Getting the Right Grip? How Understanding Electrophile Selectivity Profiles Could Illuminate Our Understanding of Redox Signaling

Marcus J. C. Long[†], Lingxi Wang[‡], and Yimon Aye[‡,*]

[‡]Swiss Federal Institute of Technology Lausanne (EPFL), 1015, Lausanne, Switzerland

[†]47 Pudding Gate, Bishop Burton, Beverley East Riding of Yorkshire, HU17 8QH, UK

[*]Correspondence: Prof. Yimon Aye. CH A2 397, EPFL, 1015, Lausanne, Switzerland. Tel: +41-(0)79 838 94 47. Fax: +41-(0)21 693 66 50. E-mail: yimon.aye@epfl.ch

Abstract: Significance: Electrophile signaling is coming into focus as a bona fide cell signaling mechanism. The electrophilic regulation occurs typically through a sensing event (i.e., labeling of a protein) and a signaling event (the labeling event having an effect of the proteins activity, association, etc.). **Recent advances:** Herein, we focus on the first step of this process, electrophile sensing. Electrophile sensing is typically a deceptively simple reaction between the thiol of a protein-cysteine, of which there are around 200,000 in the human proteome, and a Michael acceptor, of which there are numerous flavors, including enals and enones. Recent data overall paint a picture that despite being a simple chemical reaction, electrophile sensing is a discerning process, showing labeling preferences that are often not in line with reactivity of the electrophile. **Critical issues:** With a view to trying to decide what brings about highly electrophile-reactive protein-cysteines, and how reactive these sensors may be, we discuss aspects of the thermodynamics and kinetics of covalent/non-covalent binding. Data made available by several laboratories indicate that it is likely that specific proteins exhibit highly stereo- and chemoselective electrophile sensing, which we take as good evidence for recognition between the electrophile and the protein prior to forming a covalent bond. **Future directions:** We propose experiments that could help us gain a better and more quantitative understanding of the

Long, et al.

mechanisms through which sensing comes about. We further extoll the importance of performing more detailed experiments on labeling and trying to standardize the way we assess protein-specific electrophile sensing.

Key words: *Covalent labeling; stereoselectivity; electrophile signaling; affinity; mechanism; kinetic control*

Long, et al.

Man's appreciation for molecular shape, configuration, and topology in biology has had a checkered history. These vacillations are perhaps best exemplified by our understanding of enantiomers, although similar arguments hold for other minor structural changes. By the 1800's, we had started to understand that many naturally-occurring compounds had the ability to change the orientation of plane-polarized light. The examples included several naturally-derived inorganic substances, like quartz crystals (16), as well as bioactive natural products, such as tartrate (15). In the 1840's, Louis Pasteur performed his famous experiment separating racemic tartaric acid into different crystalline forms, which transpired to contain the resolved enantiomers of the compound (15). Later, it was noted that the enantiomers of borneol had distinctly-different odors, hinting that molecular recognition was linked to smell, and that our bodies can distinguish effectively between enantiomers (6). By the 1930s, the concept of enantiomorphism in biology had permeated the general conscience. A contemporary detective series writer, Dorothy L. Sayers, even used the concept that naturally-derived poisons are homochiral and synthetic versions (like would be found in a synthetic poison) are racemic as the denouement to her novel "The Documents in the Case" (48). In the book, the "poison" was muscarine, a sugar analog that is an agonist of the muscarinic acetylcholine receptor (23). Yet this appreciation for the differential biological properties of enantiomers did little to prepare scientists for the thalidomide crisis in the 1960s that has left medicinal chemists with the critical knowledge that enantiomers can be divergently bioactive. Ironically, this debacle occurred some 10 years after the observations on the potency of different enantiomers by C. Pfeiffer (5).

A silent assumption in the above examples of molecules interacting with biological receptors/enzymes/proteins, is that non-covalent binding is the key driving force of the interaction. Indeed, it is the concept of non-covalent binding and the required spatial recognition between enzyme and substrate (or protein and ligand) that has preoccupied biochemists and biologists arguably since the work of Leonor Michaelis and Maud Menten in 1913, and the development of their eponymous equation (or its precursor, the Henri equation) (11). With our current understanding that ligand recognition is a key determinant of biological processes, it is clear why the different enantiomers of different molecules can exhibit divergent

Long, et al.

biological properties, such as having different odors. Because high affinity translates to large binding energy, the stabilization of the system incurred upon ligand binding can also be leveraged to induce changes in stabilization, conformation, association, etc., that can be transmuted into meaningful changes in cellular decision-making. The binding requirements can be so stringent, that even very similar molecules that are present together in the cellular milieu can be suitably differentiated by enzyme active sites, and hence this "chemical noise" is efficiently filtered out.

The Michaelis-Menten equation also extends to situations where the enzyme forms a covalent bond to the substrate. The covalent intermediate can be harnessed for numerous possible ends, such as: to promote fission, e.g., in serine proteases; or to promote reaction with an electrophile, e.g., in thymidylate synthetase. Nevertheless, it remains generally true that it is the ability of the enzyme to accommodate the substrate(s) in the active site, and orientate reactive positions to promote reactivity that ultimately synergize to allow an overall selective process to occur. For instance, deubiquitinating enzymes mostly hydrolyze essentially "the same" substrate, i.e., an isopeptide bond formed between ubiquitin's C-terminus and a protein-lysine. Selectivity is engendered by a combination of recognition of ubiquitin chain linkage, and the protein to which the ubiquitin(s) is(are) bound.

But we now understand that there are numerous very simple biological signaling molecules that are proposed to regulate numerous important pathways. These molecules, however, appear to be assumed to have little ability to interact non-covalently with their targets. So how can, and indeed, do these molecules interact selectively? In this opinion piece, we compare and contrast the thermodynamics and kinetics of covalent and non-covalent binding, to try to better understand how reactive electrophiles work, and to try to derive some insight into how selectivity and concepts of biochemistry can be used to begin to interrogate electrophile signaling in more detail. Our ultimate hope is that this work will serve as a reminder of the importance of non-covalent interactions in marshaling even overwhelmingly strong covalent bond-forming events, and further to underline the insight that can be brought to the table by examining stereo- and chemoselectivity in biological systems.

4

Long, et al.

Thermodynamics of covalent binding compared to non-covalent bonding

Covalent bond formation and non-covalent bonding are critically different. The bond energy of an sp^3-hybridized C–H bond is around 420 kJ mol^{-1}. Even a relatively weak sp^3-hybridized bond, where orbital overlap is considered poor, such as C–S, has a bond energy of around 290 kJ mol^{-1}. Alternatively the π-bond in C=C has an energy of 266 kJ mol^{-1}. By contrast, "only" 11.5 kJ mol^{-1} represents the transition-state stabilization energy required for a reaction to proceed with 99% ee at room temperature. This energy difference is less than 5 times the corresponding available thermal energy. Less stabilization is required as the temperature decreases (8 kJ mol^{-1} at –78 °C). Even what are considered high affinity, non-covalent binders (typically K_d = 0.1–10 nM, equivalent to $\Delta G_{binding}$ = ~60 to 46 kJ mol^{-1}) have binding energies significantly lower than that of particularly weak covalent bonds. Furthermore, to achieve even "nanomolar" dissociation constants, we know that arrangement of atoms in space relative to the binding pocket must be "close to perfect". For instance, even epimerization of the hydroxyl group in hydroxyl proline renders this ligand essentially unable to interact with its cognate partner, the von Hippel-Lindau E3 ligase (7). Loss of the hydroxyl group all together decreases the affinity by 3 orders of magnitude.

To illustrate this point further, we contrast one of the strongest non-covalent interactions known (the biotin–streptavidin association) with a representative covalent bond-forming event (the Michael addition of a thiol to an α,β-unsaturated carbonyl or nitrile to generate a β-thio-carbonyl or -nitrile, respectively). The biotin–streptavidin interaction (K_d ~ 10^{-15} M; $\Delta G_{binding}$ = ~–85 kJ mol^{-1}) (17, 36, 41) is certainly stronger than the interaction of even the strongest man-made non-covalent inhibitors with their target proteins (**Figure 1**). For instance, Immucillin H, a drug approved in Japan for treating peripheral T-cell lymphoma, is an inhibitor of purine nucleoside phosphorylase with K_d = 6 x 10^{-11} M; $\Delta G_{binding}$ = –60 kJ mol^{-1} (26). Indeed, aside from some exceptions of which biotin–streptavidin is the most well-known, pM affinity is believed to be near the limit of what is accessible for solely non-covalent-binding drugs and ligands (50).

Long, et al.

By comparison, thiol Michael adductions are particularly exergonic. The enthalpy change in going from an α,β-unsaturated amide or nitrile to a β-thio amide is around –100 kJ mol^{-1}. Calculated values for ΔG of simple thiol Michael additions vary, in some instances by around 2-fold, in the literature, and are indeed difficult to calculate and require specialized methods. Nevertheless, for instance, for methyl thiol adduction to but-3-en-2-one, an average ΔG of –90 kJ mol^{-1}, with a standard deviation of 17 kJ mol^{-1}, has been reported. Similar exergonicity with similar variances were quoted for thiol adduction to nitroethene and acrylonitrile. Relatively similar energy changes have been reported for numerous other Michael-type adductions: enolate addition to α,β-unsaturated esters (ΔH ~ –88 kJ mol^{-1}) (25), nitromethane coupling with cinnamaldehyde (ΔG ~ –70 kJ mol^{-1}) (52), with similar values being reported for acetylacetone adding to the same electrophile (46); and various amines adding to parthenolides (ΔG ~ –80 kJ mol^{-1}) (62), although values quoted for the aza-Michael adduction across a large range of substrates are quite variable (13) (**Figure 1**). There are two other points that are relevant to this comparison. The first point deals with the unusually-high atom efficiency of binding of most Michael acceptors. One way of estimating the efficiency of binding of a ligand to a target in terms of atom economy is ligand efficiency, defined as the negative of the Gibbs free energy of binding divided by the number of non-hydrogen atoms. For biotin, ligand efficiency is 5 kJ mol^{-1}/non-hydrogen atom (the maximum efficacy for non-covalent binding is approximately 6 kJ mol^{-1}) (24); the efficiency for Immucillin H is around 3 kJ mol^{-1}/non-hydrogen atom. On the other hand, Michael acceptors, e.g., but-3-en-2-one and ethene-nitrile, have ligand efficiencies of 18 and 23 kJ mol^{-1}/non-hydrogen atom, respectively. From this perspective, the variation in ΔG values for Michael adduction in the literature is minimal. This is because the values of ligand efficiency for ethene-nitrile is close to the highest values for interaction per non-hydrogen atom known in biology (24), which typically corresponds to metal-ion interactions with proteins (**Figure 2A**). The second standout point of our literature analysis of Michael adduction is that energies of binding are relatively independent of substitutions, although clearly there are some variations. Similar conclusions have also been made in the case of reversible Michael adductions (21). By contrast, the biotin–streptavidin interaction is sensitive to relatively small changes in structure, such as loss of the thioether motif (K_d ~ 10^{-10} M; ΔG = –57 kJ mol^{-1})

Long, et al.

(34) (**Figure 1**), whereas changes to the urea moiety can affect binding constants by 2.5 orders of magnitude for O to S substitution, and around 4.5 orders of magnitude for O to NH substitution (2, 17), although upon protonation, the imino functionalized analog binds streptavidin quite weakly (7–10 orders of magnitude change) (43, 47).

Covalent bond-forming events to simple biological electrophiles are overall slow and show little selectivity

From the preceding discussion, there is overwhelming thermodynamic driving force for bond formation that should render Michael adduction more favored than even the most potent non-covalent binding interactions. One could therefore naïvely ask what need is there for molecular recognition during Michael adduction? This question highlights a darker side of the large thermodynamic driving force and substituent independence of covalent bond-forming events: their thermodynamic favorability is also quite independent of the nucleophile. So, what could such relatively weak interactions benefit in the face of such large driving force and substituent independence?

This question clearly ignores the amount of time required for biological processes to occur. Specifically, when we discuss electrophile modification, either by an electrophilic drug or an electrophilic signaling molecule, we typically consider biological signaling processes, similar to phosphorylation or ubiquitination. Most drugs act on enzymes, typically inhibiting their function, which is often linked to driving a specific signaling pathway required by the target disease. For a covalent drug that is intended to react with its target selectively, typically rapid second-order kinetics of binding are needed. The parameter most commonly considered is the effective second-order rate constant for forming the covalent enzyme inhibitor complex, E–I. This rate constant is described by k_{inact}/K_i, a parameter that reflects both the speed of formation of E–I from the E•I non-covalent complex (k_{inact}) as well as the binding affinity of the enzyme for the inhibitor (K_i). Typically, for an effective covalent drug, rate of inhibition must compete against

Long, et al.

metabolic processes, clearance and labeling of unintended proteins (off-target effects). Good drugs accordingly have k_{inact}/K_i above 10^4 M^{-1}s^{-1} (49), although most modern covalent kinase inhibitor anti-cancer drugs have k_{inact}/K_i's approaching close to the diffusion-controlled limit, which is around 10^8 M^{-1}s^{-1}. It is typically assumed that as the rate of the on-target interaction(s) increase(s), the specificity of the inhibitor will increase. However, there is a large amount of proteomic evidence showing that off-target labeling of reactive drugs, even those with high second-order rates of reaction with their chosen targets, can be significant. For drugs like Ibrutinib, these unintended targets can contribute to negative effects in human patients. Thus, even "highly selective" binders have to run the gauntlet when it comes to avoiding inadvertent labeling. Nevertheless, rapid on-target labeling is still an "ideal" property of covalent drugs, especially when it comes at the expense of off-target reactions.

We and others have shown that even ostensibly simple biologically-derived electrophiles can regulate cellular information flow in a similar way to covalent drugs. Because of their simple size and often higher implicit reactivity than approved drugs as well as their rapid clearance, reactive lipid-derived electrophiles must perform a delicate balancing act between rapid on-target labeling, minimizing off-target labeling, and ensuring that on-target labeling can compete against clearance (29, 40). Although there are many constraints on this system, we and others have provided evidence that several critical signaling processes, including ubiquitination and phosphorylation, are subject to electrophile regulation. The time period in which a labeling reaction must occur to be able to trigger cell signaling is difficult to define, and likely depends on what downstream processes are being intercepted. For signaling modes involving protein upregulation, time frames likely have to be considerably faster than the time taken for downstream signaling to occur, such as transcriptional upregulation, which takes a few hours. For signaling modes like apoptosis, which involves activation of pre-existing proteins, the process likely has to be faster (44). However, duration or half-life of reactive electrophiles may well be rate-limiting for signaling, as reactive small molecules appear to show metabolic-conversion and clearance mechanisms much faster than approved drugs (as may be expected given that drugs are usually optimized to be metabolically stable) (29). The

Long, et al.

half-lives of electrophiles like α,β-unsaturated carbonyls are likely no more than minutes in most cells (29,

40). Assuming that the interaction between the α,β-unsaturated carbonyl and a specific enzyme is

biologically relevant (i.e., leads to a significant selective modulation of a specific pathway), it will also

have to happen in such a way that there is selective buildup of the modified protein under conditions where

other proteins are ostensibly unaffected. But as the covalent bonding energy is so dominant over non-

covalent interactions, at equilibrium every protein cysteine and small-molecule thiol would be modified,

assuming there were enough electrophile-molecules to go around! Fortunately, the rate of reaction of free

cysteine (and by extension, most unhindered, unactivated protein cysteines) with β-substituted α,β-

unsaturated aldehydes, like the prototypical lipid derived electrophile, 4-hydroxynonenal (HNE), is slow,

around 1 $M^{-1}s^{-1}$ (40), i.e., considerably below the theoretical limit. Furthermore, with concentrations of β-

substituted α,β-unsaturated aldehyde and enzyme set to 20 μM, the half-life for labeling of a cysteine of

standard reactivity (regardless of the cysteine being incorporated into protein or being free) is

approximately 10 h. For comparison, with a rate constant of 10^6 $M^{-1}s^{-1}$ [similar to the k_{inact}/K_i for the

approved covalent kinase drugs afatinib and neratinib (49)] at the same concentration, the half-life becomes

0.05 s (note: 20 μM is a relatively high concentration for an enzyme, and these bimolecular half-lives will

increase in length as the enzyme concentration decreases, as shown in **Figure 2B**). Our considerations may

be mitigated somewhat by HNE concentration, which has been proposed to be able to reach up to 500 μM

(40). However, we stress that this would require *sustained* elevation of such concentrations (which are

known to be toxic, at least when administered from outside of the cell) and hence are unlikely to be relevant.

In any case, the pseudo first-order half-life for protein labeling under these conditions is ~25 minutes,

significantly longer than the half-life of HNE in most cells (~2 minutes).

It is commonly argued in the literature that tweaking the acidity of a thiol such that it is exclusively

a thiolate at neutral pH can cause such an increase in reactivity that these "acidic" or "hot" thiols will be

inherently able to act as "biologically relevant" nucleophiles. Unfortunately, the complete deprotonation of

small-molecule cysteine (and by extension, a non-activated cysteine on a protein) can increase the second-

Long, et al.

order rate constant by *maximally* 20-fold relative to the equilibrium mixture of cysteine and cysteine thiolate that exists under close to biological conditions (40). In absolute terms, for reaction with a signaling molecule like HNE reacting with a cysteine, this represents a jump of second-order rate constant from around 1 $M^{-1}s^{-1}$ to 20 $M^{-1}s^{-1}$. As we see from **Figure 2B**, these second-order rate constants are unlikely to enable a significant amount of product to build up, even if the proteins *and* the electrophile can *each* reach 20 μM. Indeed, if we assume that for a labeling process to be "biologically relevant" kinetics should occur around the rate of a typical enzyme reaction, one can see that a threshold of 1000 $M^{-1}s^{-1}$ is needed to be meaningful (4). As we noted above, HNE is also not particularly stable in cells (29), and hence the time frame to react under signaling conditions, may well be short (60), perhaps minutes. This constraint may make the rate threshold for HNE signaling *higher* than some enzymatic processes. In fact real physical processes also make matters worse for the 'acidic thiol' hypothesis. This is because as pK_a decreases, the nucleophilicity of the thiolate consequentially decreases, all other factors being equal, *rendering a 20-fold increase in rate not readily attainable by pK$_a$ decreases alone.* Thus deprotonation alone is also not likely to lead to signaling output, and further, given that the total cellular protein-cysteine content is around 10 mM (40), such minimal rate enhancements are unlikely to lead to either significant occupancy, or selective buildup of the modified cysteine over the other background cysteines in the cell in sufficient time to relay a signal. Throughout this mini-review, we will evaluate this hypothesis in more detail, and ultimately conclude that it is likely correct. However, we will also give some evidence that indicates that thiolate formation could well play some role in this process.

We have discussed previously different ways in which a specific protein could elevate the reactivity of one (or some) of its cysteine(s) above the maximal rate of reaction possible for a common cysteine thiolate (29, 40). In general, these methods involve a rapid non-covalent pre-binding interaction leading to positioning and/or activation of the electrophile prior to covalent bond formation (both of which can elevate reactivity many folds above background, for instance 4-oxononenal (ONE) reacts with cysteine with a rate constant of around 150 $M^{-1}s^{-1}$; furthermore, numerous organocatalytic Michael reactions to enals proceed

Long, et al.

through preformation of an iminium ion from the enal, speeding up Michael addition due to the fact that the unsaturated iminium ion has a lower LUMO than the enal (37).

Mimicking these processes by activating the electrophile via LUMO lowering, positioning the electrophile in an orientation predisposed to undergo attack by a pendant thiol, or ideally stabilizing the intermediate formed along the Michael adduction pathway (or by raising the HOMO of the nucleophile) certainly would lead to increases in reactivity much higher than what occurs through cysteine deprotonation alone (31), as is quite clear from both examples of enzymes, and the specific examples above.

Similar to the proposals of Michaelis and Menten or the lock and key/induced fit models, this logic would predict that proteins with elevated reaction rates, i.e., that are privileged sensors of electrophiles, should pre-bind their targets prior to reaction. There is evidence for HNE-binding non-covalently to its molecular targets: for instance, molecular modeling studies have indicated that HNE may have an affinity approximately 16 kJ mol^{-1} (K_d in the mM range) for its postulated target, Src kinase (22). It is worth noting that in the optimized binding mode published, the enone was in a high energy staggered conformation that is unable to undergo conjugate addition, and the proposed nucleophilic residue was postulated to be involved in the recognition between HNE and the protein through H-bonding to the OH. Only time will tell to what extent such models are realistic. Certainly, more experimental work needs to be done to understand the non-covalent aspects of protein–electrophile interaction. Experiments such as investigating blocking of modification of HNEylation by HNE-analogs that cannot undergo Michael adduction would be a fair start. But such experiments—unless carried out on proteins that are themselves kinetically highly-reactive and under conditions where binding sites are not saturated by HNE—could lead to confounding results.

There is another simple way to resolve whether something other than the "brute force" of chemical reactivity is responsible for reaction between α,β-unsaturated ketones and enzymes. One could just measure the rate of the second-order reaction between a specific protein and HNE to test if the rate constant exceeds what is possible by a common thiolate reagent, like glutathione. The higher the rate of reaction above the theoretical maximum, the more likely prebinding/enzyme behavior is at play. The enzyme glutathione-S-

Long, et al.

transferase (GST) that catalyzes conjugate addition of glutathione to Michael acceptors can give rate accelerations around 3-5 orders of magnitude above background rates (29, 40). This enzyme uses prebinding and other "enzyme tricks" to promote electrophile adduction. These tricks include: promoting correct substrate alignment, stabilization of the transition state for electrophile adduction, stabilization of the intermediate, and leaving group stabilization (where applicable) (12). Although glutathione deprotonation is also considered to be part of GST mechanism, we need to bear in mind the maximal acceleration for deprotonation alone is minimal compared to an acceleration of 5-orders of magnitude. Unfortunately, second-order rates of enzyme labeling by electrophiles have proven to be quite difficult to assess quantitatively in vitro; cell-based experiments cannot lead to directly-extractable kinetic parameters. However, there is a corollary to the question of kinetics that can be answered; that is, how selective are different electrophiles for different proteins? Should certain protein cysteines display reactivities that are elevated above that which is possible for a canonical protein/small-molecule cysteine, at least some of these reactive cysteines would be expected to show selectivity towards a specific subset of reactive small molecules. We stress that this trait is not absolutely necessary as some GSTs, for instance, are not particularly substrate discriminating. However, in cases of proteins that rapidly engage HNE and are primed to undertake biologically-relevant cell signaling roles, which we have termed "privileged first responders" (29), such selectivity would be expected because non-covalent pre-binding of the electrophile is likely employed by a subset of these proteins to promote fast kinetics. Such an interaction could in effect act as a filter priming the protein to favor a type, or in an extreme case, one specific electrophile, over others. Indeed, this question obviates thiol-specific tricks, such as thiol desolvation, that could potentially lead to an overall hyperactivity that is faster than expected than a thiolate in solution (40). These assays would be more poignant should the assayed cysteines be shown to have highly-elevated reactivity compared to the global cysteine pool.

Mass spectrometry datasets indicate some proteins show privileged reactivity with electrophiles

Long, et al.

For many years there was very little evidence showing that some proteins have elevated reaction rates with α,β-unsaturated ketones. However, there are now several groups investigating electrophile labeling that have provided good evidence for such phenomena. We will here review the work showing that proteins are highly discriminatory for various electrophiles. In many instances this question has been approached from various different angles and similar results have been obtained. We will furthermore, discuss ways in which confounding results can occur.

The Cravatt laboratory has published numerous data sets that are highly useful. These are mass spectrometry data, profiling reactivity of a pool of protein cysteines towards various electrophiles. The main method used is a venerable chemical biology approach. However, several implicit limitations for the method are noteworthy and have been discussed in detail elsewhere (29, 31, 35, 40), such as: the method being built on loss-of-signal (and hence prone to artifactual readouts, especially when reactive ligands are employed); assumption that the cell is a homogeneous entity; and only detecting a small fraction [at most ~4% (31)] of the total cellular cysteine content. Encouragingly for practitioners of this method, several variants have been recently published by other laboratories that have begun to address some of these issues, including, e.g., derivatization of acyl group (10) and photocaged electrophiles (1). The data sets from the Cravatt laboratory and those from applying the acyl-group derivatization approach for the most part agree, arguing at least within the profiled data set, comparisons and conclusions are valid.

After some thought-provoking work showing that iodoacetamide-derived ligands have a kinetic preference for reaction with a subset of protein cysteines, out of a profiled data set of around 1000 protein cysteines (58), the Cravatt lab's attention turned to investigating the interaction preference of natural electrophiles (56). 4-Hydroxynonenal (HNE) and 15-Deoxy-Delta(12,14)-prostaglandin J2 (15d-PG J2) were chosen. From our perspective, this choice was very insightful. The principal reactive motif is, in each instance, an ene-carbonyl. HNE possesses an enal; the prostaglandin possesses two enones, one being endocyclic to the 5-membered ring, the other being exocyclic and conjugated to both the ketone and another olefin (**Figure 3A**). Given the subtle differences in terms of electronic properties of the two structures, it is

13

Long, et al.

difficult to predict which of the two molecules would be more reactive. Although the rate constants for reaction with model nucleophiles has not been systematically disclosed, both are known to be depleted by glutathione upon addition to cells (3), and both react with serum albumen in vitro (28). However, these two molecules have similar reactive chemotypes and prostaglandin possesses enones in both the *s*-cis and *s*-trans configurations, that can mimic the two reactive, low-energy conformations of HNE (**Figure 3A**). Thus, one may naïvely predict that there would at least be a "decent" level of overlap between the adducted proteins in cells treated with these two small molecules, if only reactivity and olefin orientation were considered. Conversely, the connectivity and topology of the two molecules is quite distinct. HNE possesses a pendant hydroxyl group, which is not present in prostaglandin. Furthermore, we know, for instance, from Von Hippel Lindau (7) and studies from Src (22), that the hydroxyl function within HNE may be important for binding in some instances. On the other hand, the ring core of prostaglandin and its pendant carboxylic acid are standout differences in the carbon-carbon connectivity relative to HNE.

The profiling experiments showed that fast-reacting proteins for HNE and 15d-PG J2 did not overlap at all (56). Indeed, many of the proteins that were efficiently labeled by HNE did not react appreciably with 15d-PG J2, and vice versa (**Figure 3B**). Unfortunately, the kinetics of binding/labeling of the electrophiles for their respective targets were not calculated. We thus strictly speaking have no clear idea how reactive the proteins labeled selectively are, or how much of these rate enhancements could be explained by thiol deprotonation alone; however, given the duration of the experiment (1 h), and that permeation of reactive electrophiles into cells is relatively slow, it seems unlikely to us that deprotonation alone is sufficient to explain the data. [This is because, based on the data discussed above, with deprotonation alone, the absolute maximal second-order rate achievable would be 20 $M^{-1}s^{-1}$; and at this rate, reaction half-life would be too long (~100 minutes in the second-order limit; or 30 minutes assuming a first-order reaction in HNE at a constant concentration of 20 μM) for 4–20 μM protein/electrophile concentrations and incompatible with experimental timescale]. Of course, the use of cells to probe these labeling interactions may complicate interpretation, and indeed, no direct in vitro comparisons between

Long, et al.

labeling of sensitive proteins by the different electrophiles were shown. However, some hits were inhibited upon addition of the electrophiles used, at least validating the general points of the assay.

The Wang lab developed a different version of activity-based protein profiling (ABPP) originally established by the Cravatt laboratory (**Figure 4A**). The Wang lab's method makes use of the carbonyl function integrated into the protein post Michael adduction to HNE, as a handle to identify HNE-modified proteins (**Figure 4B**) (10). This method detected enhanced labeling of several proteins, some of which were detected by the Cravatt lab in their previous competitive iodoacetamide-based ABPP experiments (56), e.g., ZAK (C22), HMOX2 (C282) and RTN4 (C1001) (10). Repetition of these experiments using acrolein as an electrophile showed a different set of sensitive protein cysteine residues and indeed only overall 30% of sites were common between HNE and acrolein. Intriguingly, several proteins showed multiple acrolein modifications, e.g., P53 (6 acrolein-adduction events) [not present in HNE set], AHNAK (5 acrolein-adduction events) [not main sensor of HNE], ANLN (4 acrolein-adduction events) [not present in HNE set], EEF2 (4 acrolein-adduction events) [not a sensor of HNE] (9).

One of the clear issues with the above approaches is that modified proteins were identified under conditions where excess of electrophile was administered to the cells for periods likely longer than "signaling conditions" persist. The Wang lab proceeded to investigate proteins affected by electrophile modification under ferroptosis (a process known to involve lipid oxidation) (10). Because the Wang method is based on enrichment of specific, electrophile-modified peptides, the method was able to be applied not only to target ID, but also to identifying specific modifications [in this case, HNE and 4-oxononenal (ONE)] (18). Once again, such comparisons are instructive because ONE is 100-fold kinetically more reactive than HNE to cysteines, even though ONE and HNE have very similar connectivities. Thus, one would expect the targets of HNE and ONE to overlap. Intriguingly, in spite of the similarities between the two electrophiles, the Wang lab identified several proteins that were labeled selectively by HNE and ONE, again indicating that there are some very efficient and specific protein sensors of electrophiles. We recognize: (1) that it is possible that HNE adducts can be converted to ONE adducts post adduction and

Long, et al.

vice versa (even during downstream processing), and (2) we are cognizant of the fact that many HNE adducts do not give free aldehydes for conjugation as required to be detected by this method, but ONE modifications do not have the ability to form cyclic structures and hence are more likely to be detected by the method; and (3) we do not strictly know if there are different compartments of formation for ONE and HNE; given the differences in reactivities, however, we can speculate that the diffusion distances of the two molecules are quite different that could lead to different protein labeling spectra. Nevertheless, these data are far from expected and further indicate that proteins may well be very discerning when it comes to electrophile reactivity, and furthermore can countermand reactivity (ostensibly by 100's of fold) even between electrophiles of seemingly similar structures.

Data from REX technologies are also consistent with protein-ligand recognition

We ourselves have focused on protein-specific modification events. In order to achieve this goal, we developed the first protein-specific electrophile modification strategy that functions in cells and whole organisms, called targetable reactive electrophiles and oxidants (T-REX) (**Figure 5A**) (14, 27, 30, 32, 33, 38, 39, 42, 51, 54, 61). These experiments have told a similar story to that of the Cravatt and Wang labs' experiments, but from the perspective of protein-specific modifications by specific electrophiles. We stress that because of the way our experiments are set up, many of the caveats of the Cravatt and Wang experiments are assuaged by T-REX. Thus, the data sets are in fact very complementary. As recently discussed at length (29, 39, 40, 42), T-REX does have its own caveats, which are principally limited to the fact that an ectopic fusion protein is required to perform T-REX, although careful implementation of built-in technical T-REX controls (42), and the use of 'non-fusing protein' control (40, 42) and hypomorphic mutants (32, 61), altogether allow any off-target responses/potential mis-interpretations to be ruled out rigorously in T-REX-experiments. Notably, the combination of T-REX and G-REX (42), (a system to release low concentrations of a specific reactive electrophile unfettered into the cellular milieu to identify

16

Long, et al.

the best native sensors) (29), has assuaged the concerns underlying overexpression of fusion proteins quite considerably.

We first compared how different electrophiles were able to trigger downstream signaling upon Keap1 labeling (27, 30, 38, 39). Keap1 is one of the master regulators of the linchpin antioxidant response (AR) regulator, Nrf2 (19). Keap1 typically binds Nrf2, but upon electrophile adduction, this binding is impaired, leading to an increase in Nrf2 and upregulation in AR, which can typically be read out by reporter assays, qPCR or western blot (39). Expecting that Keap1 may prove differentially susceptible to electrophiles under protein-specific adduction, we compared the extent of labeling of Keap1 and downstream AR triggered under Keap1-specific modification in living cells, to unstimulated levels and bolus dosing levels for a number of α,β-unsaturated electrophiles (27). Perhaps, disappointingly, both the extent of electrophile labeling and the overall AR upregulated was similar for all electrophiles (aside for electrophiles that transpired to be incompatible with the T-REX strategy). One surprise was that AR upregulation upon T-REX did not differ markedly from AR upregulation observed under bolus dosing (27, 38), implying that both positive and negative regulation of AR may be incurred upon bolus electrophile treatment conditions. We found such antagonistic behavior for numerous other systems in subsequent studies (30).

We elected to investigate the molecular underpinnings of Keap1-specific versus global AR more closely by identifying residue occupancy under the different conditions. At this level of scrutiny, the outputs were quite different. T-REX, delivering a cyclohexenyl electrophile (CHE, **Figure 5B**) in cells, hit C613 and C489; whereas upon bolus electrophile dosing of cells, CHE hit C513/C518 and C273 (**Figure 6A**) (27). On the other hand, for HNE, T-REX in cells labelled C513 and C518; bolus HNE treatment (100 µM) of cells labelled C23, C226, C273, and C368. But only C23 and C368 were hit when cells were treated with 25 µM HNE (38). In vitro, C77, C151, C226, C273, C319, and C368 were hit by HNE upon bolus dosing (38). Just as in the data from ABPP above, the mass spectrometry analyses cannot give a complete picture of the different residues modified, so we recognize that it is possible that some of these conclusions are

17

Long, et al.

biased by ionization propensities, solubilities/stabilities of the product peptides or resolution of the mass spectrometer. Nonetheless, these data suggest that different cysteines within Keap1 have different preferences for different electrophiles.

Pleasingly, these data agree with other reports that different reactive electrophiles and oxidants interact with specific subsets of Keap1 cysteines (8). However, in our case the electrophiles are quite similar and in the case of T-REX are delivered from a similar position relative to Keap1; the electrophiles are also delivered at the same concentrations, under conditions where off-target effects and negative impacts, permeation and other issues are side-stepped. Notably, irrespective of whether HaloTag is fused to N- or C-terminus of Keap1 protein, the same residue is modified (C613) upon T-REX-mediated CHE-targeting of Keap1 in live cells (39). It is also critical to bear in mind that labeling efficiency of Keap1 by CHE and HNE under T-REX conditions were similar (27). Thus, the T-REX data imply that reactive cysteines in Keap1 are highly sensitive to chemical change.

We also investigated how different isoforms of different proteins sense HNE. After a medium-throughput screen, we identified Akt, an enzyme which has three isoforms as an interesting model system for further study (32). These proteins share >75 % similarity overall, and the pairwise similarity comparisons are shown in **Figure 6B**. It quickly became apparent that Akt1 is insensitive to HNE. On the other hand, Akt2 is weakly sensitive to HNE, but Akt3 is one of the most sensitive HNE sensors we have characterized. Mass spectrometry showed that a specific residue (C119) that is unique to Akt3 sensed HNE. Mutation of this residue to serine (C119S) ablated Akt3 HNE sensing. Intriguingly, the sensor residue, C119 in Akt3, lies on a flexible linker region between the two principal domains of Akt, the PH and kinase domains (**Figure 6B**). This linker is highly divergent between the three isoforms. Furthermore, Akt2 also houses a reactive cysteine (C124) in this linker region, which is believed to react with peroxide (57). Although we have shown that the primary structures surrounding these two sensor cysteines [Akt3(C119) and Akt2(C124)] are different, the question remains unanswered whether or not primary sequence differences alone are sufficient to cause such a divergence in selectivity. However, it is likely reactivity

Long, et al.

induced by primary sequence changes are promoted by pK_a decreases, increasing the amount of thiol anion formed. Such changes favor interaction with both oxidants and electrophiles similarly, and the margin for improvement for cysteine is relatively small (~20-fold) (40) in each case. Thus, given that Akt3 was at least 4-fold more reactive to HNE than Akt2, and Akt2 is significantly more reactive to peroxide than Akt3, it seems unlikely that *only* pK_a effects are at work. Of course, the background labeling of Akt2 by HNE could certainly represent a contribution from increased thiol anion, although the comparison in this instance is Akt1 that does not possess a cysteine and hence is not a fair comparison. In other work, we have mutated a putative general base that is spatially close to the privileged cysteine in HSPB7 (51). This led to a significant decrease in HNE sensing, arguing that deprotonation may be part of the story, likely if other important aspects of activation are also present.

Proteins may show enantioselectivity, although this is far from clear

Cellular studies comparing (*R*)- and (*S*)-HNE have indicated that the two enantiomers have different biological properties (40). One study, assessing the effects of the two compounds in cultured mouse hepatocytes found that (S)-HNE was better able to initiate apoptosis and several pro-apoptotic signaling pathways than its enantiomer. However, other studies in colorectal cells have found very little difference between the two enantiomers (59). Of course, such different behaviors could reflect different targets expressed in the two cells, or different metabolizing enzymes present. Nevertheless, such data are interesting, if somewhat preliminary.

Enantioselective modification of different proteins by HNE has been reported. One of the best examples is glyceraldehyde-3-phosphate dehydrogenase (GAPDH). This is inactivated over 3-times faster by (*S*)-HNE than (*R*)-HNE (20). However, it should be ceded that these rates of reaction are quite slow (6 $M^{-1}s^{-1}$ vs. 2 $M^{-1}s^{-1}$) and the difference in energy for these two rate constants is not much greater than the

Long, et al.

thermal energy (~ 3 kJ mol^{-1} [1] vs. ~ 2.5 kJ mol^{-1}, assuming T = 298K). For instance, this pales into insignificance when we consider enantiospecificity of approved drugs. (*S*)-ibuprofen is over 100-fold more potent an inhibitor of cyclooxygenase than its antipode (corresponding to a difference in energy of over 11 kJ mol^{-1} for the different bound states).

One may predict that as the reactivity of different cysteines increases (i.e., the rate of reaction is increased above the background rate, and hence a larger interaction between electrophile and protein may be expected to be required), the enantioselectivity increases. Unfortunately, this is certainly not always true. Thioredoxin possesses two HNE-reactive cysteines, C73 and C32. C73 displays faster labeling kinetics than C32. However, C73 displays little preference for different enantiomers. On the other hand C32 displays significant preference for (*R*)-HNE (55). Although, since no kinetic parameters were discussed in the paper and it is possible that neither cysteine is particularly reactive, these data do certainly give a cautionary tale that study of electrophile interactions with proteins is likely to be complex and throw up some confounding information. However, it is worthwhile to ask if we have chosen the correct proteins for these in vitro assays. Indeed, most of the proteins we discuss in this section, and the larger list from the literature of in vitro studied protein modifications have been carried out on proteins that are not detected as represented as HNE selective in any of the above HNE profiling assays (40). It is possible that these specific proteins have little to teach us about how HNE interacts with privileged sensor proteins as they do not have the mechanisms that lead to meaningfully enhanced rates of HNE adduction. Hopefully more modern methods can be applied to these problems, and more downstream analysis can be used to help shed light on these problems.

Conclusion

---

[1] Value derived using: $\ln(k_2/k_1) = (\Delta E)/RT$ at 298K.

Long, et al.

We are a long way from understanding how lipid-derived electrophiles (LDEs) interact semi-selectively with proteins. Our tools to study these processes are just on the verge of having the ability to scrutinize these systems well enough to ask meaningful questions that need to be answered. However, the data from the laboratories we have discussed above are starting to show that there is huge diversity in terms of lipid sensing ability and indeed selectivity among reactive-lipid-sensing proteins. Indeed, there is evidence that proteins not only exhibit kinetic selectivity for specific reactive ligands, but also that these proteins can buck the trend of typical reactivity. The signs are encouragingly pointing to the fact that genuine enzyme-like sensing of electrophiles may be occurring. As we have also argued, if sensing occurs between specific proteins and specific electrophiles with defined molecular architectures, i.e., through interaction with/occupancy of a specific binding pocket, this opens the doors for gain-of-function events, amplification of downstream signaling events and other properties that have been proposed to occur upon electrophile modification.

As we have argued previously, to really understand these processes, there is really a need for more quantitative and rigorous analysis of these data. It is apparent that gleaning this level of understanding will require a more diverse skillset than is currently applied to most papers dealing with reactive molecule signaling. As usual, resolution will ultimately need raising of the bar for the way the field in general goes about characterizing electrophile-protein interactions. However, in our opinion, this extra effort will be worth it. We lay out our thoughts below.

From our introduction, it is clear that understanding kinetics of labeling is essential. Cell-based experiments unfortunately do not offer anywhere near precise enough measurements of sensing to give us a good platform to rank and understand sensing quantitatively. We have for instance measured externally-administered HNE labeling of the soluble cellular proteome. There was a significant lag phase (around 5-10 minutes) between identification of labeling and HNE dosing. However, when HNE was released from inside the cell, maximal labeling of the proteome occurred in no more than 3 minutes (61)(**Figure 7**). This labeling was similar to what was observed following bolus HNE treatment (30-mins). Thus, permeation is

Long, et al.

almost totally rate-limiting in this situation, although admittedly the labeled proteomes were not directly compared. Indeed, it is known that under bolus dosing conditions, the cell experiences a gradient of reactive molecules (29), with concentrations of the reactive molecule being highest at the membrane and falling off towards the nucleus. Furthermore, it is believed that for this gradient to be overcome some of the detoxification machinery needs to be compromised. Our experiments indicate that intracellular release of the electrophile may not be so subject to such issues. Regardless, it is important to question just what is leading to selectivity under bolus dosing, is it really genuine HNE sensitivity or is it proximity, diffusivity, stability or (an)other parameter(s) that dominate(s) over kinetics and masks the key sensors? It is also worth noting that it is unlikely that external bolus dosing models particularly model well endogenous signaling conditions, that likely arrive from intracellular cues/electrophile release, unless significant stress is levied on the system. Thus, more time should be given to careful, and traditional characterization of labeling, or inhibition (k_{inact}/K_i) kinetics of identified proteins by low concentration of electrophile on purified isolated sensor proteins. Such experiments should be performed at the expense of performing time-independent experiments with excess of electrophile and protein, which give little indication of kinetics or ability of proteins to react at a reasonable (or any) rate.

Of course, such considerations do not explain why different proteins show differential reactivity/susceptibilities to subtly different electrophiles. It is indeed this unexpected observation which is particularly exciting. But again, in vitro validation and quantitative assessment of such parameters need to be carried out, as confounding factors such as differential diffusion, distribution, or potentially metabolism, could confound our interrogations. Conversely, as we have suggested elsewhere, T-REX may be able to better serve this purpose. This is because in vitro analysis is complex and T-REX is ostensibly an unbiased measure of a protein's sensitivity to electrophiles.

One issue where we have yet to really cover decent ground is enantioselectivity. Such a gap in our understanding is significant, as it is enantioselectivity, and the "natural" resolution of racemic products that is most indicative of biological processes. Hence, one could imagine that enantioselective reaction of

Long, et al.

different proteins with an electrophile would be most indicative of an "enzyme-like" recognition process occurring. Unfortunately, this area is also not immune to artifacts caused by metabolism. For instance, rat GSTα4-4 metabolizes (*S*)-HNE 4-times faster at V_{max} than (*R*)-HNE; the K_m was 3-fold higher for (*R*)-HNE than the (*S*)-enantiomer. This corresponds to a 12-fold increase in second-order rate of reaction for the (*S*)-enantiomer than for the (*R*)-enantiomer (20). To further confound matters, it seems that although many GSTs show S-selectivity, other GST isoforms may show the opposite sensitivity and other HNE detoxification processes such as NAD^+ dependent oxidation may show (*R*)-HNE metabolism preference. Thus administration of single enantiomers of HNE could lead to different enantiomers being enriched within different locales, etc (53).

We thus hope the field will not continue to neglect the quintessential aspects of biochemistry and recognition in its continued efforts to understand redox signaling. We feel that careful, and controlled experiments, such as those offered by several techniques available these days, can be used to interrogate electrophile sensing more accurately, and to generate data sets that are more transposable between laboratories. It is our feeling that the data hint that electrophile sensing is incredibly nuanced and we hope that further experimental proof can be provided in the near future.

Long, et al.

Author Disclosure Statement.

No competing financial interests exist.

Long, et al.

List of abbreviations.

15d-PG J2: 15-deoxy-delta(12,14)-prostaglandin J2

ABPP: activity-based protein profiling

AR: antioxidant response

CHE: cyclohexenone-derived LDE

GAPDH: glyceraldehyde-3-phosphate dehydrogenase

G-REX: genome-wide profiling of sensors responsive to reactive electrophiles and oxidants

GST: glutathione-S-transferase

HNE: 4-hydroxynonenal

IA: iodoacetamide

isoTOP-ABPP: isotopic tandem orthogonal proteolysis–ABPP

LDE: lipid-derived electrophile

m-APA: *m*-aminophenylacetylene

ONE: 4-oxononenal

POI: protein of interest

TEV: tobacco etch virus protease

T-REX: targetable reactive electrophiles and oxidants

Long, et al.

References.

1. Abo M, Bak DW, and Weerapana E. Optimization of Caged Electrophiles for Improved Monitoring of Cysteine Reactivity in Living Cells. *ChemBioChem* 18: 81–84, 2017.

2. Athappilly FK and Hendrickson WA. Crystallographic analysis of the pH-dependent binding of iminobiotin by streptavidin. *Protein Sci* 6: 1338–1342, 1997.

3. Balogh LM and Atkins WM. Interactions of glutathione transferases with 4-hydroxynonenal. *Drug Metab Rev* 43: 165–178, 2011.

4. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, and Milo R. The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* 50: 4402–4410, 2011.

5. Barlow R. Enantiomers: how valid is Pfeiffer's rule? *Trends Pharmacol Sci* 11: 148–150, 1990.

6. Bentley R. The Nose as a Stereochemist. Enantiomers and Odor. *Chem Rev* 106: 4099–4112, 2006.

7. Bondeson DP, Mares A, Smith IED, Ko E, Campos S, Miah AH, Mulholland KE, Routly N, Buckley DL, Gustafson JL, Zinn N, Grandi P, Shimamura S, Bergamini G, Faelth-Savitski M, Bantscheff M, Cox C, Gordon DA, Willard RR, Flanagan JJ, Casillas LN, Votta BJ, den Besten W, Famm K, Kruidenier L, Carter PS, Harling JD, Churcher I, and Crews CM. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol* 11: 611–617, 2015.

8. Bryan HK, Olayanju A, Goldring CE, and Park BK. The Nrf2 cell defence pathway: Keap1-dependent and -independent mechanisms of regulation. *Biochem Pharmacol* 85: 705–717, 2013.

9. Chen Y, Liu Y, Hou X, Ye Z, and Wang C. Quantitative and Site-Specific Chemoproteomic Profiling of Targets of Acrolein. *Chem Res Toxicol*, 2019.

10. Chen Y, Liu Y, Lan T, Qin W, Zhu Y, Qin K, Gao J, Wang H, Hou X, Chen N, Friedmann Angeli JP, Conrad M, and Wang C. Quantitative Profiling of Protein Carbonylations in Ferroptosis by an Aniline-Derived Probe. *J Am Chem Soc* 140: 4712–4720, 2018.

11. Cornish-Bowden A. The origins of enzyme kinetics. *FEBS Lett* 587: 2725–2730, 2013.

Long, et al.

12. Deponte M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim Biophys Acta BBA - Gen Subj* 1830: 3217–3266, 2013.

13. Desmet GB, D'hooge DR, Omurtag PS, Espeel P, Marin GB, Du Prez FE, and Reyniers M-F. Quantitative First-Principles Kinetic Modeling of the Aza-Michael Addition to Acrylates in Polar Aprotic Solvents. *J Org Chem* 81: 12291–12302, 2016.

14. Fang X, Fu Y, Long MJC, Haegele JA, Ge EJ, Parvez S, and Aye Y. Temporally controlled targeting of 4-hydroxynonenal to specific proteins in living cells. *J Am Chem Soc* 135: 14496–14499, 2013.

15. Flack HD. Louis Pasteur's discovery of molecular chirality and spontaneous resolution in 1848, together with a complete review of his crystallographic and chemical work. *Acta Cryst Sect A* 65: 371–389, 2009.

16. Glazer AM and Stadnicka K. On the origin of optical activity in crystal structures. *J Appl Crystallogr* 19: 108–122, 1986.

17. Green NM. Thermodynamics of the binding of biotin and some analogues by avidin. *Biochem J* 101: 774–780, 1966.

18. Groeger AL and Freeman BA. Signaling Actions of Electrophiles: Anti-inflammatory Therapeutic Candidates. *Mol Interv* 10: 39–50, 2010.

19. Hayes JD and Dinkova-Kostova AT. The Nrf2 regulatory network provides an interface between redox and intermediary metabolism. *Trends Biochem Sci* 39: 199–218, 2014.

20. Hiratsuka A, Hirose K, Saito H, and Watabe T. 4-Hydroxy-2(E)-nonenal enantiomers: (S)-selective inactivation of glyceraldehyde-3-phosphate dehydrogenase and detoxification by rat glutathione S-transferase A4-4. *Biochem J* 349 Pt 3: 729–735, 2000.

21. Jackson PA, Widen JC, Harki DA, and Brummond KM. Covalent Modifiers: A Chemical Perspective on the Reactivity of α,β-Unsaturated Carbonyls with Thiols via Hetero-Michael Addition Reactions. *J Med Chem* 60: 839–885, 2017.

Long, et al.

22. Jang EJ, Jeong HO, Park D, Kim DH, Choi YJ, Chung KW, Park MH, Yu BP, and Chung HY. Src Tyrosine Kinase Activation by 4-Hydroxynonenal Upregulates p38, ERK/AP-1 Signaling and COX-2 Expression in YPEN-1 Cells. *PLOS ONE* 10: e0129244, 2015.

23. Jin Z. Muscarine, imidazole, oxazole, and thiazole alkaloids. *Nat Prod Rep* 28: 1143–1191, 2011.

24. Kuntz ID, Chen K, Sharp KA, and Kollman PA. The maximal affinity of ligands. *Proc Natl Acad Sci U S A* 96: 9997–10002, 1999.

25. Kwan EE and Evans DA. Intermolecular Michael Reactions: A Computational Investigation. *Org Lett* 12: 5124–5127, 2010.

26. Lewandowicz A, Ringia EAT, Ting L-M, Kim K, Tyler PC, Evans GB, Zubkova OV, Mee S, Painter GF, Lenz DH, Furneaux RH, and Schramm VL. Energetic Mapping of Transition State Analogue Interactions with Human and *Plasmodium falciparum* Purine Nucleoside Phosphorylases. *J Biol Chem* 280: 30320–30328, 2005.

27. Lin H-Y, Haegele JA, Disare MT, Lin Q, and Aye Y. A generalizable platform for interrogating target- and signal-specific consequences of electrophilic modifications in redox-dependent cell signaling. *J Am Chem Soc* 137: 6232–6244, 2015.

28. Liu Q, Simpson DC, and Gronert S. The reactivity of human serum albumin toward trans-4-hydroxy-2-nonenal: Reactions of HSA with HNE. *J Mass Spectrom* 47: 411–424, 2012.

29. Liu X, Long MJC, and Aye Y. Proteomics and Beyond: Cell Decision-Making Shaped by Reactive Electrophiles. *Trends Biochem Sci* 44: 75–89, 2019.

30. Long MJ, Lin H-Y, Parvez S, Zhao Y, Poganik JR, Huang P, and Aye Y. β-TrCP1 Is a Vacillatory Regulator of Wnt Signaling. *Cell Chem Biol* 24: 944-957.e7, 2017.

31. Long MJC and Aye Y. Privileged Electrophile Sensors: A Resource for Covalent Drug Development. *Cell Chem Biol* 24: 787–800, 2017.

32. Long MJC, Parvez S, Zhao Y, Surya SL, Wang Y, Zhang S, and Aye Y. Akt3 is a privileged first responder in isozyme-specific electrophile response. *Nat Chem Biol* 13: 333–338, 2017.

Long, et al.

33. Long MJC, Urul DA, Chawla S, Lin H-Y, Zhao Y, Haegele JA, Wang Y, and Aye Y. Precision Electrophile Tagging in Caenorhabditis elegans. *Biochemistry* 57: 216–220, 2018.

34. Magalhães MLB, Czekster CM, Guan R, Malashkevich VN, Almo SC, and Levy M. Evolved streptavidin mutants reveal key role of loop residue in high-affinity binding. *Protein Sci* 20: 1145–1154, 2011.

35. Maurais AJ and Weerapana E. Reactive-cysteine profiling for drug discovery. *Curr Opin Chem Biol* 50: 29–36, 2019.

36. Miyamoto S and Kollman PA. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches. *Proteins Struct Funct Bioinforma* 16: 226–245, 1993.

37. Mukherjee S, Yang JW, Hoffmann S, and List B. Asymmetric Enamine Catalysis. *Chem Rev* 107: 5471–5569, 2007.

38. Parvez S, Fu Y, Li J, Long MJC, Lin H-Y, Lee DK, Hu GS, and Aye Y. Substoichiometric hydroxynonenylation of a single protein recapitulates whole-cell-stimulated antioxidant response. *J Am Chem Soc* 137: 10–13, 2015.

39. Parvez S, Long MJC, Lin H-Y, Zhao Y, Haegele JA, Pham VN, Lee DK, and Aye Y. T-REX on-demand redox targeting in live cells. *Nat Protoc* 11: 2328–2356, 2016.

40. Parvez S, Long MJC, Poganik JR, and Aye Y. Redox Signaling by Reactive Electrophiles and Oxidants. *Chem Rev* 118: 8798–8888, 2018.

41. Piran U and Riordan WJ. Dissociation rate constant of the biotin-streptavidin complex. *J Immunol Methods* 133: 141–143, 1990.

42. Poganik JR, Long MJC, and Aye Y. Interrogating Precision Electrophile Signaling. *Trends Biochem Sci*, 2019.

43. Raphael MP, Rappole CA, Kurihara LK, Christodoulides JA, Qadri SN, and Byers JM. Iminobiotin Binding Induces Large Fluorescent Enhancements in Avidin and Streptavidin Fluorescent Conjugates and Exhibits Diverging pH-Dependent Binding Affinities. *J Fluoresc* 21: 647–652, 2011.

Long, et al.

44. Rehm M, Huber HJ, Dussmann H, and Prehn JHM. Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein. *EMBO J* 25: 4338–4349, 2006.

45. Ren J, Wen L, Gao X, Jin C, Xue Y, and Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Res* 19: 271–273, 2009.

46. Rokob TA, Hamza A, and Pápai I. Computing Reliable Energetics for Conjugate Addition Reactions. *Org Lett* 9: 4279–4282, 2007.

47. Ruff LE, Sapre AA, Plaut JS, De Maere E, Mortier C, Nguyen V, Separa K, Vandenbogaerde S, Vandewalle L, Esener SC, and Messmer BT. Selection of DNA nanoparticles with preferential binding to aggregated protein target. *Nucleic Acids Res* 44: e96–e96, 2016.

48. Sayers DL. *The Documents in the Case*.

49. Schwartz PA, Kuzmic P, Solowiej J, Bergqvist S, Bolanos B, Almaden C, Nagata A, Ryan K, Feng J, Dalvie D, Kath JC, Xu M, Wani R, and Murray BW. Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. *Proc Natl Acad Sci* 111: 173–178, 2014.

50. Smith AJT, Zhang X, Leach AG, and Houk KN. Beyond Picomolar Affinities: Quantitative Aspects of Noncovalent and Covalent Binding of Drugs to Proteins. *J Med Chem* 52: 225–233, 2009.

51. Surya SL, Long MJC, Urul DA, Zhao Y, Mercer EJ, Elsaid IM, Evans T, and Aye Y. Cardiovascular Small Heat Shock Protein HSPB7 Is a Kinetically Privileged Reactive Electrophilic Species (RES) Sensor. *ACS Chem Biol* 13: 1824–1831, 2018.

52. Świderek K, Nödling AR, Tsai Y-H, Luk LYP, and Moliner V. Reaction Mechanism of Organocatalytic Michael Addition of Nitromethane to Cinnamaldehyde: A Case Study on Catalyst Regeneration and Solvent Effects. *J Phys Chem A* 122: 451–459, 2018.

53. Tsuchiya Y, Okuno Y, Hishinuma K, Ezaki A, Okada G, Yamaguchi M, Chikuma T, and Hojo H. 4-Hydroxy-2-nonenal-modified glyceraldehyde-3-phosphate dehydrogenase is degraded by cathepsin G. *Free Radic Biol Med* 43: 1604–1615, 2007.

Long, et al.

54. Van Hall-Beauvais A, Zhao Y, Urul DA, Long MJC, and Aye Y. Single-Protein-Specific Redox Targeting in Live Mammalian Cells and C. elegans. *Curr Protoc Chem Biol* 10: e43, 2018.

55. Wakita C, Maeshima T, Yamazaki A, Shibata T, Ito S, Akagawa M, Ojika M, Yodoi J, and Uchida K. Stereochemical Configuration of 4-Hydroxy-2-nonenal-Cysteine Adducts and Their Stereoselective Formation in a Redox-regulated Protein. *J Biol Chem* 284: 28810–28822, 2009.

56. Wang C, Weerapana E, Blewett MM, and Cravatt BF. A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat Methods* 11: 79–85, 2014.

57. Wani R, Qian J, Yin L, Bechtold E, King SB, Poole LB, Paek E, Tsang AW, and Furdui CM. Isoform-specific regulation of Akt by PDGF-induced reactive oxygen species. *Proc Natl Acad Sci* 108: 10550–10555, 2011.

58. Weerapana E, Wang C, Simon GM, Richter F, Khare S, Dillon MBD, Bachovchin DA, Mowen K, Baker D, and Cravatt BF. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* 468: 790–795, 2010.

59. West JD, Ji C, Duncan ST, Amarnath V, Schneider C, Rizzo CJ, Brash AR, and Marnett LJ. Induction of Apoptosis in Colorectal Carcinoma Cells Treated with 4-Hydroxy-2-nonenal and Structurally Related Aldehydic Products of Lipid Peroxidation. *Chem Res Toxicol* 17: 453–462, 2004.

60. Winterbourn CC, Hampton MB, Livesey JH, and Kettle AJ. Modeling the Reactions of Superoxide and Myeloperoxidase in the Neutrophil Phagosome: Implications for Microbial Killing. *J Biol Chem* 281: 39860–39869, 2006.

61. Zhao Y, Long MJC, Wang Y, Zhang S, and Aye Y. Ube2V2 Is a Rosetta Stone Bridging Redox and Ubiquitin Codes, Coordinating DNA Damage Responses. *ACS Cent Sci* 4: 246–259, 2018.

62. Zoubir M, Zeroual A, El Idrissi M, El Haib A, Moumou M, Hammal R, Mazoir N, Benharref A, and El Hajbi A. Understanding the Chemoselectivity and Stereoselectivity in Michael Addition Reactions of β-Hydroxyparthenolides and Amines such as Pyrrolidine, Morpholine, Piperidine and 1-Methylpiperazine: a DFT Study. *J Mater Environ Sci* 8: 990–996, 2017.

Long, et al.

Figure Legends.

Figure 1. Comparison of covalent bond formation and non-covalent binding in terms of *approximate* free energy (ΔG) or enthalpy (ΔH) of binding at ambient temperature (RT= 8.314×298). For nitromethane coupling with cinnamaldehyde, an approximate average of all four calculated ΔG values in the cited literature (52) was presented; and for amine conjugate addition to parthenolides, the value shown is an approximate mean of 3 different 6-membered-cyclic amines (62).

Figure 2. Thermodynamics and kinetics of ligand–target interactions. **(A)** Change in ΔG as a function of K_d; color-coded from dark to light grey corresponding to: metal-enzyme interactions, covalent binding, high-affinity binders like biotin, highest-affinity inhibitors like Immucillin H, and then down to lower-affinity binders. **(B)** Half-life of bimolecular reactions (e.g., covalent bond formation) versus rate constant for different concentrations of reagents, assuming equal concentrations between the two reacting species. Five concentrations are chosen arbitrarily to represent high (20 µM) and low (0.05 µM) concentrations of reagent and protein (40). The arrow beside the y-axis denotes an estimate of the lifetime of HNE in cells. The arrow below the x-axis refers to the lower limit threshold of second-order rate constants considered to lead to biologically-relevant enzyme-catalyzed reactions. Fully-dark shapes indicate a likely biologically-relevant signaling process (half-life < 10 minutes); half-dark, half-grey shapes indicate a possibly biologically-relevant signaling process (half-life between 10 and 80 minutes); and grey shapes indicate likely non-biologically-relevant signaling process (half-life > 80 minutes).

Figure 3. Conformational flexibilities of select LDEs and divergent targets reported for HNE and 15d-PG J2. **(A)** Chemical structures of HNE, ONE, 15d-PG J2 and acrolein. Relevant *s*-trans or *s*-cis conformations are highlighted in red and purple, respectively. Shaded regions highlight unique structural features of these molecules discussed in the review. **(B)** Data available in the literature (56) are plotted using GraphPad Prism8. Human MDA-MB-231 proteome was treated with 100 µM HNE or 15d-PG J2. Reactivity of cysteines in proteome was studied with isotopic tandem orthogonal proteolysis–activity-based protein profiling (isoTOP-ABPP) platform. Cysteines with R values (isoTOP-ABPP ratios) greater than 5 were

Long, et al.

considered to be highly sensitive to compound added. In both graphs, the resulting hits from experiments using HNE vs. 15d-PG J2 are labeled, respectively, in blue and red.

Figure 4. Representative indirect and direct profiling of LDE-modified proteins. **(A)** Competitive isoTOP-ABPP platform. The proteomes are treated either with a lipid-derived electrophile (LDE) or DMSO and then treated with iodoacetamide-alkyne (IA-alkyne) which serves as a proxy electrophilic probe (56). The cysteines reactive to LDE are blocked from IA-alkyne labeling (figure drawn assuming that IA-alkyne also reacts faithfully with all LDE-reactive cysteines – see: DMSO-treated samples), so they are "lost" in the LDE-treated proteome compared to DMSO-treated proteome. (Note: figure further reflects the fact that not all remaining cysteines that have not been modified by LDE react with IA-alkyne – see: LDE-treated samples). The IA-alkyne labeled proteins are conjugated to isotopically-labeled tobacco etch virus protease (TEV)-cleavable biotin tag, and proteomes labeled with light and heavy isotopes are mixed in 1:1 ratio. After enrichment with streptavidin beads and on-bead trypsin digestion, the probe-labeled peptides are quantified by LC-MS/MS analysis. **(B)** A quantitative protein carbonylation profiling method using aniline-derived probe *m*-aminophenylacetylene (*m*-APA) (10). Proteome of interest is treated with *m*-APA. LDE-modified proteins (where carbonyl groups are still retained) can be labeled with *m*-APA and conjugated to acid-cleavable biotin tags. The labeled proteins are enriched with streptavidin pulldown, followed by trypsin digestion and acid cleavage. Target profiling is performed using LC-MS/MS analysis.

Figure 5. T-REX method interrogates individual protein- and LDE-chemotype-specific cellular response. **(A)** Targetable reactive electrophiles and oxidants (T-REX). Cells expressing HaloTagged-POI (protein of interest) are treated with a photocaged precursor of LDE of interest, housing a HaloTag-targeting motif (grey triangle) which stoichiometrically and irreversibly binds to HaloTag, and unbound precursors are washed out. Light exposure releases LDE within the proximity of POI. If POI is a privileged sensor of the LDE, the POI will be labeled specifically. Downstream signaling events triggered by labeling of POI by the LDE, as well as the identity of modified residue(s) and LDE-ligand occupancy can be analyzed (39, 42,

Long, et al.

54). **(B)** Chemical structures of T-REX photocaged precursor of LDEs, and two representative LDEs (HNE and CHE) (27).

Figure 6. Sensing sites on Keap1 vary under different experimental conditions, although T-REX faithfully labels specific sites regardless of N- or C-terminus Halo fusion and shows chemotype-specific changes of sensing site within Keap1; Akt-isoform specific sensing arises from a single cysteine residue within Akt3 linker region. **(A)** Domain structure of human Keap1 protein. Cysteines detected to be modified by HNE or CHE under indicated conditions are illustrated (27, 38). **(B)** Domain structure of human Akt1, Akt2 and Akt3. The amino acid sequences of the linker region are shown. Akt2(C124) (bolded and underlined) is a peroxide-sensing residue (57). Akt3(C119) (bolded and underlined) is an HNE-sensing residue (31, 32). Domain structures were generated with software DOG 2.0 (45).
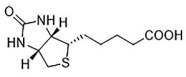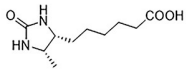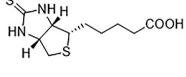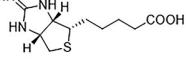
Figure 7. Comparison between external administration of reactive molecules and in-situ release in G-REX method (61).

Long, et al.
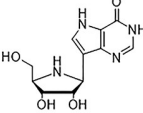
Figure 1

Long, et al.

Figure 2
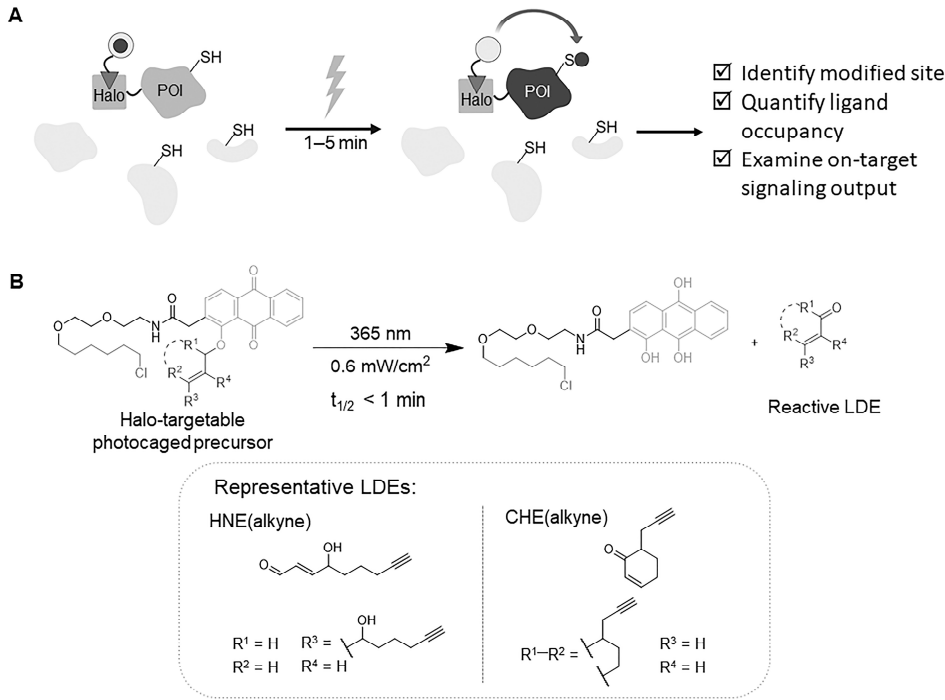
Long, et al.

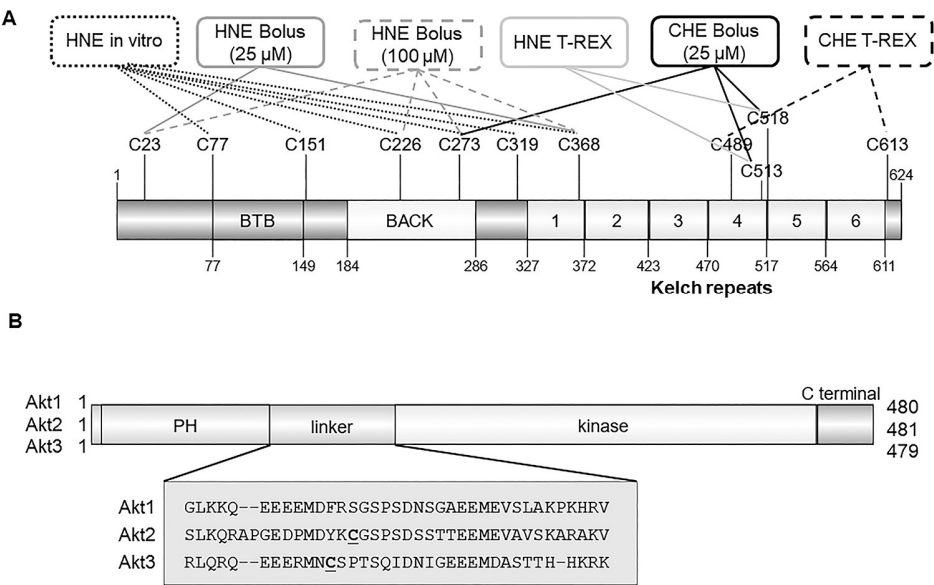**Figure 3**

37
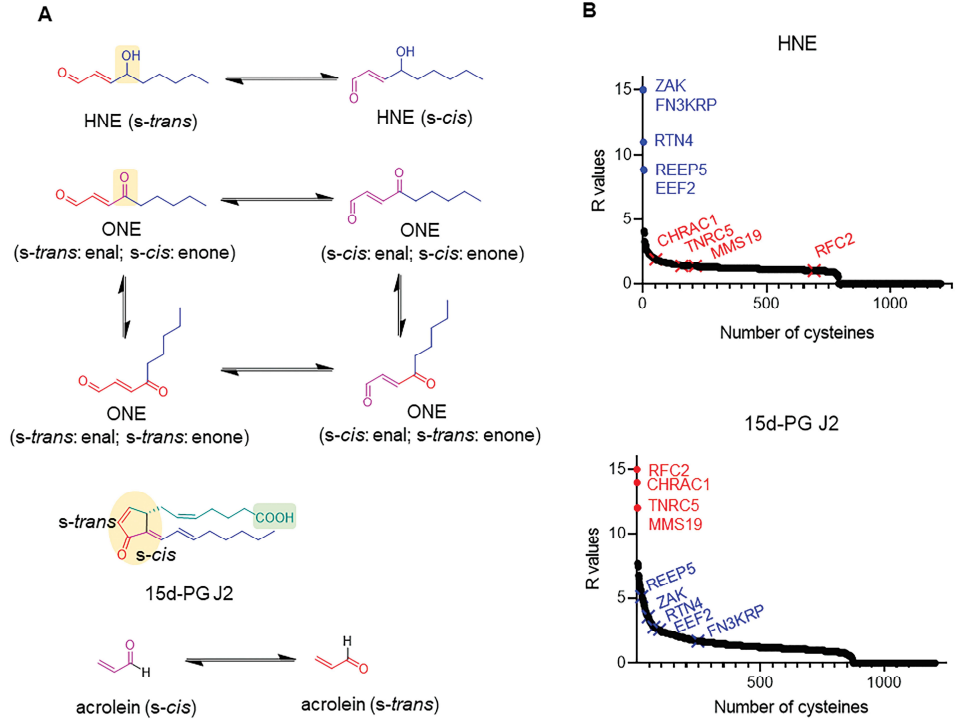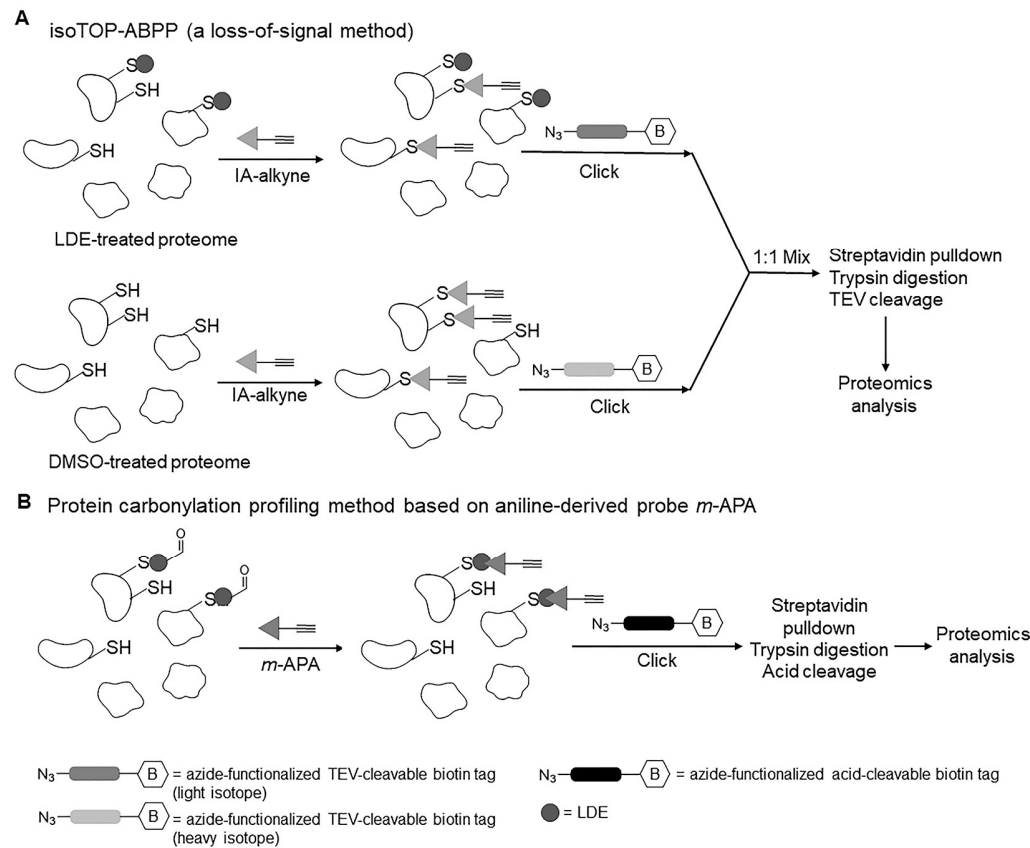
Long, et al.

Figure 4

Long, et al.

Figure 5

Long, et al.

Figure 6

Long, et al.

Figure 7

External administration — 5–10 min lag phase — Concentration gradient — = cellular proteins — = reactive LDE (e.g., HNE)

Challenges: diffusion distribution metabolism …

In situ release — Maximum labeling < 3 min — Halo — Halo — = photocaged LDE — = light