of the varied and complex tissue structures and cellular diversity, but also because these methods typically require knock-in, or transgenesis, that can take much longer to generate than simple lines in cultured cells. Moreover, different organisms require different growth conditions, including different feedstuffs and temperatures. The latter can severely affect the activity of TURBO-ID, for instance.[26]

Proximity labeling lends itself to both tissue–specific and protein/subcellular–locale–specific labeling, due to the short radius of coverage of the reactive biotin species and the possibility of the biotinylating enzyme to be expressed in specific locales.[27] One recent example of protein–specific interactome mapping at the tissue specific level is the use of TURBO-ID to identify proteins interacting with non–centrosomal microtubule–organizing centers (ncMTOCs), specifically in *C. elegans* gut (Figure 2B). ncMTOCs are structures, where microtubules grow and localize, in differentiated cells. These structures are distinct from the centrosome, where microtubules usually localize in non–differentiated cells. Using a *C. elegans* strain that expresses a microtubule minus–end binding protein that localizes to the apical ncMTOC, PTRN-1, fused to TURBO-ID expressed selectively in the gut, versus a strain that expresses non-PTRN-1-fused TURBO-ID, under an identical promoter, the authors investigated ncMTOC-associating proteins. Imaging of targeted biotinylation in *C. elegans* embryos showed, gut–specific, but different patterns of biotin–labeling for the two strains. The strain encoding non-fused TURBO-ID showed diffuse labeling of gut–specific cells, whereas that encoding PTRN-1–TURBO-ID showed apical–specific biotinylation patterns, consistent with the localization of PTRN-1. Moreover, from whole *C. elegans* lysates, proteins differentially labeled in nematodes expressing the two different TURBO-ID constructs, were able to be enriched using streptavidin beads. Proteomics analysis of proteins selectively enriched from PTRN-1-TURBO-ID-expressing animals showed a large number of cytoskeletal proteins, constituting potential ncMTOC interactors. Among those captured hits, VAB-10B and WDR-62, two proteins that had not been previously linked to ncMTOC, were shown to be important for ncMTOC, but not centrosomal function.

Key elements of these experiments were the use of biotin–depleted bacteria as a food source for *C. elegans*, to limit labeling background; the expression of TURBO-ID in the gut not fused to PTRN-1 to account for non-specific labeling by TURBO-ID; and the fact that labeling was detectable in whole lysates as opposed to isolated tissues. The authors, however, also noted that strong overexpression of PTRN-1 was necessary to achieve high labeling efficiency, which could potentially have affected homoestasis, although no growth defects were reported. Thus, the activity of TURBO-ID likely can be limiting, at least in some instances. It should be further noted that as TURBO-ID was constitutively active during embryonic development (and proteomics experiments were performed on L4 animals), one cannot be sure at which growth stage specific proteins were labeled. Of technical note, to ensure optimal comparison between two samples, it is usually better to compare data from a fusion protein (PTRN-1-TURBO-ID, in this example) and a split fusion protein, such as one bearing a P2 A within the linker region (e.g., PTRN-1-P2A-TURBO-ID). Beyond this case study from *C. elegans* gut, other recent studies have shown applicability to almost all *C. elegans* tissues.[28]

Proximity labeling has also been applied to mice.[25,29,30] Mice inducibly expressing TURBO-ID in CamK2a neurons were used to identify local protein abundance and pathway signatures in CamK2a neurons in different regions of the brain (Figure 3A).[29] A modified ELISA detection protocol was also used to detect changes in phosphorylation of specific proteins in specific regions. In the same study, by expressing TURBO-ID either in astrocytes, the predominant form of glial cell in the brain, or CamK2a neurons, protein components within these two cell types were delineated. Critically, these experiments were performed in live mice that were shown to undergo minimal perturbation of brain function due to TURBO-ID expression.

One important question to arise from these experiments is how to account for variations in TURBO-ID expression levels across different regions and tissues. In the studies above, data were normalized for TURBO-ID expression, a logical approach that assumes that TURBO-ID expression is dominant for total proteome labeling. This conclusion has been drawn in several different model organisms, including *Xenopus laevis*.[31] However, how the activity of TURBO-ID varies depending on pH, ionic strength, and other microenvironment-specific factors that could change across different cell types, is less well known. These are important parameters to define and will help better normalization protocols in the future.

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH

14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Chemistry Europe
European Chemical Societies Publishing

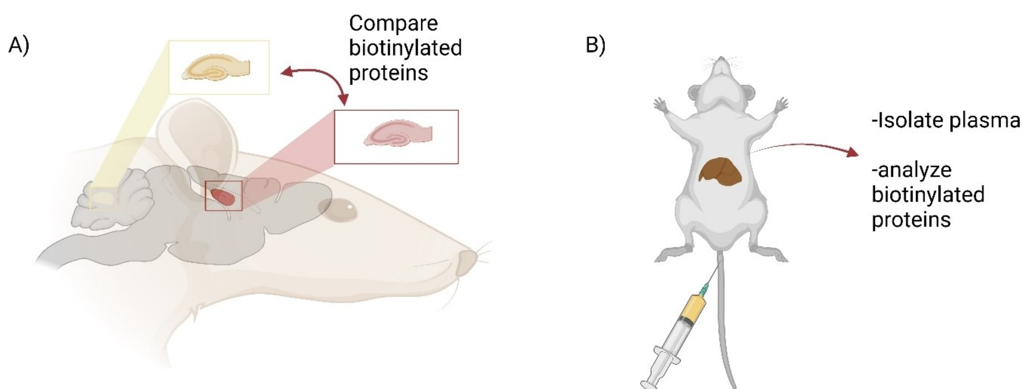

Figure 3. A) Expression of TURBO-ID in all CamK2a neurons followed by excision of different regions of the brain allows mapping of tissue-specific expression patterns.[29] B) Introducing Sec61b–TURBO-ID expression in the liver through tail vein injection of adenovirus, allows profiling of the liver secretome in live mice.[30]

The mice used in the above experiments were derived from a transgenic line where TURBO-ID expression was induced in specific cells by treatment with tamoxifen that drove tissue–specific expression of a recombinase. As noted above, this is a common way to avoid the constitutive activity of TURBO-ID. TURBO-ID has also been expressed in mice livers using adenoviral viral delivery via tail vein injection, a common way to introduce transgenes in the liver. Using this method, Sec61b-TURBO-ID was introduced into mice livers (Figure 3B).[30] Sec61b is a central component of the protein translocation apparatus, and as such, proteins that are secreted pass in close proximity to this protein. In the Sec61b-TURBO-ID mice, proteins secreted from the liver will thus be biotinylated. Several days post infection, biotinylation of the secreted proteome in plasma was measured. Although 81 % of the secreted proteins were similar between studies from primary hepatocytes, several proteins were uniquely detected in those derived from mice liver samples. Subsequent applications of this approach identified elevated secretion of proteins associated with diabetes and insulin resistance in a mouse model of diabetes.

As with all non–targeted methods, it is of course possible that some of the proteins identified within the secretome could have been attributed to aberrant viral incorporation. Perhaps more importantly, the effect of TURBO-ID on liver–cell viability is relatively difficult to determine from this experiment. The authors showed that there was little upregulation of cleaved Caspase 3, a marker of apoptosis, under their conditions. However, upregulation of lytic cell–death pathways could have biased secretomes, and this may not have been detected by the respective controls deployed.

Thus, biotinylation methods effectively constitute the molecular version of crawling into a person's skin. Unsurprisingly, their deployment will likely become ever more intricate, and promises to lead to improved understanding of several important biological processes through defining precise interactomes as functions of specific stimuli and contexts.

Walking around in the Subject's Skin: Using Natural Electrophiles to Profile Proximity

It is at this point where we need to consider all the words uttered by Atticus Finch: he urged his daughter not only to crawl within the skin, but also to walk around in it. Although a stretch, the latter part of the advice, is an instruction that to really understand your subject you need to learn how your subject functions, not just identifying what lies around them. We have pointed out that although biotinylation methods do a great job of informing on changing interactomes and cellular components, there is little information on what function these interactomes have, and, perhaps more importantly how these locale-specific functions could be perturbed and probed. Unfortunately, in many complex systems, identifying the specific roles of the (associating) protein(s) can be rate–limiting, and this key question is often addressed by approaches involving bulk/uncontrolled pharmacological treatments of the cells or organisms, inducing global changes to the proteome. We cede that gleaning such information is by no means simple, and affecting locale proteomes selectively is almost always not possible using current technologies. On the other hand, affecting proteins in a cell–type specific manner is possible, although this is not perhaps within the scope of most laboratories. Moreover, knockout of specific proteins does not necessarily inform on pharmacological interventions targeting those proteins.

We were struck by the fact that a relatively small subsection of bioactive small molecules are inherently reactive electrophiles,[1,32,33] that have defined but relatively broad and varied interactomes. Using the technologies we have developed (Figure 4A) that are applicable to cells,[34] fish,[35] and worms,[36] these reactive small molecules can be photocaged and released on demand with rapid kinetics,[37] similar to biotinylation methods. However, unlike the small–molecule species used in biotinylation experiments that are all non–drug–like in nature, and indiscriminately reactive, and often inherently unstable, reactive electrophiles are relatively stable,[38] often affect specific functions of the proteins they label,[39–46] and critically, can

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH

14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Chemistry Europe
European Chemical Societies Publishing

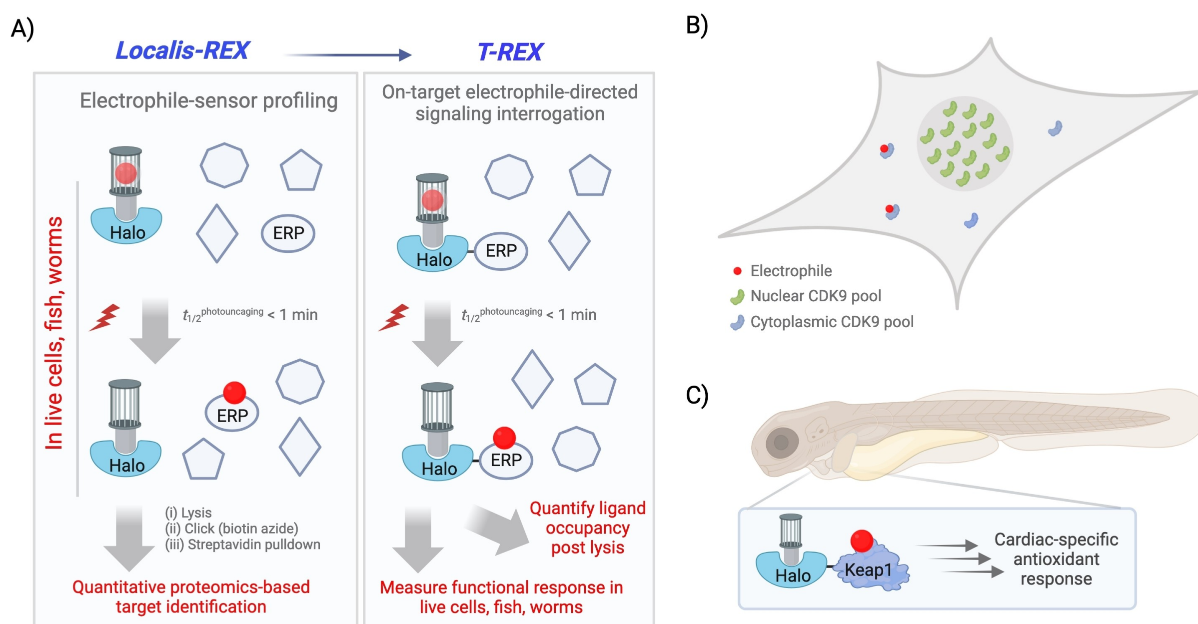

Figure 4. A) In the current setup of REX-technologies,[34,35,37] following administration of Halo–targetable photocaged electrophile probes (Ht–PEs) to live cells/animals, the cell–permeable Ht–PE binds Halo, or Halo fused to a protein of interest (POI), expressed in a specific locale. Expression can be restricted to a specific locale using a requisite localization sequence fused to Halo/Halo-POI, or to specific tissues by driving expression with native promoters in model organisms. Following washout of the excess unbound Ht–PE, only a 1:1 complex between Ht–PE and Halo/Halo-POI remains. Low–energy light illumination (0.5 mW/cm^2, 365 nm) at a preordained time enables rapid release ($t_{1/2} < 1$ min[45]) of the desired reactive electrophile (**red sphere**) within the proximity of Halo (in Localis-REX, left–panel) or Halo–POI (in T-REX, right–panel). In our recently–established function–guided target-ID tool, Localis-REX (left–panel),[39] electrophile molecules rapidly–liberated in limited dosage at a defined locale, are competitively 'intercepted' by the fast–responding native electrophile–responsive proteins (ERPs) within the vacinity of Halo. The non-invasive integration of an alkyne unit within the Ht–PE (that is carried away in the released electrophile) allows Click–chemistry–enabled biotin–pulldown–based target enrichment. In unison with state–of–the–art SILAC, TMT, and LFQ–based target-ID workflows, Localis–REX enables quantitative target-ID of ligand-responsive local proteome (left–panel). Target/ligand-pairs so-discovered are subsequently fed forward to precision functional validations, using T-REX (right–panel). T-REX stands to date as the only technology capable of interrogating protein–specific reactive–ligand–guided signaling responses in living systems. In T-REX, should POI be a privileged responder to the liberated electrophile, the POI can react prior to electrophile irreversibly diffusing away from the POI. Providing a given electrophile–ligand occupancy on the POI is sufficient to elicit a response, we have shown in numerous living models, how T-REX–and oftentimes used in tandem with Localis–REX–delivers a unique means to unambiguously map precision electrophile-signaling mechanisms; to score downstream on–target cellular responses using established quantitative cell-based readouts or organismal phenotypes; and to uncover novel pathway players and their mode of action. B) Applying Localis–REX as described in the text discovered cytosol–specific electrophile–responsive function of the kinase CDK9, despite the nucleus being CDK9's primary subcellular residence. Such an ability to seive out a functional property from minority pool of a given protein (here, minority cytoplasmic CDK9 pool) is not achievable using any other technologies. (note: Studies using T-REX in combination with other genetics and biochemical approaches show that the lack of electrophile-sensing action of nuclear CDK9 is due to its nuclear-spencific binder, cyclin T1, blocking the electrophile-sensing site within CDK9). C) Applying Z-REX (zebrafish version of T-REX) where Halo-POI is selectively expressed in the heart of transgenic fish larvae as described in the text, shows tissue–specific antioxidant response upregulation as a result of POI (Keap1)-specific electrophile labeling, in an otherwise largely unperturbed backdrop of $>200{,}000$ unique protein–cysteines in the heart proteome or that of other cells/tissues in the animal.

inform on drug design.[47,48] Indeed, many approved drugs now contain a reactive electrophilic Michael–acceptor–based appendage that is reminiscent of reactive electrophilic molecules.[49–51] In some ways, therefore, reactive electrophiles are also highly applicable to proximity-profiling experiments, but can afford more functional information than reactive biotin probes.

Unfortunately, there are several potential downsides to using reactive electrophiles for proximity profiling. One of these is that reactive electrophiles' diffusion distances in cells are highly context dependent, due to their intrinsic metabolic vulnerability to small–molecule nucleophiles such as GSH and also enzymatic degradation.[38] Thus, spatial resolution remains a large question mark when deploying electrophiles for proximity profiling. Nonetheless, we will discuss below how we have used electrophiles released in specific sub–locales to profile proteins that are reactive to specific electrophiles in defined locales. This result overall bodes well for future experiments probing reactivity in specific subcellular locales. The above limitation is much less pertinent to tissue-/organ–specific profiling experiments.

One other potential downside with using reactive electrophiles for proximity mapping is that they are inherently only reactive with a subset of proteins,[1] hence their local–proteome coverage is lower than that of reactive biotinylation strategies, which can label a large amount of local proteins. On the flip side, it is the discriminatory nature of reactive electrophiles' reactivity profiles that furnishes information on function, and can give information on subtle structural or associational changes, for instance, in response to a specific biological cue. One means to circumvent the issue of low coverage is to use multiple electrophiles bearing divergent reactive chemotypes.

Using photocaged electrophiles that are anchored to specific locales because they can bind irreversibly to HaloTag

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH
14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Perspective
doi.org/10.1002/cbic.202400005

protein that was expressed in either the cytosol or the nucleus, we recently profiled specific proteins that are electrophile responsive in those locales, harnessing Localis–REX (Figure 4A).[39] Validation showed that all top hit proteins sensed electrophiles best in the locales identified. One protein, CDK9, emerged to be a cytosol–specific electrophile sensor, even though it resides canonically in the nucleus (Figure 4B). Using CDK9-specific labeling (enabled by T-REX in live cells, Figure 4A) and bolus dosing with electrophiles, relative to electrophile–labeling–resistant but otherwise–kinase–active mutants, we showed that CDK9 labeling by the bioactive lipid–derived signaling electrophile, 4-hydroxynonenal, inhibits RNA-polymerase-II-mediated transcriptional initiation.[39] Thus, localized precision electrophile labeling techniques have the ability to resolve organelle–specific sensing and can inform on means to inhibit proteins in a locale–specific manner.

One useful aspect of photocaged electrophiles is that once an electrophile–sensor protein is identified, these sensor proteins can be fed directly into experiments to assay specific protein's function in specific locales or tissues. For instance, we have examined how tissue–specific electrophile modification of a key electrophile–sensor protein, Keap1, regulates antioxidant response in zebrafish heart. To enable this pursuit, we engineered transgenic (Tg) zebrafish where heart–specific expression of Halo–tagged Keap1 is driven by the endogenous heart-specific promoter, myl7, that naturally drives the expression of zebrafish heart-specific myosin light chain 7 protein (Figure 4C).[35] Keap1 is a principal negative regulator of the antioxidant response (AR).[52] Electrophile targeting of this protein inactivates AR inhibition, upregulating a battery of cytoprotective AR-driven genes. These Tg fish were treated with a biocompatible, non–toxic, and animal/cell–permeable photocaged precursor to a bioactive signaling electrophile. Following an embryo–growth period allowing for consequential washout of the excess non–Halo–bound photocaged probe, photouncaging was performed. Specifically in the heart, we observed a 2-fold increase in AR.[35] Importantly, when Halo–Keap1 was expressed in the majority of tissues using mRNA injection, we saw predominantly upregulation of AR in the tail fin, which was not observed when Halo–Keap1 was expressed in the heart. This result, together with other protein–specific examples of targeted electrophile delivery from our laboratory (such as delivery to Akt3,[47] Ube2V2[43]) indicates that locale–specific perturbation of proteins is indeed possible.

Conclusions: Everybody's Gotta Learn, Nobody is Born Knowin'

We end our piece with a quote that should remind us that, as biological scientists, we use empirical information to unmask a world that is uniquely organized, but derived from infinite possibilities. Indeed, we are those who have to peer into the unknown and learn what constitutes the fabric of ourselves. Tools to unmask this world are hugely important and can have enormous consequences for how we understand biological systems. Although it is hard to imagine now, it is indeed the ability to delve deeper into the inner workings of biological systems across a broad range of subfields that is common to a host of fundamental methods used in biology. Such methods include all forms of microscopy, genetics, and omics. Interactome profiling methods are now taking their place within this canon, offering insights that few other methods can deliver.

Looking at the above trends across the years, it is discernable that groundbreaking techniques are starting to become more interdisciplinary. Indeed, this trend is clear even when one considers developments that have occurred across new sub–disciplinary techniques, such as (electron)microscopy, and various omics techniques. In this way, proximity–labeling techniques represent a logical extension of previous trends in the life sciences. The fact that proximity–labeling approaches are chemical genetic in nature, has arguably aided their integration into the broader canon, as their deployment fits well into traditional biological experiments, as we have seen for generation of transgenic *C. elegans* and other whole–animal research models.

Nonetheless, as proximity–labeling approaches leverage aspects of several fields to fill gaps in our understanding, they potentially open new avenues for translational gain. In other words, we are no longer using light (as in microscopy) to help us peer into the void, or using mutants and genetic crosses to decode protein function and in which pathways they lie. We are now tapping into small–molecule chemistries, our most fertile source of biological tools and ultimately drugs, to do the same. We strongly believe that it is crucial that we take advantage of the new opportunities that such methods bring to feedforward to applied science. REX technologies,[34,35,37] and other similar incarnations have begun to tap into such aspects. However, a lot more work needs to be done both to evaluate drug–like molecules as functional proximity probes and in what areas/locales they can be most useful.

Of course, we can only speculate on these points as the field is only starting to move into these areas. However, given that many reactive molecules, including reactive drugs target off–active–site reactive residues, it seems less likely that *C. elegans* and simple model organisms would be particularly fruitful hunting grounds for functionally–manipulable proteins. Mice seem to be a logical option, as much of their physiology is conserved to man. However, mice do not fit in so perfectly with photouncaging techniques that are leveraged by many of the existing proximity electrophile–delivery techniques. Indeed, although optogenetics tools have been proven successful in mice, these remain complex and are limited to accessible regions, such as the brain. This is of course one area where reactive biotinylation–based proximity–profiling methods are particularly beneficial. In the case of photocaged reactive electrophiles, although there are many options, it would seem that zebrafish, as we discussed above, including potentially adult Casper zebrafish, are quite useful options. Alternatively, other methods of uncaging reactive electrophiles (admittedly that are not currently in use), such as enzymatic, or pH, could also be deployed to perform proximity labeling.

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH

14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14397633, 2024, 8, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cbic.202400005 by Bcu Lausanne, Wiley Online Library on [22/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Acknowledgements

EPFL and UNIL. Open Access funding provided by École Polytechnique Fédérale de Lausanne.

Conflict of Interests

There are no conflicts to declare.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Keywords: Proximity/interactome mapping tools · Function-guided spatial-proteomics · Electrophile signaling · REX-technologies · Bio-ID and variants · APEX · Protein-protein interactions

[1] S. Parvez, *Chem. Rev.* **2018**, *118*, 8798–8888.
[2] S. Sukenik, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6776–6781.
[3] R. Schairer, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 3093–3102.
[4] M. J. C. Long, *Macromolecular Protein Complexes IV: Structure and Function* (eds. J. R. Harris *et al.*) **2022**, 155–197 (Springer International Publishing, Cham; 2022).
[5] J. S. Lin, E. M. Lai, *Methods Mol. Biol.* **2017**, *1615*, 211–219.
[6] Y. Fu, *Nat. Chem. Biol.* **2018**, *14*, 943–954.
[7] P. J. Thul, *Science* **2017**, *356*.
[8] O. Söderberg, *Nat. Methods* **2006**, *3*, 995–1000.
[9] C. Grefen, *Curr Protoc Neurosci* **2007**, *5(5.27)*.
[10] C. Bignon, *Int. J. Mol. Sci.* **2022**, *23*.
[11] L. M. Orre, *Mol. Cell* **2019**, *73*, 166–182.e167.
[12] M. Ma, *Nat. Commun.* **2022**, *13*, 7736.
[13] W. Qin, *Nat. Methods* **2021**, *18*, 133–143.
[14] C. P. Seath, *Chem. Soc. Rev.* **2021**, *50*, 2911–2926.
[15] C. L. Chen, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 12093–12098.
[16] E. A. Veal, *Mol. Cell* **2007**, *26*, 1–14.
[17] F. Antunes, P. M. Brito, *Redox Biology* **2017**, *13*, 1–7.
[18] D. I. Kim, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2453–E2461.
[19] L. Kubitz, *Communications Biology* **2022**, *5*, 657.
[20] M. Artan, *J. Biol. Chem.* **2021**, *297*.
[21] Y. Liu, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2025299118.
[22] S. Y. Lee, *Nat. Methods* **2023**, *20*, 908–917.
[23] T. C. Branon, *Nat. Biotechnol.* **2018**, *36*, 880–887.
[24] E. Holzer, *PLoS Genet.* **2022**, *18*, e1010150.
[25] S. A. Shuster, *Neuron* **2022**, *110*, 3882–3896.e3889.
[26] D. Arora, *Plant Cell* **2020**, *32*, 3388–3407.
[27] A. D. Sanchez, *Curr. Biol.* **2021**, *31*, 3586–3600.e3511.
[28] M. Artan, *J. Biol. Chem.* **2021**, *297*, 101094.
[29] S. Rayaprolu, *Nat. Commun.* **2022**, *13*, 2927.
[30] K. E. Kim, *Nat. Commun.* **2021**, *12*, 5204.
[31] C. R. Kanzler, *Dev. Biol.* **2022**, *492*, 133–138.
[32] F. J. Schopfer, *Chem. Rev.* **2011**, *111*, 5997–6021.
[33] A. T. Jacobs, L. J. Marnett, *Acc. Chem. Res.* **2010**, *43*, 673–683.
[34] S. Parvez, *Nat. Protoc.* **2016**, *11*, 2328–2356.
[35] K. T. Huang, *Nat. Protoc.* **2023**, *18*, 1379–1415.
[36] M. J. C. Long, *Biochemistry* **2018**, *57*, 216–220.
[37] M. J. C Long *Acc. Chem. Res.* **2021**, *54*, 618–631.
[38] X. Liu, *Trends Biochem. Sci.* **2019**, *44*, 75–89.
[39] Y. Zhao, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2120687119.
[40] A. Van Hall-Beauvais, *eLife* **2022**, *11*, e83373.
[41] J. R. Poganik, *Nat. Commun.* **2021**, *12*, 5736.
[42] J. R. Poganik, *The FASEB Journal* **2019**, *33*, 14636–14652.
[43] Y. Zhao, *ACS Cent. Sci.* **2018**, *4*, 246–259.
[44] M. J. Long, *Cell Chem. Biol.* **2017**, *24*, 944–957.e947.
[45] H.-Y. Lin, *J. Am. Chem. Soc.* **2015**, *137*, 6232–6244.
[46] S. Parvez, *J. Am. Chem. Soc.* **2015**, *137*, 10–13.
[47] M. J. Long, *Nat. Chem. Biol.* **2017**, *13*, 333–338.
[48] X. Liu, *ACS Cent. Sci.* **2020**, *6*, 892–902.
[49] L. Boike, *Nat. Rev. Drug Discovery* **2022**, *21*, 881–898.
[50] M. J. C. Long, *Curr. Opin. Chem. Biol.* **2019**, *51*, 48–56.
[51] M. J. C. Long, Y. Aye, *Cell Chem. Biol.* **2017**, *24*, 787–800.
[52] A. Cuadrado, *Nat. Rev. Drug Discovery* **2019**, *18*, 295–317.

Manuscript received: January 2, 2024
Revised manuscript received: February 7, 2024
Version of record online: March 21, 2024

© 2024 The Authors. ChemBioChem published by Wiley-VCH GmbH