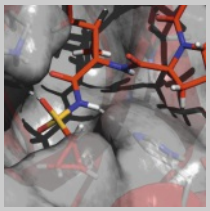
doi:10.2533/chimia.2022.598 Chimia 76 (2022) 598–600



Medicinal Chemistry and Chemical Biology Highlights

Division of Medicinal Chemistry and Chemical Biology
A Division of the Swiss Chemical Society

Bringing Functional Context to Emerging Proximity-mapping Proteomics Tools

Yimon Aye*

*Correspondence: Prof. Yimon Aye, (yimon.aye@epfl.ch), Swiss Federal Institute of Technology Lausanne (EPFL)

Abstract: If one considers chemical-biology toolsets that have had the greatest impact on numerous fields of life sciences over the most recent years, proximity-labeling tools, such as APEX, and Bio-ID arguably lead the way. This article reflects upon the current state-of-the-art and discusses key limitations underlying these emerging approaches, in particular, the limited functional knowledge they provide in understanding local proteomes / interactomes. This limitation is directly linked to the use of non-biologically- or non-pharmaceutically-relevant reactive intermediates in the course of covalently labeling the local proteomes. As such, these methods cannot report on specific functions of localized protein players, nor can they scrutinize whether the specific functions of such proteins/interactomes can be directly manipulated by pharmacologically-relevant small-molecule ligands. The latest data hint that precision localized electrophile delivery concept ushers a means to address this limitation with high spatiotemporal resolution, and ultimately, in relevant live animals.

The interests, and at times even the needs, to peer into proteins present in specific subcellular locales of a cell have continued to grow in many research areas across the gamut of life science fields. Thus, tools such as Bio-ID[1] and APEX[2], have rapidly gained acceptance and have been broadly deployed. Beyond addressing fundamental research questions, understanding local interactomes/behaviors at subcellular-, cell-type-, or context-specific levels can open a new lens into the development of targeted therapies with improved efficacies and potentially reduced off-target effects. On a broader level, emerging widespread applications of single-cell RNA-Seq, single-cell proteomics, and related spatial omics share a similar, if not entirely, common goal of searching for context-specific changes in local players. This opinion piece highlights the current state of proximity-mapping proteomics techniques. Along the way, it highlights critical technological and knowledge gaps, and identifies ways to facilitate the otherwise largely-unmet need to add new dimensions to these techniques, specifically focusing on functional aspects, *i.e,* locale-specific chemical actionability or druggability (Fig. 1). Since the biological applications of mapping techniques have been recently reviewed by us[3,4] and others[5-7], this discussion is limited to conceptual frameworks underlying existing tools and how they could be improved.

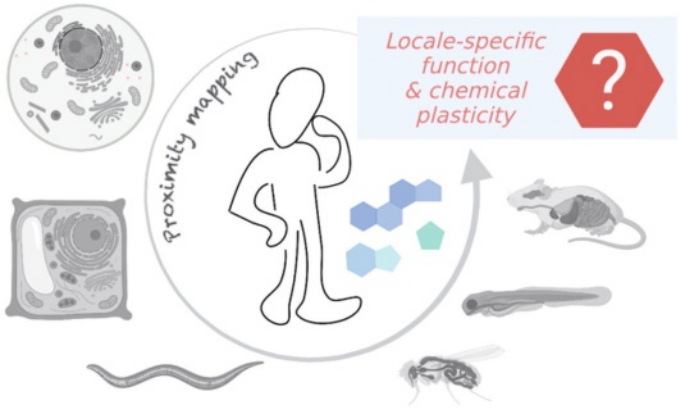


Fig. 1. Despite being a huge boon to numerous research fields, the emerging proximity-labeling tools are limited because they can only index the identity of local protein players / subcellular interactomes using non-biologically or non-pharmaceutically-relevant reactive small-molecule intermediates. This opinion piece highlights the criticality of understanding locale-specific functions (especially to gain an ability to selectively manipulate the local proteome at will with drug-like small-molecule ligands), and how chemical biology innovations could address these limitations, with the goal to ultimately achieve proximity mapping guided by designer small-molecule pharmacophores, and to be able to probe locale-specific actionability in mulitple relevant live research models, such as cultured cells, nematode worm, fruit fly, zebrafish, mouse.

Developing chemical biology tools to probe the make-up of local proteomes often requires a combination of genetic and chemical engineering as well as a means to enrich the local proteome against global noise for mass-spectrometry-based target identification. This necessity largely stems from the fact that entirely chemical methods to localize small molecules are insufficient relative to: (*i*) using a known native subcellular localization sequence that can be genetically encoded to a protein of interest (POI), allowing expression at a designated organelle or a subcellular locale, or (*ii*) using a native cell-type/tissue-specific promoter that can restrict the expression of a POI to defined tissues/cells in animals. Bio-ID[1,8] (and later variants such as Turbo-ID[9]) and APEX[2,10] achieve locale specificity of their corresponding engineered proteins, BirA* and related promiscuous biotin ligases, and the APEX-peroxidase, respectively. Thus, in the case of Bio-/Turbo-ID, following whole-cell treatment with Biotin, the activated ester, Biotin-AMP, is enzymatically generated, in the presence of cellular ATP and BirA*, only in the locale where BirA* resides. Proximal proteins bearing surface lysines (within a ~50 nm radius of the activating protein) are non-specifically biotinylated. Following cell lysis, an established streptavidin pulldown and digest mass-spectrometry permit the identity of local proteomes to be determined, against defined control samples/biological conditions. Incorporation of quantitative

proteomics methods such as SILAC[11] or TMT[12], further allows enriched protein targets to be quantitatively ranked, approaches common to all proximity mapping coupled proteomics-target-ID workflows. Turbo-ID/miniTurbo improve labeling efficiency of the original Bio-ID approach (by ~20-fold): the several hours or days of labeling needed with Bio-ID have been reduced to minutes using, for instance, miniTurbo.[9]

APEX[2,10] stands as a major breakthrough parallel to Bio-ID, and is among the most widely used tools by life scientists to ID the local proteome. When restricted to a specific subcellular locale, the engineered APEX peroxidase catalytically converts biotin phenol (fed from outside the cells) to the corresponding phenoxy radical in the presence of peroxide, which must be added, typically at cytotoxic concentrations, to activate the APEX enzyme. The phenoxy radical produced has a shorter intrinsic half-life[13] than that of Biotin-AMP, and hence the labeling radius attained via APEX is inherently smaller than that of Bio-ID. However, intrinsic cellular factors, such as enzymatic hydrolysis of Biotin-AMP, likely render this comparison context-dependent. As we discussed in an earlier perspective[14], high concentrations of peroxide required to kick-start APEX can negatively influence cell physiology, and inadvertent alterations in subcellular locales and trafficking *etc,* as a result of peroxide treatment, render targets captured under this condition difficult to parse from genuine hits. Indeed, part of the motivation behind the latest development of miniTurbo/Turbo-ID is to address this specific shortfall underpinning APEX, while also addressing the prolonged labeling period required in Bio-ID.[9] Likely also because of peroxide treatment, APEX-based proximity mapping is largely limited to cultured cells, or excised tissues, instead of real-time proximity mapping in intact live animals. Nonetheless, the capabilities of APEX to couple with transcriptomics tools as in APEX-Seq, have further ushered new opportunities to unearth the identity of locale-specific changes in different categories of subcellular RNAs.[15]

One of the most recent inventions in the field of proximity mapping uses the use of carbenes to label local proteomes. This method, termed μMap,[5,16] that can ID the local interactome at microscale, as a result of intrinsically short-lived (~ns) nature of carbenes. This results in a compact labeling radius of ~5 nm. In the presence of light and iridium-photocatalyst (that can be anchored to a specific intracellular protein tag, or an antibody at the cell surface), small-molecule ligands housing a diazirine unit (fed from outside the cell), can generate carbenes within the proximity of the protein tag (or the antibody). Indeed, it has become popular to compare diffusion distances of specific methods to understand their spatial resolution and hence rank each method. Although such comparisons are interesting, comparing diffusion distances and half-lives of reactive species is, as alluded to above, complex since these parameters can vary depending on specific biological microenvironments:[13] lipidation, pH, redox status, as well as the local proteome itself. Furthermore, although the shorter labeling radius may provide a more 'selective' and 'higher resolution' mapping, given the average diameter of proteins is ~5 nm, too short a radius may pose a limit, or be less relevant in certain contexts, for instance, in mapping large complexes or synapses.

Despite such revolutionary pioneering approaches to ID the local proteome, transcriptome, and/or interactomes discussed above and other closely-related strategies reviewed elsewhere,[3,5-7] all of these approaches are only able to inform on the identity of local players. It would be transformative to gain the ability to directly *and simultaneously* discern functional aspects, while probing for novel localized players. This posit goes with the appreciation that functional information is what is crucial to understand and *control* cell responses, behaviors, and ultimate decision making. Functional knowledge thus proffers golden opportunities to manipulate the local proteome for targeted therapies. Indeed, target-ID and functional validations long continue to be segregated, in the context of drug discovery and functional biological experiments. It is of the author's opinion that an all-in-one approach can be offered if proximity mapping tools can be (re)devised to enable mapping guided by function.

Our laboratory has recently shown a generalizable means toward integrating the aspects of 'chemical actionability' into proximity mapping and spatial omics technologies. In a method termed "Localis-REX",[17] we adapted our original REX technologies[18,19] and leveraged the underlying localized electrophile delivery concept,[14,20] to enable controllable delivery of a defined dosage of chemically-defined natural electrophiles,[13] to specific subcellular locales with precise time and spatial resolution. The approach allows quantitative capture of the electrophile-actionable local protein players. Briefly, the Halo-protein tag that can be expressed in a specific locale in a cell, or a specific cell/tissue in an animal, covalently binds the Halo-recognition ligand appended to the designer photocaged electrophile probe. The latter cell-permeable bifunctional probe is administered from outside the cell/animal.[18,21-23] Following washout to remove excess non-Halo-bound probe, light is used to unleash the reactive electrophile within the proximity of Halo, in an amount maximally stoichiometric to intracellular concentration of Halo. Under this regimen, targets uncovered from localis-REX were unaffected by local protein abundance.[17,24] indicating that the method effectively dampens global proteome 'noise' to allow pinpointing of chemotype-specific functional responders. Indeed, natural and preternatural electrophilic fragments that REX methods can deliver are validated pharmacophores in emerging precision therapeutics, and play a role in many of the emerging fragment-based drug discovery and medicinal chemistry programs.[20,25] Thus, uncovering druggability/chemical responsivity of the local proteome to individual electrophilic chemotypes within such native signals, provides a means to directly inform on novel target/ligand pairs of medicinal relevance, and furthermore the responsivities so uncovered are also likely Nature's 'pre-validated' regulatory mechanisms.

Indeed, we have shown that target/ligand pairs uncovered from precision electrophile delivery concepts and harnessing nature's electrophiles, can directly guide precision medicine development. For instance, we have established drug-like small-molecule lead compounds that can specifically target endogenous disease-relevant proteins,[21,25,26] bypassing the need for Halotag, light, and photocage probes (prerequisite of REX methods[20]). Critically, mechanism-of-action of lead molecules so developed[26] directly recapitulate phenotypes induced by electrophile-specific protein engagement.[21] Such levels of mechanistic insights are further delivered by the parallel application of T-REX method[18] to targets identified from Localis-REX,[17] allowing functional consequences of specific electrophilic ligand/target engagement in a cellular context to be deciphered. Because the REX methods have been generally adapted to suit applications in different whole organisms,[20-23] this versatility can be further harnessed in Localis-REX, to not only interrogate but chemically control local behaviors following nature's electrophile regulation that is likely conserved across evolution and tuned to a specific pathophysiological context.

Thus, combination of Localis-REX (function-guided proximity mapping)[17] and T-REX (precision signaling interrogations)[18,21,22] offers a rigorous approach to inform on functional plasticity of the local proteome in live cells and animals, hijacking naturally-occurring electrophile-regulatory circuits/networks/interactomes.[13] The current limitations of Localis-REX lie in the fact that unlike APEX, Bio-ID/Turbo-ID, μMap and other related methods, the electrophile dosage is limited in Localis-REX, which likely aids identification of "the best" sensors, but could also restrict the depth of electrophile-responder proteins one can unearth at a specific locale. We are actively addressing this limita-

tion while striving to expand the tools to suit a wider spectrum of model organisms (Fig. 1) and a broader chemical space/chemotype, beyond the current toolset representing cultured bacterial and mammalian cells, fish, and worms, and Michael-acceptor-based natural electrophiles.

In sum, the most recent years have witnessed literature explosion regarding the growing inventory of local protein players, transcripts, and their context-specific changes. These discoveries were directly enabled by the remarkable advances in proximity-labeling tools, the innovations of which have also underlined the power of interdisciplinarity in scientific inventions. Indeed, assimilation of fundamental, and often orthogonal, concepts underpinning organic synthetic methodology, photochemistry, genetics, and bioengineering, has gone into the successful design and evolution of these tools. Going forward, new capabilities to execute proximity mapping with customizable biologically-relevant ligands are much in demand. Such an ability is anticipated to unveil spatially-coordinated functional cell signaling behaviors with precise spatial and temporal control, and importantly, shed a direct light to better manipulate the local interactome across wide-ranging medicinal and therapeutic research programs globally. We thus hope that as the proximity mapping tools mature, the coming years will see either revamping of the existing growing toolbox or design of new creative biocompatible methods that will collectively help us get closer to simultaneous mapping of locale-specific functions beyond the identification of locale proteome / interactome.

Received: May 7, 2022

[1] Roux, K.J, Kim, D.I, Raida, M. & Burke, B., *J. Cell. Biol.* **2012**, *196*, 801, https://doi.org/10.1083/jcb.201112098.
[2] Rhee, H.W. Zou, P. Udeshi, N.D. Martell, J.D. Mootha, V.K. Carr, S.A. & Ting, A.Y., *Scienc*e **2013**, *339*, 1328, https://doi.org/10.1126/science.1230593.
[3] Long, M.J.C, Zhao, Y. & Aye, Y., *RSC Chem. Biol.* **2020,** *1*, 42, https://doi.org/10.1039/D0CB00041H.
[4] Long, M.J.C, Miranda Herrera, P.A. & Aye, Y., *Chem. Res. Toxicol.* **2022,** https://doi.org/10.1021/acs.chemrestox.2c00006.
[5] Seath, C.P, Trowbridge, A.D, Muir, T.W. & MacMillan, D.W.C., *Chem. Soc. Rev.* **2021**, *50*, 2911, https://doi.org/10.1039/D0CS01366H.
[6] Qin, W, Cho, K.F, Cavanagh, P.E. & Ting, A.Y., *Nat. Methods,* **2021**, *18*, 133, https://doi.org/10.1038/s41592-020-01010-5.
[7] Scorrano, L, De Matteis, M.A, Emr, S, Giordano, F, Hajnóczky, G, Kornmann, B, Lackner, L.L, Levine, T.P, Pellegrini, L, Reinisch, K, Rizzuto, R, Simmen, T, Stenmark, H, Ungermann, C. & Schuldiner, M., *Nat. Commun*. **2019**, *10*, 1287, https://doi.org/10.1038/s41467-019-09253-3.
[8] Kim, D.I, Jensen, S.C, Noble, K.A, Kc, B, Roux, K.H, Motamedchaboki, K. & Roux, K.J.. *Mol. Biol. Cell*. **2016,** *27*, 1188, https://doi.org/10.1091/mbc.E15-12-0844.
[9] Branon, T.C, Bosch, J.A, Sanchez, A.D, Udeshi, N.D, Svinkina, T, Carr, S.A, Feldman, J.L, Perrimon, N. & Ting, A.Y., *Nat. Biotechnol.* **2018**, *36*, 880, https://doi.org/10.1038/nbt.4201.
[10] Lam, S.S, Martell, J.D, Kamer, K.J, Deerinck, T.J, Ellisman, M.H, Mootha, V.K. & Ting, A.Y.. *Nat. Methods,* **2015**, *12*, 51, https://doi.org/10.1038/nmeth.3179.
[11] Mann, M., *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 952, https://doi.org/10.1038/nrm2067.
[12] Li, J, Van Vranken, J.G, Pontano Vaites, L, Schweppe, D.K, Huttlin, E.L, Etienne, C, Nandhikonda, P, Viner, R, Robitaille, A.M, Thompson, A.H, Kuhn, K, Pike, I, Bomgarden, R.D, Rogers, J.C, Gygi, S.P. & Paulo, J.A., *Nat. Methods* **2020**, *17*, 399, https://doi.org/10.1038/s41592-020-0781-4.
[13] Parvez, S, Long, M.J.C, Poganik, J.R. & Aye, Y., *Chem. Rev.* **2018**, *118*, 8798, https://doi.org/10.1021/acs.chemrev.7b00698.
[14] Long, M.J, Poganik, J.R. & Aye, Y.*, J. Am. Chem. Soc.* **2016**, *138*, 3610, https://doi.org/10.1021/jacs.5b12608.
[15] Padrón, A, Iwasaki, S. & Ingolia, N.T., *Mol. Cell*. **2019**, *75*, 875, https://doi.org/10.1016/j.molcel.2019.07.030.
[16] Geri, J.B, Oakley, J.V, Reyes-Robles, T, Wang, T, McCarver, S.J, White, C.H, Rodriguez-Rivera, F.P, Parker, D.L, Jr, Hett, E.C, Fadeyi, O.O, Oslund, R.C. & MacMillan, D.W.C., *Science* **2020**, *367*, 1091, https://doi.org/10.1126/science.aay4106.
[17] Zhao, Y, Herrera, P.A.M, Chang, D, Hamelin, R, Long, M.J.C. & Aye, Y., *Proc. Natl. Acad. Sci.* **2022**, *119*, e2120687119, https://doi.org/10.1073/pnas.2120687119.
[18] Parvez, S, Long, M.J, Lin, H.Y, Zhao, Y, Haegele, J.A, Pham, V.N, Lee, D.K. & Aye, *Nat. Protoc.* **2016**, *11*, 2328, https://doi.org/10.1038/nprot.2016.114.
[19] Poganik, J.R, Long, M.J.C. & Aye, Y., *Trends Biochem. Sci.* **2019***, 44*, 380, https://doi.org/10.1016/j.tibs.2019.01.006.
[20] Long, M.J.C, Rogg, C. & Aye, Y., *Acc. Chem. Res.* **2021**, *54*, 618, https://doi.org/10.1021/acs.accounts.0c00537.
[21] Long, M.J, Parvez, S, Zhao, Y, Surya, S.L, Wang, Y, Zhang, S. & Aye, Y. . *Nat. Chem. Biol.* **2017**, *13*, 333, https://doi.org/10.1038/nchembio.2284.
[22] Poganik, J.R, Huang, K.T, Parvez, S, Zhao, Y, Raja, S, Long, M.J.C. & Aye, Y., *Nat. Commun.* **2021**, *12*, 5736, https://doi.org/10.1038/s41467-021-25466-x.
[23] Long, M.J.C, Urul, D.A, Chawla, S, Lin, H.-Y, Zhao, Y, Haegele, J.A, Wang, Y. & Aye, Y. *Biochemistry* **2018**, *57*, 216-220, https://doi.org/10.1021/acs.biochem.7b00642.
[24] Liu, X, Long, M.J.C. & Aye, Y. Proteomics and Beyond: Cell Decision-Making Shaped by Reactive Electrophiles. *Trends Biochem. Sci.* **2019***, 44*, 75, https://doi.org/10.1016/j.tibs.2018.09.014.
[25] Long, M.J.C. & Aye, Y., *Cell. Chem. Biol.* **2017***, 24*, 787, https://doi.org/10.1016/j.chembiol.2017.05.023.
[26] Liu, X, Long, M.J.C, Hopkins, B.D, Luo, C, Wang, L. & Aye, Y., *ACS Cent. Sci.* **2020**, *6*, 892, https://doi.org/10.1021/acscentsci.9b00893.