Where Electrophile Signaling and Covalent Ligand–Target Mining Converge

Yimon Aye*

Abstract: Interests in learning how to engineer most effective covalent ligands, identify novel functional targets, and define precise mechanism-of-action are rapidly growing in both academia and pharmaceutical industries. We here illuminate the establishment of a multifunctional platform that offers new capabilities to logically engineer covalent ligands and dissect 'on-target' bioactivity with precise biological context and precision hitherto inaccessible. Broadly **aimed at non-specialist readers**, this opinion piece is aimed to stoke the interest of emerging chemists and biologists/bioengineers, but the underlying technological and conceptual topicality is anticipated to **also appeal to experts** leading ligand–target mining, validation, and -discovery research programs.

Keywords: Covalent ligands · Electrophile signalling · Ligand–Target mining

Yimon Aye read chemistry at University of Oxford (UK) (2000–04), and received her PhD degree in organic chemistry with David Evans (Harvard, USA) and postdoctoral training in biochemistry with JoAnne Stubbe (MIT, USA), where she established the mode-of-action of therapeutics targeting ribonucleotide reductase. In her independent career that began mid-2012, she set out to understand the detailed mechanisms of electrophile signaling. This impetus culminated in development of REX technologies (T-REX™ delivery and G-REX™ profiling). In a parallel research program, she studies genome-regulatory pathways, including mechanisms of anticancer drugs. She currently leads the Laboratory of Electrophiles and Genome Operation (LEAGO *https://leago.epfl.ch/*).

Cysteome at the Drug Discovery Frontiers

The renaissance of covalent drugs over the recent years[1–5] has kindled renewed excitement in mining the cellular cysteome – a catch-all term for the 250,000+ unique protein cysteines[4] present in the human genome. Many of these individual protein cysteines could be harnessable for novel target discovery and development of covalent therapeutics. Covalent adduction of protein cysteines manifesting soft nucleophilicity typically occurs *via* conjugate addition to the soft β-carbon of α,β-unsaturated carbonyl-derived appendages.[4] These so-called 'Michael acceptor' electrophilic motifs are present in many approved and emerging covalent drugs. Indeed, modern cysteome-targeting covalent drugs can be classified into two groups, all of which typically harbor Michael acceptors: *Group 1* comprises compounds wherein core target-binding-affinity-attributing components bearing relatively intricate architectures targeting specific proteins, are fused to Michael-acceptor units primed for covalent cysteine-tagging. Examples include pioneering synthetic kinase drugs, such as afatinib and osimertinib, and also compounds inspired by natural products to which Michael acceptors have been embedded, such as bardoxolone (Fig. 1a). *Group 2* constitutes emerging blockbuster compounds with little or no structural complexity that likely do not boast appreciative binding affinity to their targets, but show heightened electrophilicity with respect to *Group 1*. Tecfidera, Vumerity,[6] and plant-derived compounds such as curcumin and sulphoraphane belong to Group 2 (Fig. 1a). Generally, from reactivity–selectivity principles, Group 2 compounds of magnified electrophilicity are anticipated to cover a significantly broader on-/off-target spectra.

When Natural and Unnatural Worlds Converge

Interestingly, Nature has endowed us with a repertoire of electrophilic small-molecule metabolites having similar properties to Group 2 molecules. These include lipid-derived electrophiles (LDEs)[7–9] and (by)products of essential metabolic processes such as fumarate and itaconate, that altogether represent endogenously-produced Michael-acceptor-based reactive electrophiles native to humans (Fig. 1b).[9] Interest in such reactive metabolites has intensified over the recent years, with many reports documenting novel signaling pathways and targets, or pathway intersections, regulated by reactive metabolites.[9–15] When my independent laboratory began mid-2012, I was struck by the fact that levels of endogenous electrophiles are dependent on multifactorial cellular processes, such as metabolism, stress, development, and aging, and thus LDEs interact in a context-specific manner with endogenous proteins.[9] I proposed that such a situation may have evolved to an extent where context-specific engagement of an LDE with specific protein-cysteines, among the 250,000 available,[4] has become integral to cell-decision making.[16] If this were the case, a means to comprehensively decode such endogenous electrophilic-metabolite-directed regulatory patterns would be a gateway to understanding nuanced mechanisms of cell signaling. Furthermore, the similarity between the Michael-acceptor units within LDEs and many successful modern covalent drugs, renders deciphering natural electrophile signaling codes an untapped avenue toward leveraging nature in the design of novel covalent pharmacophores and novel target discovery.[17]

Druggability and Coverage: Target ID Without Biological Sufficiency in Drug Response?

Given the number of billion-dollar drugs, one could question why bother pursuing more small-molecule-based therapies? Such a question may seem particularly relevant today where biologics such as antibody-conjugates and engineered cells are gaining traction. However, historically and even up to the present day, small-

Correspondence: Prof. Y. Aye, E-mail: yimon.aye@epfl.ch
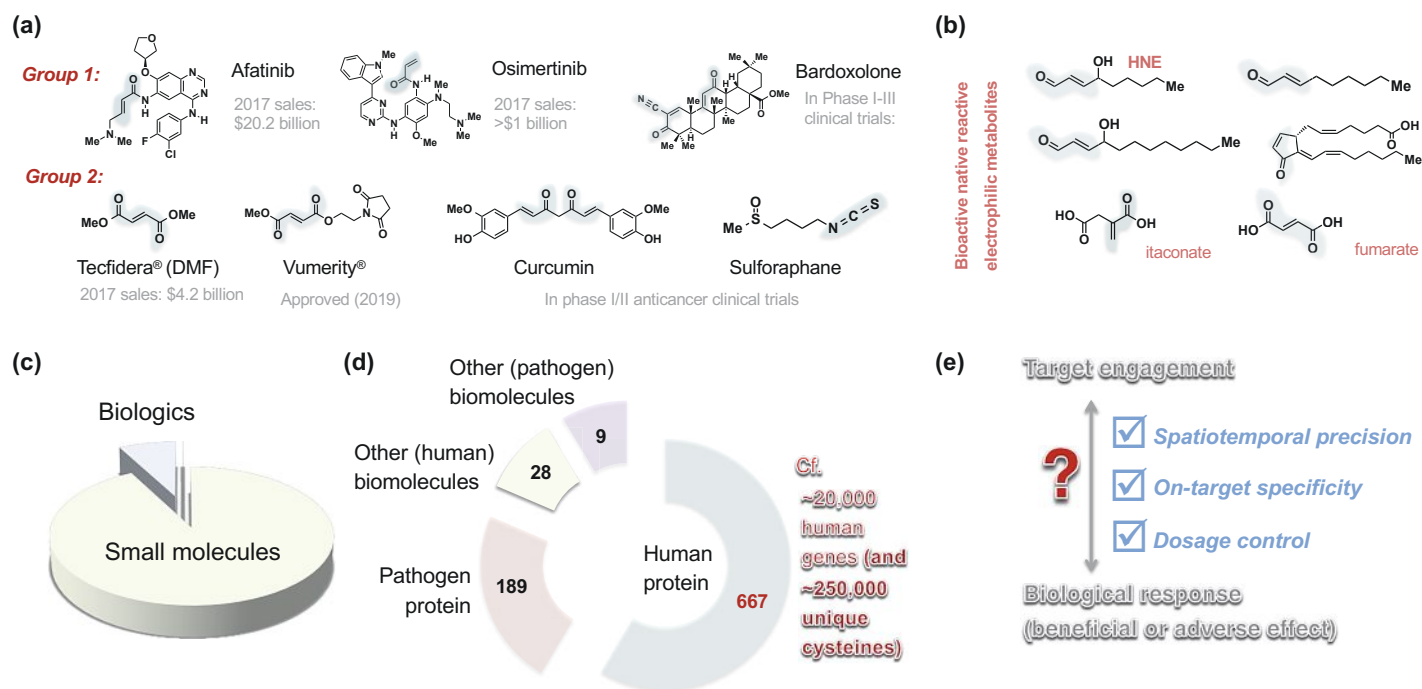Swiss Federal Institute of Technology Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Fig. 1. Challenges in modern drug discovery and the unique roles played by REX technologies. (a) Emerging covalent pharmaceuticals in clinical use and trials. See text for Group 1 and 2 classification. Afatinib and osimertinib are kinase inhibitors used in the treatment of non-small cell lung carcinomas, and ongoing trials of bardoxolone methyl assess treatment of kidney disease, hypertension, and cancer.[3,4,9] Tecfidera and vumerity are both approved to treat relapsing multiple sclerosis, but with poorly-understood mode(s) of action.[6] Plant-derived natural reactive electrophile curcumin and sulforaphane are in various stages of anticancer trials.[9] (b) Endogenously-produced reactive electrophiles that regulate human pathophysiology in a context-specific manner (HNE, 4-hydroxynonenal).[7–9] (a,b) All molecules bear electrophilic Michael-acceptor motifs shaded in blue.[4] (c) To date, small-molecule-based therapies are still a major contributor to the total number of FDA-approved drugs, compared to biologics.[18] (d) Absolute number of targets of FDA-approved drugs categorized by target class[18] shows the relevance of human proteins as molecular targets among all approved therapies but the number of druggable human proteins is <4% of ~20,000 genes in human, and is miniscule considering ~250,000 unique cysteines in the human proteome[4] that could be potential molecular targets of covalent medicine. (e) REX technologies[33] bridge the missing knowledge gaps between target-binding/target-ID analysis and functional outputs, by their ability to simultaneously profile target-ID and precisely assay for ligand-binding-induced responses with biological context and resolution hitherto inaccessible. See also Fig. 3 and 4.

molecule-based therapies continue to make up the lion's share of drug discovery (Fig. 1c). Small molecules also can reach targets often inaccessible to biologics, and hence they cover a wide range of human diseases. Critically, our ability to mine druggable human proteins falls desperately short of the demand: <700 out of the 20,000 human genes are currently targeted by approved drugs;[18] in other words, less than 4% of human proteins have been harnessed (Fig. 1d), even though a large proportion of proteins have drug-like properties. *Thus, effective mining of functionally-useful and therapeutically-profitable targets remains a grand challenge.* Our continued inability to surmount this challenge is tied to technological limitations, both in terms of profiling targets of small molecules but also importantly in functionally identifying the precise consequences of target engagement of individual ligand–target pairs in a biologically-relevant context. These limitations are magnified for broad-specificity electrophiles and covalent drugs, where we have been – at least to some extent – misguided by the concept that target identification (ID) is 'simple'. Existing state-of-the-art proteomics profiling are typically indirect and these approaches only examine at best <~6,000 targets, representing <~3% coverage of the whole human cysteome.[4] Furthermore, many of the targets identified are not strongly linked to phenotypic changes induced by the small molecules these methods purport to investigate. Thus, we are yet to even scratch the surface of the trove of drug-relevant applications of the cysteome. Many of these issues would be remedied by a general acknowledgement that ligand binding does not ensure a biological response (desired or undesired), and further embracing the logical conclusion that it is crucial to ensure the focus be on profiling 'functionality', beyond target ID.

In mid-2012, our laboratory – which has devoted its time roughly equally to the study of electrophile signaling[9] and the study of DNA damage response and genome maintenance centered around the enzyme ribonucleotide reductase[19–21] (Fig. 2) – launched a method-driven research program, more specifically, a means through which we can begin to directly link individual profiled targets of specific LDEs to specific biological outcomes triggered upon the target-specific engagement with a specific covalent ligand. We were acutely aware of the need to forge this missing link in the context of living systems. We reasoned that such a technology would be uniquely enabling to gain insight into electrophile signaling. This logic could also be extended to covalent drugs/drug-candidates, and unsolved challenges in establishing the *biologically-sufficient* mechanisms-of-action of the broad class of Group 2 compounds, including approved blockbuster drugs like Tecfidera.[6] The mystery shrouding such drugs continues to obfuscate target/ligand-optimization, defining contraindications, and management of clinical side effects, *e.g.*, idiosyncratic drug-induced liver injuries.

Challenges in Assessing Modern Reactive Covalent Drugs

Reactivity, Promiscuity, and Considerations of Permeation, Distribution, and Metabolism

At the outset of our program, all strategies to profile covalent-ligand-responsive targets, or studying the biological actions of reactive electrophiles, involved application of reactive electrophiles from outside of cells/animals, typically in large excess quantities (high μM to mM) and/or over prolonged time (hrs). Such uncontrolled bolus procedures have afforded increased insight into the
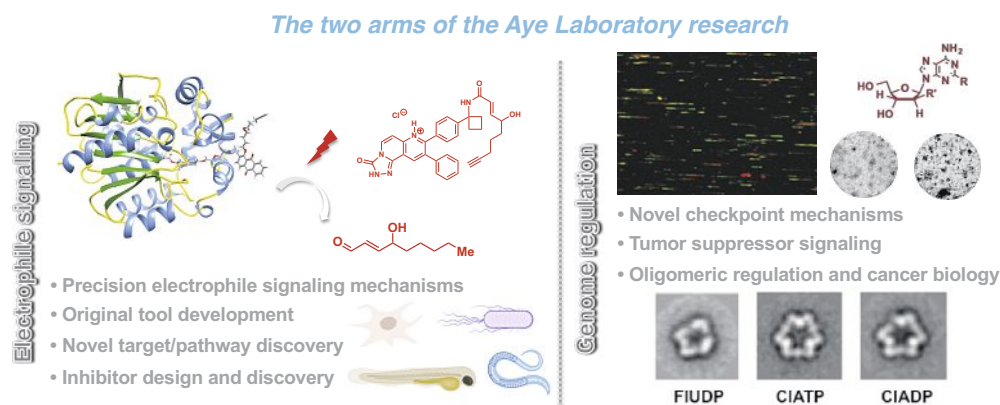
Fig. 2. Parallel research programs in the author's laboratory (established mid-2012). This article highlights the Aye laboratory research on electrophile signaling and stress response (left panel), wherein precision chemical biology tools are engineered in house aimed at addressing unmet biological questions of significant pharmaceutical relevance.[4,6,9,16,17,22–33,35–44] We apply these tools in combination with organic chemistry, biochemistry, and genetic principles in relevant *in vivo* models illustrated. The figure depicts general concept underlying REX-technologies: Halo-tag-protein covalently bound to photocaged-HNE enables light-driven on-demand delivery of covalent small-molecule such as HNE, Fig. 1b, *in vivo*, with precise timing and spatial and dosage controls (see also Fig. 3 and 4). The molecule on top-right shows a representative isoform-specific kinase inhibitor derived from application of REX-technologies (see also Fig. 5). The Aye laboratory also has longstanding interest and expertise in investigating the regulatory mechanisms of a validated drug-target enzyme ribonucleotide reductase (RNR) (right panel). Our work in this arena[19–21] includes discoveries of novel RNR-dependent pathways of importance in genome maintenance,[19,20] and molecular mechanisms of approved anti-cancer therapeutics targeting RNR.[21] Top left image shows dual-color DNA-fiber combing assay analyzing the effect of RNR on the rate of DNA-synthesis; top right image shows a generic adenosine-derived structure of RNR-targeting approved antileukemic nucleoside prodrugs, and the two circular images show results from focus formation assays analyzing effects of RNR (and its subcellular-localized variants) on tumor transformation. Negative-stained EM images show RNR-α-subunit hexamers formed as a result of binding of the nucleoside drugs, such as diphosphate of fludarabine (FlUDP) and tri-/diphosphates of cladribine (ClAT/DP).[19–21] Figures pertaining to ribbon structure, dual-color DNA-fiber combing image, images from focus formation assay, and EM electron micrographs, are adapted from the author's original research publications cited above, and the following publishers are acknowledged: American Chemical Society Elsevier, and Springer Nature.

pathophysiological effects of electrophilic species, and in several cases, led to identification of potential functional targets/pathways. However, bolus approaches do not function *in an otherwise largely-unperturbed backdrop*, hence the precise consequences of modifying an individual protein with a specific reactive ligand remain obscured. For a reactive ligand such as a native metabolite, 4-hydroxynonenal (HNE), a well-known LDE in humans,[7–9] or a reactive covalent drug Tecfidera (DMF),[6] even a brief whole-cell exposure in sub-nM quantities results in labeling/modification of thousands of proteins.[22] Furthermore, because any two small molecules could possess different levels of permeation and intracellular/intra-/inter-organellar distribution, metabolic lability, or undergo different metabolic processing, we are indeed blinded which protein target is exposed to a given electrophile and when. We are also unaware whether the chemotype of the drug is still the same at the point of target engagement from what it was at the point of cell/animal administration.[23] One may simply submit that cacophony that occurs during bolus processing is what happens when a medication is administered. However, without basic understanding of precise target engagement, not well-informed and ultimately costly decisions can be made. It also transpires that target engagement, and potentially locale specificity and other parameters not controlled by bolus dosing,[24,25] can indeed be controlled through '*drug hybridization*'[26,27] as I will discuss below.

Could Traditional Genetic Approaches Intervene?

Treating live specimens with reactive electrophiles also often leads to acute cellular stress, and if left unchecked, cell death.[9] Even with modest treatment time/dose, perturbation of redox environments and subcellular redox homeostasis occurs in ways that are not easily accountable or predictable. Furthermore, because the interaction between an electrophile and a target is typically irreversible, off-target labeling and its ramifications accrue in a time-dependent manner,[23] giving rise to muddied outputs that render precise outcomes hard to replicate and may well not reflect the conditions target cells experience in whole organisms. Notably, the largely irreversible engagement between a reactive

electrophilic metabolite, or a covalent drug, and a protein is typically not assisted by an enzyme, in stark contrast to how classical post-translational modifications (PTMs), such as phosphorylation, choreograph cell signal rewiring.[28] Thus, genetic knockdown/knockout of PTM-mediator-enzymes – such as knocking down a kinase in dissecting phosphosignaling dynamics – is not available to those studying electrophile signaling. Genetic knockdown/knockouts of the postulated modified protein also do not shut off the accumulation of off-target effects occurring during bolus dosing. Indeed, as many proteins are intrinsically sensitive to a specific electrophile, several proteins can also be responsible for the observed response. These factors hugely confound targets captured following bolus electrophile treatment (both in terms of masking key on-target players and highlighting proteins whose kinetics of labeling are inadequate), and complicate assignments of phenotypes. Indeed, it is quite commonly seen in the literature where top-enriched protein-targets from bolus treatment-based proteomics target-profiling experiments fail to account for the observed biological responses; namely, upon targeted knockdown/knockout of that protein, the same responses are largely maintained following drug treatment in the knockdown/knockout background as in wild-type-protein background.[6,23] Of course, knocking down the postulated electrophile-modified protein does not just simply remove the modification but it also affects the basal functional networks and interactome of that protein. What we thus needed is a system that would cleanly modify a protein with an electrophile on demand, allowing us to precisely monitor the consequences of such a modification event, in the backdrop of a largely-unperturbed cell.

Considerations of Conventional Interpretations and Analyses

Gaining an atomic resolution of ligand-binding sites is extremely useful for drug optimization. The field has historically relied hugely on 3D structures of ligand-bound proteins and more recently mass-spectrometric-based identification to glean such insights. Both these assays can also assign residue-specificity

of labeling for covalent binding. However, both methods suffer from artifacts, which are magnified for highly-reactive promiscuous ligands. For instance, taking the LDE, HNE, as an example, this molecule has multiple reactive motifs.[9] Such molecules in isolated systems – often deployed to study binding by the above methods – will label many exposed cysteines, and, with time, other nucleophilic residues, *e.g.,* histidine, and lysine, stochastically and substoichiometrically,[9] rendering data acquired hard to rationalize. Direct mass-spectrometry measurements of labile thiol modifications are further complicated because relative, and especially quantitative, assessments are limited since intrinsic properties, such as ionization and stability, often dominate. Modern mass-spectrometry profiling methods have sought to address these issues, but success has been achieved at the expense of limited proteome coverage, or other assay limitations we have discussed at length elsewhere.[9,23]

In terms of validation, point mutagenesis of a specific residue postulated to be modified – provided mutant protein is biologically functional – is a valid strategy to interrogate the precise role of covalent molecule sensing. However, some proteins with multiple functionally-responsive cysteines are not amenable to such validations, as many proteins with cysteines eventually react with the electrophile available in excess under bolus conditions.[29,30] This method also cannot account for multiple redundant sensor-residues when experiments are performed in living systems. Last but not least, many studies also try to identify sensitivity or site-specificity using purified proteins in isolated systems.[9] However, it remains excessively challenging to recapitulate the subcellular micro-environment(s) (and associated redox status) where a protein resides, or even the seemingly-simpler parameters such as protein's concentration, conformation, which altogether in turn affect propensity of a cysteine's responsivity to an electrophile.[9] These issues ultimately confound the interpretations derived from exclusively in vitro measurements of electrophile/covalent-ligand labeling. pK_a and structure-based predictions of surface-exposed cysteine

also arouse concerns/limitations. In fact, *a quick evidence-based calculation already indicates pK_a may not be the over-riding factor determining a given cysteine's reactivity.*[9,22]

Engineering T-REX: A Tool that Can Simultaneously Perturb and Validate Precision Target Engagement *in vivo*

Our laboratory conceived a chemical-biology concept which we later termed 'breaking the tether'.[31] This concept was at the time unheard off and carried a lot of risk, especially for a starting laboratory. Chemical biologists had invested a large amount of effort developing linkers and scaffolds to shepherd a specific unnatural entity to a protein of interest (POI), or POI fused to an engineered protein/peptide-tag.[31] But we wanted to design a system where we would undo all this work at a specified time. Furthermore, for this model to work, the reacting cysteine from a POI would need to have partially or fully grabbed the covalent small-molecule signal liberated by breaking of the tether,[31] in the split second *prior to diffusion away from the POI's coordination shell* (Fig. 3).

We proposed that when a reactive molecule encounters a protein that is tuned to react with it, reaction would be so fast,[32] the molecule (or a large proportion of it) would not be able to get away from the protein before reaction occurred.[16] Imagining this in a series of snap-shots, we would begin with a reactive metabolite becoming available within the POI's microenvironment, directly forming an encounter complex; within this ephemeral complex, the ligand would undergo partitioning between irreversible labeling of the POI and irreversible diffusion away from the POI, the result reflecting a ratio of reactivity and collapse of the encounter complex.[16] Thus, we designed the system – now known as REX technologies[33] – to enact liberation of (1) specific dosage and (2) at specific time, of a predetermined electrophile juxtaposed to the POI. Of course, we wanted to be able to functionally profile the precise consequences of ligand–target
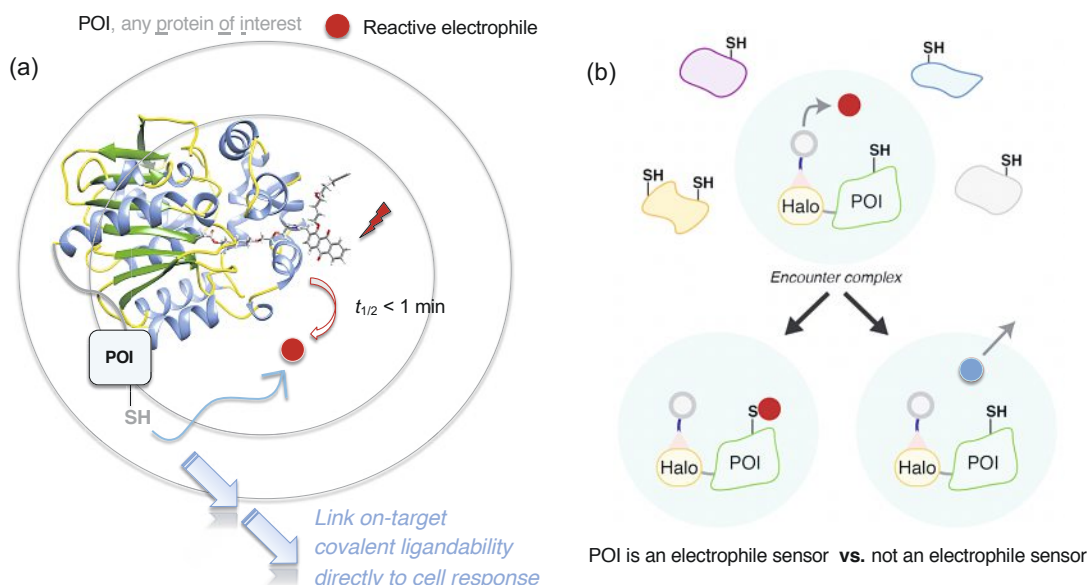

Fig. 3. T-REX probes precision on-target POI-electrophile engagement *in vivo*.[36,38,43] (a) Model of Halo-POI fusion protein wherein Halo (ribbon structure) is covalently bound to anthraquinone-based photocaged probe precursor to a representative native electrophile HNE (see Fig. 1b). Provided standard suite of technical controls are implemented, T-REX can quantitatively assay, for any POI in living systems, at precise space and time: (1) to what extent the POI is sensitive to a given electrophile; (2) if the POI is labeled by the electrophile, what the on-target ligand occupancy is; and (3) what response is triggered as a result of POI-electrophile target engagement *in vivo*. (b) The 'Breaking the Tether' concept[31] underlying T-REX: the electrophile rapidly liberated ($t_{1/2} < 1$ min) within the proximity of Halo-POI has a choice to be irreversibly-captured by the POI, or diffuse away from the encounter complex at its native diffusion rate. (Scenario on the left shows POI as an electrophile sensor where some level of labeling occurs; the right shows the case where the POI is not a sensor of the electrophile). See also Fig. 4. NOTE: maximum concentration of electrophile released is stochiometric to intracellular concentration of Halo-POI (in T-REX) or Halo (in G-REX) (measured to be < 5 µM typically in cells.[39] but also tunable), thus mimicking native electrophile signaling under limited-electrophile dosage. Figures are adapted from the author's original research publications cited above, and the following publishers are acknowledged: American Chemical Society Elsevier, and John Wiley and Sons.

.engagement for numerous POIs without the need to perform large amount of construct cloning, *etc.* When we came across a Halo-tagged-mediated protein-degradation method called 'HyT' from the Crews laboratory,[34] our interest in Halo-tagging was piqued. This was especially because unlike other protein tags, there is a Halo ORFeome library covering 20,000 human and mouse genes. So, we realized that if the general idea were to prove successful, the Halo-ORFeome would facilitate testing of almost any POI from human or mouse origins. Accordingly, we designed Halo-targetable small-molecule photocaged probes for what we later termed as T-REX (Targetable reactive electrophiles and oxidants) technology that can release a given electrophile at a preordained time. We designed the probe and validated photouncaging $t_{1/2}$, permeability, lack of toxicity, and stoichiometric binding to Halo in living cells[26,29,30,35–43] and subsequently live animals[38,39,41,43,44] (Fig. 4).[33]

The next question was what we should choose as the test protein to be the recipient of a representative LDE (HNE) released from Halo. Thanks to a senior technician scientist from a colleague's laboratory, we first became aware of the quintessential electrophile-regulated Keap1/Nrf2 antioxidant response pathway[45] around the time the idea of T-REX was being proposed in late 2011 as part of the author's then upcoming independent career. When this pathway later became a target of the approved drug, Tecfidera (and now also Vumerity),[6] and that this pathway was (and continues to be) hotly pursued by several pharma industries

and medicinal-chemistry/drug-discovery-driven academic labs, the author's interest was further piqued. For us, the established nature of this pathway, and its broader relevance meant it was the perfect initial test ground for T-REX when our independent laboratory began in mid-2012. We were excited to see that T-REX to Halo-fused-Keap1 could transfer around 40% of released HNE, from photocaged-HNE bound to Halo, to Keap1 post photouncaging; few other cellular proteins were labeled.[35,36] A stark contrast was noted in specificity offered by T-REX-assisted delivery of HNE (Fig. 4b, inset) vs. following whole-cell treatment with HNE.[29,30,35,36] Through the author's reading of the literature in the redox-related field, we came across PTEN, a tumor-suppressor, that is also stress-responsive.[46] This was the second protein validated to be compatible with T-REX in our laboratory's early years.[35,36] Building on these inaugural proof-of-concept studies, we have now shown successful electrophile delivery to >10 POIs *via* T-REX.[26,2930,35–44] Many of these proteins were shown to be highly electrophile reactive *in vitro*.[39,40]

With target-specificity validation in hand, we were eager to apply T-REX in two unique ways. One way was, for a known POI, to map on-target responses precisely, and in the process discover nuanced pathway mechanisms and intersections that are otherwise hidden under bolus conditions. The second way was to discover novel electrophile-responsive POIs whose modifications gives unexpected signaling outputs. Over the past 8 years, through pending and published papers, we successfully demonstrated the
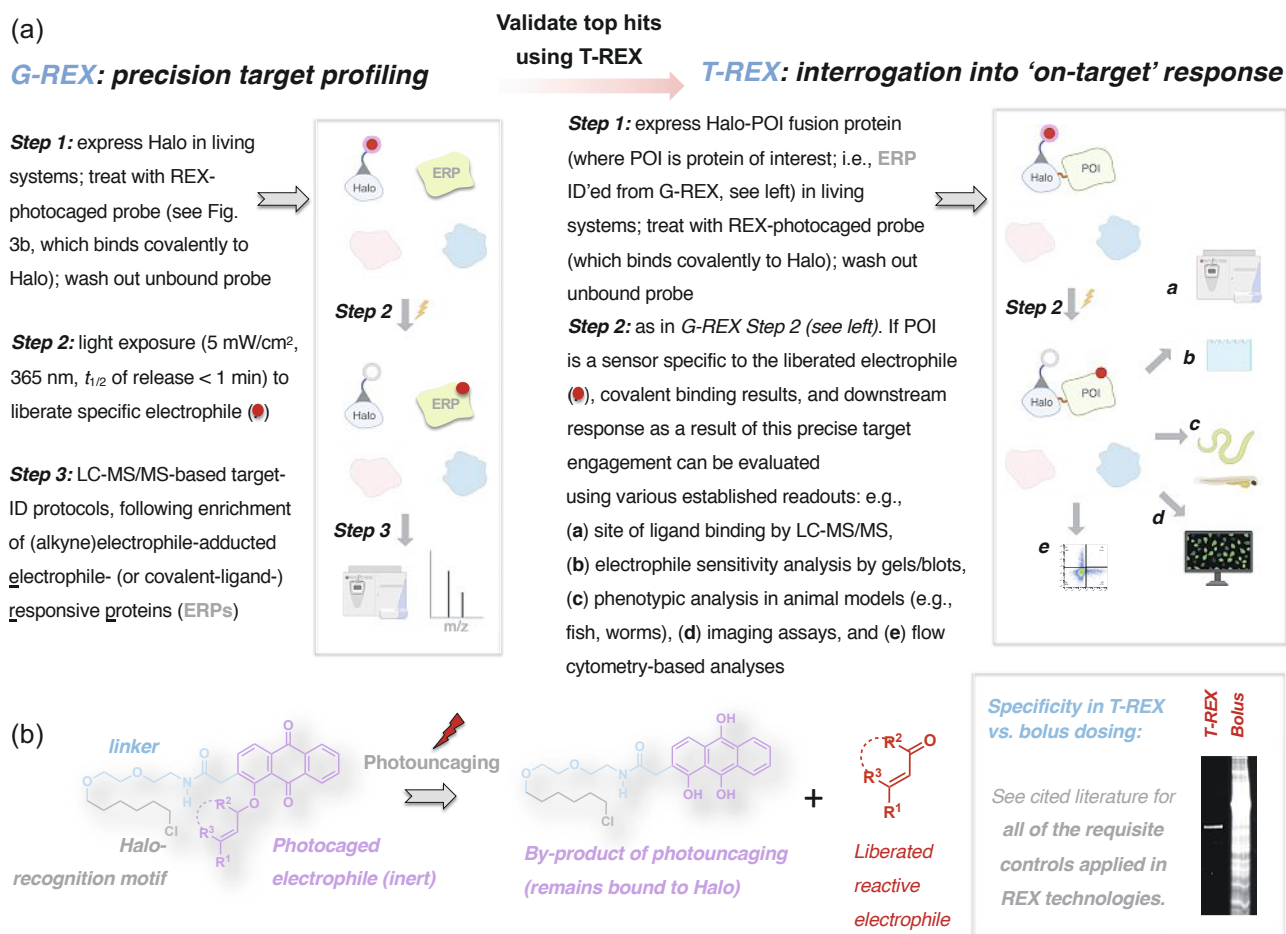

(a)

Validate top hits using T-REX

G-REX: *precision target profiling* → **T-REX:** *interrogation into 'on-target' response*

Step 1: express Halo in living systems; treat with REX-photocaged probe (see Fig. 3b, which binds covalently to Halo); wash out unbound probe

Step 2: light exposure (5 mW/cm², 365 nm, $t_{1/2}$ of release < 1 min) to liberate specific electrophile (●)

Step 3: LC-MS/MS-based target-ID protocols, following enrichment of (alkyne)electrophile-adducted electrophile- (or covalent-ligand-) responsive proteins (**ERPs**)

Step 1: express Halo-POI fusion protein (where POI is protein of interest; i.e., **ERP** ID'ed from G-REX, see left) in living systems; treat with REX-photocaged probe (which binds covalently to Halo); wash out unbound probe

Step 2: as in *G-REX Step 2 (see left)*. If POI is a sensor specific to the liberated electrophile (●), covalent binding results, and downstream response as a result of this precise target engagement can be evaluated using various established readouts: e.g., (**a**) site of ligand binding by LC-MS/MS, (**b**) electrophile sensitivity analysis by gels/blots, (**c**) phenotypic analysis in animal models (e.g., fish, worms), (**d**) imaging assays, and (**e**) flow cytometry-based analyses

(b)

linker — Photouncaging — *By-product of photouncaging (remains bound to Halo)* + *Liberated reactive electrophile*

Halo-recognition motif — *Photocaged electrophile (inert)*

Specificity in T-REX vs. bolus dosing:

See cited literature for all of the requisite controls applied in REX technologies.

Fig. 4. Workflow of REX technologies[31] and underlying chemistry. (a) Left: G-REX functional proteome/subproteome profiling of native electrophile/covalent-ligand-responsive proteins (ERPs). G-REX offers space, time, (intracellular/subcellular) dosage controls in electrophile delivery, parameters that are not controllable in other existing proteomics target-ID strategies. Right: Top hits from G-REX are validated by T-REX. (b) Light-driven electrophile liberation underpinning the REX technologies. Inset shows data from an in-gel fluorescence analysis following T-REX-assisted target-specific covalent ligand (specifically, HNE, see Fig. 1d) modification of Keap1 protein in live cells, in direct comparison against the outcomes from live cells bolus treated with HNE. See literature cited within the text for experimental details and accompanying data from relevant control experiments. Fig. 4a and inset within Fig. 4b are reused/adapted from the author's original research publications and the respective publishers, Elsevier and American Chemical Society, are acknowledged.

versatility of REX technologies in both ways, and documented the nuanced roles of electrophile signaling in multiple biologically-important responses such as cell death[26,38] and DNA damage response,[39] beyond antioxidant response.[29,30,36,37,41,42] Importantly these data have given the first direct evidence that low-occupancy electrophile modifications on a specific POI is biologically sufficient to elicit a functional response. This observation was made with Keap1/Nrf2 AR regulation,[29,30,36,37] PTEN/Akt/FOXO cascade,[38] Ube2V2/DNA-damage signaling,[39] and Ube2V1/NFκB signaling,[39] representing both dominant inhibitory and stimulatory modalities of signal amplification.[47] These data furnish compelling evidence that natural electrophilic signals regulate cell signaling in the same way as canonical PTMs such as phosphorylation and ubiquitination, targeting specific sentinel proteins, or 'privileged first responders'.[4,23] In the process, we discovered novel pathway cross-talk and unanticipated codes of 'cellular currency exchange'[28] between reactive electrophiles and canonical PTM signals.[48] *These multiple independent lines of evidence allowed us to consider natural electrophiles as potential pharmacophores[17] and that understanding their signal rewiring mechanisms, will uniquely enable target ID and novel ligand design principles.*[4]

Engineering G-REX: Profiling Low-occupancy Covalently-ligandable Proteins with High Spatiotemporal Resolution

In our initial years, we were largely preoccupied with testing the scope of T-REX, setting up suites of controls, and extending the method to multiple model organisms.[33] Recently, we have also given much consideration toward addressing what applied insights can be leveraged from T-REX, and how REX technologies can be utilized for translational medicine. Here, I devote the remaining sections to highlight how we have moved forward in addressing both questions. In terms of the first perhaps obvious limitation, T-REX is a one-by-one method, and although we have successfully shown medium-throughput screening capabilities in identifying novel functional electrophile sensor proteins, we soon realized that this is inherently limited, and so we set about developing a 'proteome-wide' screening approach. We thus recently developed G-REX technology[33,39] that allows quantitative mapping of functional sub-proteome responsive to a given covalent ligand in low-occupancy (Fig. 4). Distinct from the state-of-the-art proteomics profiling of covalent ligand responsive proteins, G-REX is the first and the only tool that can directly profile functional targets of covalent ligands with subcellular-level spatiotemporal control in intact living models.[23] Of course, spatiotemporal-controlled subproteomics indexing of localized proteins have lately surged in the literature.[25] However, none of these innovative methods profile *function* in a local-specific manner. Thus G-REX, as in T-REX, remains the first and only method for precision covalent ligand–target profiling/validation. Our accumulating data from applying G-REX-&-T-REX in tandem (Fig. 4a) in pending publications promise new opportunities to researchers in both industry and academia alike whereby functional targets can be directly screened in an unbiased way with precise timing and biological context. Of course, if the covalent pharmacophores do not bear enal/enone-type frameworks, a different photocaging strategy must be designed. We consider these challenges as value-added opportunities for us to innovate novel biocompatible chemical platforms and thus such directions are currently in progress.

REX Technologies Usher Logical Engineering of Covalent Inhibitors

In terms of the second direction, namely, translational science, burgeoning evidence allowed us to propose that privileged electrophile sensor proteins unearthed by REX technologies are an untapped resource for covalent drug discovery.[4] Indeed, since early 2017, we began to ponder how accumulating new insights

and fundamental principles underpinning precision electrophile signaling code can be translated ultimately to precision medicine development. From the perspective of tool-developers, this is also a real-world-relevant question since intrinsic requirement of ectopic expression, use of non-native Halo protein, and light-driven photouncaging are, to say the least, cumbersome for therapeutic applications.[4] We thus launched a proof-of-principle project to apply our T-REX data to medicinal chemistry problems. Our idea was to leverage our recent identification of Akt3 as an isoform-specific sensor of native reactive metabolite (HNE), and subsequent extensive pathway validations in both cultured cells and zebrafish embryos,[38] toward developing first-in-class isoform-specific covalent inhibitors targeting Akt3, one of the three non-redundant isoforms of Akt kinase. PI3K/Akt-cascade is a validated druggable nexus interfacing oncogenic signaling and cancer metabolism[46,49] with two small-molecule PI3K-inhibitors approved for targeted cancer therapies. Conversely, eight small-molecule Akt inhibitors are currently in clinical trials, although none are single-isoform-specific; some exhibit isoform-bias toward Akt1 or Akt2.[50]

T-REX-assisted studies in cells and fish designated C119 as the residue unique to Akt3 that senses HNE[38] (Fig. 5a). HNEylation of this residue promotes superstoichiometric inhibition of Akt3, leading us to propose that there is a dominant-negative effect. Homology modeling of Akt3 based on the available full-length protein structure of human Akt1, shows C119 residing in a flexible linker region (Fig. 5b),[4] bridging the PH and the kinase domains, featuring high sequence divergence across the three Akt-isoforms. C119 is in closer proximity to a known allosteric site (targeted by several small-molecule candidates in trials) than the ATP-binding site.[26] Thus, we proposed a chimera of an allosteric-site targeting drug candidate, conferring general Akt-selectivity, and a pendant HNE-derived arm, conferring irreversible Akt3-binding, would overall give us an irreversible Akt3 inhibitor. Among the eight Akt-inhibitors currently in clinical trials, three of them target the allosteric site.[51,52] All bear a primary amine, to which facile appendage of an HNE-derived motive was anticipated, especially because SAR-data indicate the amine to be largely dispensable. MK-2206 was selected because it is the first orally-effective allosteric reversible pan-Akt inhibitor that reached clinical trials and continues to boast several ongoing/active Phase-I/II-trials targeting a large spectrum of solid and blood tumors. Thus, significant amount of useful data are available for ultimately benchmarking against our own inhibitors. Our choice was further motivated by treatment-related hyperglycemia, a side effect common to many targeted cancer drugs,[53] and in the case of MK-2206 hinders its clinical progress in the treatment of stage 1–3 breast cancers.[54–57] Novel isoform-specific targeting together with the general benefits of covalent druggability could potentially alleviate such clinically-observed adverse effects.[4]

The molecules so designed emerged to efficiently and selectively label Akt3, as we predicted. We then established that Akt3-selectivity manifested by our chimeras, *e.g.*, MK-H(F)NE (Fig. 5c), is traceable to the HNE-like appendage.[26] A pseudo-regioisomer of MK-H(F)NE, MK-G (Fig. 5c) showed much lower covalent ligandability, and showed enhanced Akt2-specificity. Removal of the Michael-acceptor motif in MK-HNA muted covalent binding as expected (Fig. 5c). Notably, 2nd-order kinetics of MK-H(F)NE for Akt3 inhibition (Fig. 5c) approach those reported for some successful covalent kinase drugs in the market.[58] Further consistent with irreversible binding, compound withdrawal experiments demonstrated persistent inhibitory and growth-suppressive effects in cultured cells. Unexpectedly, in terms of mechanism-of-action at the pathway level, the hybrids, MK-H(F)NE, phenocopy HNE-specific pathway-regulatory features[38] more so than those of MK-2206: (i) MK-H(F)NE shows dominant-negative pathway suppression mimicking the behavior of HNE previously
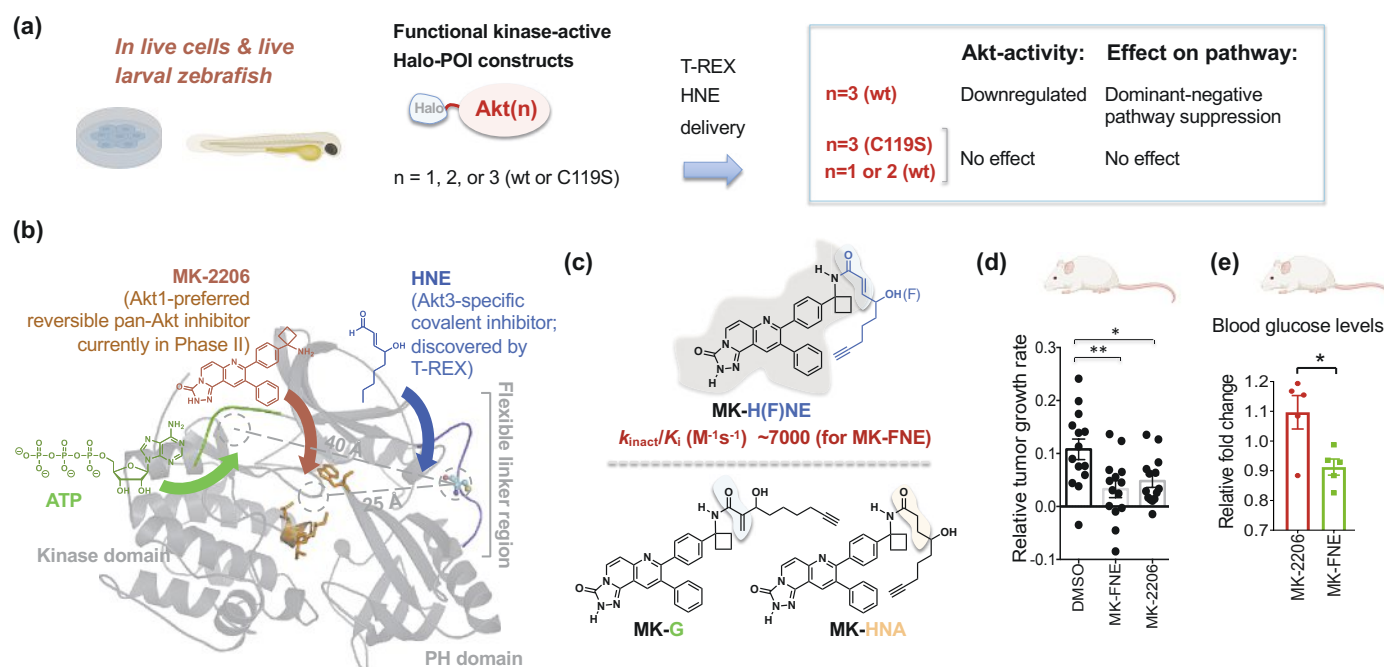
Fig. 5. Application of T-REX in translating the precision electrophile signaling code. (a) Akt3-isoform-specific electrophile signaling discovery enabled by T-REX: briefly, in cells and fish, T-REX-assisted target-specific HNEylation of Akt3 [and not Akt1/2 or Akt3(C119S)] results in downregulation of Akt-kinase activity and consequent dominant-negative suppression of Akt(n)-dependent oncogenic signaling. (b) Homology-model of Akt3-kinase derived from Akt1-crystal-structure (PDB:3O96), featuring ATP-binding-catalytic site, proposed MK-2206-targeting allosteric site (common to all Akt-isoforms) featuring key interacting residues, and HNE-sensing site (located in the flexible linker region) with C119-residue (unique to Akt3) in ball-and-stick representation. (c) Rationally-designed Akt3-specific covalent inhibitor, a hybrid of MK-2206 and HNE (see Fig. 1b, 4b), exhibiting an inhibition mechanism more akin to how HNE interacts with Akt3, than MK-2206. Lower row: Regioisomeric MK-G does not show efficacy, ruling out electrophilicity alone being attributable to specificity/efficacy. MK-HNA housing no Michael-acceptor motif shows no covalent labeling nor Akt3-specificity. (d) MDA- MB-468 xenograft mice were orally administered indicated compounds or vehicle DMSO (twice weekly) and resultant relative fold change in tumor size monitored over 9 days is shown. (e) Unlike MK2206, MK-FNE (analog of MK-HNE, but with higher *in vivo* efficacy) did not raise blood-sugar levels, a major off-target phenotype hampering clinical progress of MK2206. Briefly, MDA-MB-468 xenograft mice were orally administered either MK-2206 or MK-FNE or vehicle DMSO (twice weekly) and resultant measurement of blood glucose levels was undertaken. Data show blood glucose levels at Day 21 (4 days post termination of compound administration). All p values are from two-tailed unpaired *t*-test. Error bars indicate s.e.m. 5 mice per group (10 tumors). (Treatment conditions used were modeled on similar studies in xenografts models completed for the current pan-Akt inhibitor in Phase II, MK-2206).[26,27] Fig. 5b, 5d, and 5e are reproduced/adapted from the author's original research publication cited above and the publisher American Chemical Society is acknowledged.

reported by us using T-REX;[38] (ii) whereas MK-2206 treatment blocks Akt-T305-phosphorylation (a key Akt-activating event), the hybrid compound does not, mirroring what we reported for HNE-metabolite signaling and pathway downregulation *via* Akt3-covalent modification.[38]

Profiling of various cancer cell lines led us to identify that breast cancer (BC) lines – especially BCs deprived of critical tumor-suppressor protein *PTEN*[46,49] (*i.e.*, *pten*-null BCs) – are more sensitive to MK-H(F)NE over MK-2206. Enhanced sensitivity for MK-H(F)NE is most prominent (~5 fold) in MK-2206-resistant lines where resistance arises from overexpression of Akt3 (the isoform for which MK-2206 has the least affinity).[26] These observations are consistent with MK-H(F)NE targeting Akt3. We further confirmed on-target isoform-specificity by showing that our chimeric inhibitors synergized exclusively with Akt3-knockdown, whereas MK-2206 synergized exclusively with depletion of Akt1 (its preferred target). MK-H(F)NE also proved useful in uncovering novel downstream genes whose phosphorylation status is altered in an Akt-isoform-dependent manner and distinct from those modulated in response to MK-2206 administration. The lack of Akt3-specific inhibitors had rendered such investigations hitherto inaccessible. These data provided compelling confidence to extend our investigations to *in vivo* disease models.

In mice liver microsomes, MK-H(F)NE underwent a faster rate of degradation than MK-2206 ($t_{1/2}$ < 30 min compared to ~86 min for the non-covalent congener, MK-2206).[26] Nonetheless, T-REX and our related validations demonstrated that low-occupancy Akt3(C119)-covalent engagement is sufficient for pathway inhi-

bition. Furthermore, our second-order kinetics of Akt3-labeling were favorable, and for targeted covalent therapeutics, achieving precise target engagement is often of greater relevance than metabolic lability.[4,59] We thus progressed to profile *in vivo* efficacy in MDA-MB-468 xenograft mice, whereby oral administration of MK-FNE emerged to be at least as equally effective at suppressing triple-negative BCs as MK-2206 (Fig. 5d). Importantly, MK-FNE-treatment also resulted in a better management of hyperglycemia, one of the current dose-and-efficacy-limiting side effects of Akt-inhibition *in vivo*[54–57] (Fig. 5e).

Synopsis and Outlook

In sum, the REX technologies we have built and rigorously road-tested have revealed some of the most nuanced and fundamental, and at times, unexpected facets of cell decision making. These accumulating data are also simultaneously helping to shape the ways we understand cysteine responsivity. Our ongoing projects further integrate precision cysteome understanding along with phylogeny, and transcriptomics and other systems-level approaches. When such analyses are coupled with G-REX proteomics mapping and T-REX precision interrogations, we hope to paint a more comprehensive picture of functional outputs, both at global and single-cysteine-specific level. Unique capabilities to control and investigate at will, specific proteins and covalent-ligand pairs are also expected to bear fruit as industrially-adaptable avenues for profiling targets, dissecting regulatory networks that can also be harnessed to modulate cell behavior, and ultimately amend disease states.

Competing interests

The REX technologies and Akt inhibitors derived from applications of these technologies were filed for patent applications by Cornell University, USA, where formative years of the author's laboratory were spent.

Received: June 23, 2020

[1] J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discov*. **2011**, *10*, 307.

[2] M. Visscher, M. R. Arkin, T. B. Dansen, *Curr. Opin. Chem. Biol*. **2016**, *30*, 61.

[3] A. Chaikuad, P. Koch, S. A. Laufer, S. Knapp, *Angew. Chem. Int. Ed.* **2018**, *57*, 4372.

[4] M. J. C. Long, Y. Aye, *Cell Chem. Biol.* **2017**, *24*, 787.

[5] A. J. Maurais, E. Weerapana, *Curr. Opin. Chem. Biol.* **2019**, *50*, 29.

[6] J. R. Poganik, Y. Aye, *Front. Aging Neurosci.* **2020**, 12.

[7] F. J. Schopfer, C. Cipollina, B. A. Freeman, *Chem. Rev.* **2011**, *111*, 5997.

[8] A. T. Jacobs, L. J. Marnett, *Acc. Chem. Res.* **2009**, *43*, 673.

[9] S. Parvez, M. J. C. Long, J. R. Poganik, Y. Aye, *Chem. Rev.* **2018**, *118*, 8798.

[10] A. Luengo, K. L. Abbott, S. M. Davidson, A. M. Hosios, B. Faubert, S. H. Chan, E. Freinkman, L. G. Zacharias, T. P. Mathews, C. B. Clish, R. J. DeBerardinis, C. A. Lewis, M. G. Vander Heiden, *Nat. Commun.* **2019**, *10*, 5604.

[11] B. K. Park, A. Boobis, S. Clarke, C. E. P. Goldring, D. Jones, J. G. Kenna, C. Lambert, H. G. Laverty, D. J. Naisbitt, S. Nelson, D. A. Nicoll-Griffith, R. S. Obach, P. Routledge, D. A. Smith, D. J. Tweedie, N. Vermeulen, D. P. Williams, I. D. Wilson, T. A. Baillie, *Nat. Rev. Drug Discov.* **2011**, *10*, 292.

[12] V. Ramesh, J. W. Locasale, *Nat. Immunol.* **2020**, *21*, 497.

[13] L. A. J. O'Neill, M. N. Artyomov, *Nat. Rev. Immunol.* **2019**, *19*, 273.

[14] M. J. Bollong, G. Lee, J. S. Coukos, H. Yun, C. Zambaldo, J. W. Chang, E. N. Chin, I. Ahmad, A. K. Chatterjee, L. L. Lairson, P. G. Schultz, R. E. Moellering, *Nature* **2018**, *562*, 600.

[15] M. Sciacovelli, E. Gonçalves, T. I. Johnson, V. R. Zecchini, A. S. H. da Costa, E. Gaude, A. V. Drubbel, S. J. Theobald, S. R. Abbo, M. G. B. Tran, V. Rajeeve, S. Cardaci, S. Foster, H. Yun, P. Cutillas, A. Warren, V. Gnanapragasam, E. Gottlieb, K. Franze, B. Huntly, E. R. Maher, P. H. Maxwell, J. Saez-Rodriguez, C. Frezza, *Nature* **2016**, *537*, 544.

[16] J. R. Poganik, M. J. C. Long, Y. Aye, *Bioessays* **2018**, *40*, e1700240.

[17] M. J. C. Long, X. Liu, Y. Aye, *Curr. Opin. Chem. Biol.* **2019**, *51*, 48.

[18] R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologa, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea, J. P. Overington, *Nat. Rev. Drug Discov.* **2017**, *16*, 19.

[19] Y. Fu, M. J. C. Long, S. Wisitpitthaya, H. Inayat, T. M. Pierpont, I. M. Elsaid, J. C. Bloom, J. Ortega, R. S. Weiss, Y. Aye, *Nat. Chem. Biol.* **2018**, *14*, 943.

[20] M. J. C. Long, Y. Zhao, Y. Aye, *Cell Chem. Biol.* **2020**, *27*, 122.

[21] S. Wisitpitthaya, Y. Zhao, M. J. C. Long, M. Li, E. A. Fletcher, W. A. Blessing, R. S. Weiss, Y. Aye, *ACS Chem. Biol.* **2016**, *11*, 2021.

[22] M. J. C. Long, L. Wang, Y. Aye, *Antioxid. Redox Signal.* **2019**, doi: 10.1089/ars.2019.7894.

[23] X. Liu, M. J. C. Long, Y. Aye, *Trends Biochem. Sci.* **2019**, *44*, 75.

[24] M. J. C. Long, J. R. Poganik, S. Ghosh, Y. Aye, *ACS Chem. Biol.* **2017**, *12*, 586.

[25] M. J. C. Long, Y. Zhao, Y. Aye, *RSC Chem. Biol.* **2020**, DOI: 10.1039/d0cb00041h.

[26] X. Liu, M. J. C. Long, B. D. Hopkins, C. Luo, L. Wang, Y. Aye, *ACS Cent. Sci.* **2020**, DOI:10.1021/acscentsci.9b00893.

[27] Y. Aye, X. Liu, M. J. C. Long, PCT/US18/36210, filed 06/**2018**.

[28] M. J. C. Long, Y. Aye, *Chem. Res. Toxicol.* **2016**, *29*, 1575.

[29] S. Parvez, Y. Fu, J. Li, M. J. Long, H. Y. Lin, D. K. Lee, G. S. Hu, Y. Aye, *J. Am. Chem. Soc.* **2015**, *137*, 10.

[30] H. Y. Lin, J. A. Haegele, M. T. Disare, Q. Lin, Y. Aye, *J. Am. Chem. Soc.* **2015**, *137*, 6232.

[31] M. J. C. Long, J. R. Poganik, Y. Aye, *J. Am. Chem. Soc.* **2016**, *138*, 3610.

[32] M. J. C. Long, D. A. Urul, Y. Aye, *Meth. Enzymol.* **2020**, *633*, 203.

[33] J. R. Poganik, M. J. C. Long, Y. Aye, *Trends Biochem. Sci.* **2019**, *44*, 380.

[34] T. K. Neklesa, H. S. Tae, A. R. Schneekloth, M. J. Stulberg, T. W. Corson, T. B. Sundberg, K. Raina, S. A. Holley, C. M. Crews, *Nat. Chem. Biol.* **2011**, *7*, 538.

[35] X. Fang, Y. Fu, M. J. Long, J. A. Haegele, E. J. Ge, S. Parvez, Y. Aye, *J. Am. Chem. Soc.* **2013**, *135*, 14496.

[36] S. Parvez, M. J. Long, H. Y. Lin, Y. Zhao, J. A. Haegele, V. N. Pham, D. K. Lee, Y. Aye, *Nat. Prot.* **2016**, *11*, 2328.

[37] M. J. Long, H. Y. Lin, S. Parvez, Y. Zhao, J. R. Poganik, P. Huang, Y. Aye, *Cell Chem. Biol.* **2017**, *24*, 944.

[38] M. J. C. Long, S. Parvez, Y. Zhao, S. L. Surya, Y. Wang, S. Zhang, Y. Aye, *Nat. Chem. Biol.* **2017**, *13*, 333.

[39] Y. Zhao, M. J. C. Long, Y. Wang, S. Zhang, Y. Aye, *ACS Cent. Sci.* **2018**, *4*, 246.

[40] S. L. Surya, M. J. C. Long, D. A. Urul, Y. Zhao, E. J. Mercer, E. I. IM, T. Evans, Y. Aye, *ACS Chem. Biol.* **2018**, *13*, 1824.

[41] J. R. Poganik, M. J. C. Long, M. T. Disare, X. Liu, S. H. Chang, T. Hla, Y. Aye, *FASEB J.* **2019**, *33*, 14636.

[42] J. R. Poganik, A. K. Van Hall-Beauvais, M. J. C. Long, M. T. Disare, Y. Zhao, Y. Aye, *Helv. Chim. Acta* **2020**, *103*, e2000041.

[43] A. Van Hall-Beauvais, Y. Zhao, D. A. Urul, M. J. C. Long, Y. Aye, *Curr. Protoc. Chem. Biol.* **2018**, *10*, e43.

[44] M. J. Long, D. A. Urul, S. Chawla, H.-Y. Lin, Y. Zhao, J. A. Haegele, Y. Wang, Y. Aye, *Biochem.* **2018**, *57*, 216.

[45] J. D. Hayes, A. T. Dinkova-Kostova, *Trends Biochem. Sci.* **2014**, *39*, 199.

[46] C. C. Dibble, L. C. Cantley, *Trends Cell Biol.* **2015**, *25*, 545.

[47] N. E. Hynes, P. W. Ingham, W. A. Lim, C. J. Marshall, J. Massagué, T. Pawson, *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 393.

[48] M. J. C. Long, X. Liu, Y. Aye, *Front. Chem.* **2019**, 7.

[49] G. Hoxhaj, B. D. Manning, *Nat. Rev. Cancer* **2020**, *20*, 74.

[50] L. Quambusch, I. Landel, L. Depta, J. Weisner, N. Uhlenbrock, M. P. Müller, F. Glanemann, K. Althoff, J. T. Siveke, D. Rauh, *Angew. Chem. Int. Ed.* **2019**, *58*, 18823.

[51] K. Guo, W. Tang, H. Zhuo, G. Zhao, *Chem. Sel.* **2019**, *4*, 9040.

[52] J. S. Brown, U. Banerji, *Pharmacol. Therapeut.* **2017**, *172*, 101.

[53] J. W. Goldman, M. A. Mendenhall, S. R. Rettinger, *Oncologist* **2016**, *21*, 1326.

[54] A. Akcakanat, F. Meric-Bernstam, *Ann. Transl Med.* **2018**, *6*, S57.

[55] C. X. Ma, V. Suman, M. P. Goetz, D. Northfelt, M. E: Burkard, F. Ademuyiwa, M. Naughton, J. Margenthaler, R. Aft, R. Gray, A. Tevaarwerk, L. Wilke, T. Haddad, T. Moynihan, C. Loprinzi, T. Hieken, E. K. Barnell, Z. L. Skidmore, Y.-Y. Feng, K. Krysiak, J. Hoog, Z. Guo, L. Nehring, K. B. Wisinski, E. Mardis, I. S. Hagemann, S. Vij, S. Sanati, H. Al-Kateb, O. L. Griffith, M. Griffith, L. Doyle, C. Erlichman, M. J. Ellis, *Clin. Cancer Res.* **2017**, *23*, 6823.

[56] C. Simioni, L. M. Neri, G. Tabellini, F. Ricci, D. Bressanin, F. Chiarini, C. Evangelisti, A. Cani, P. L. Tazzari, F. Melchionda, P. Pagliaro, A. Pession, J. A. McCubrey, S. Capitani, A. M. Martelli, *Leukemia* **2012**, *26*, 2336.

[57] K. Kalinsky, J. A. Sparano, X. Zhong, E. Andreopoulou, B. Taback, L. Wiechmann, S. M. Feldman, P. Ananthakrishnan, A. Ahmad, S. Cremers, A. N. Sireci, J. R. Cross, D. K. Marks, P. Mundi, E. Connolly, K. D. Crew, M. A. Maurer, H. Hibshoosh, S. Lee, D. L. Hershman, *Clin. Transl Oncol.* **2018**, *20*, 1474.

[58] P. A. Schwartz, P. Kuzmic, J. Solowiej, S. Bergqvist, B. Bolanos, C. Almaden, A. Nagata, K. Ryan, J. Feng, D. Dalvie, J. C. Kath, M. Xu, R. Wani, B. W. Murray, *Proc. Natl Acad. Sci. USA* **2014**, *111*, 173.

[59] R. A. Copeland, *Nat. Rev. Drug Discov.* **2016**, *15*, 87.