Activity-based protein profiling for mapping and pharmacologically interrogating proteome-wide ligandable hotspots
Allison M Roberts, Carl C Ward and Daniel K Nomura

Despite the completion of human genome sequencing efforts nearly 15 years ago that brought with it the promise of genome-based discoveries that would cure human diseases, most protein targets that control human diseases have remained largely untranslated, in-part because they represent difficult protein targets to drug. In addition, many of these protein targets lack screening assays or accessible binding pockets, making the development of small-molecule modulators very challenging. Here, we discuss modern methods for activity-based protein profiling-based chemoproteomic strategies to map ‘ligandable’ hotspots in proteomes using activity and reactivity-based chemical probes to allow for pharmacological interrogation of these previously difficult targets. We will showcase several recent examples of how these technologies have been used to develop highly selective small-molecule inhibitors against disease-related protein targets.

Address
Departments of Chemistry, Molecular and Cell Biology, and Nutritional Sciences and Toxicology, University of California, Berkeley, Berkeley, CA 94720, United States

Corresponding author: Nomura, Daniel K (dnomura@berkeley.edu)

Current Opinion in Biotechnology 2017, 43:25–33
This review comes from a themed issue on Analytical biotechnology
Edited by Jurre Kamphorst and Ian Lewis

http://dx.doi.org/10.1016/j.copbio.2016.08.003
0958-1669/© 2016 Elsevier Ltd. All rights reserved.

Introduction
While many disease-modifying protein targets have been discovered, most of these targets have remained untranslated as many of these proteins are considered to be ‘undruggable’ or difficult to target with small-molecule drugs. Indeed, most of the proteome is devoid of pharmacological tools, hindering and oftentimes paralyzing both basic and translational research efforts. Studies have convincingly demonstrated that the development of high-quality chemical tools for proteins of interest catalyze research into the function and therapeutic exploitation of those proteins, thus correlating the development of chemical tools for specific proteins with their associated research activity [1]. As Scott Dixon and Brent Stockwell pointed out in their poignant review article, only ~2% of all predicted human gene products are currently targeted with small-molecule drugs and only 10–15% of all human genes are thought to be ‘druggable,’ with only a 25% overlap between druggable protein targets and known disease-modifying targets [2]. Thus, developing approaches that enable the discovery of pharmacological tools for every protein in the proteome would radically expand our ability to understand protein function and accelerate the drug discovery process to cure complex diseases. Here, we will discuss established, emerging, and potential applications of activity-based protein profiling (ABPP) as a powerful chemoproteomic strategy to map ‘ligandable’ hotspots or sites within proteins that can be pharmacologically interrogated for drug discovery. We will also discuss the advantages of coupling ABPP platforms with the development of irreversible small-molecule modulators of protein targets to facilitate drug discovery efforts.

ABPP with active-site directed probes to develop active-site directed inhibitors
With the completion of human genome sequencing efforts over a decade ago, there has been a widening gap between the number of promising genes and their encoded proteins that have been linked to health and disease and pharmacological tools and drugs to interrogate these targets for biological characterization and disease therapy [2,3]. This widening gap has formed in-part because many of these disease-relevant protein targets are considered ‘undruggable,’ in which these proteins may not have obvious binding pockets that can be pharmacologically interrogated or do not possess high-throughput screening amenable functional assays for identifying small-molecule modulators against these targets. Compounding upon these challenges, many targets that may be linked to disease may also be uncharacterized as to their biochemical roles, thus hindering both their mechanistic characterization as well as the ability to assay for inhibitors against these targets.

Chemical proteomics or chemoproteomics has arisen as a powerful technology to broadly assess protein functionality, and through doing so, has enabled strategies to not
only assay protein function for even uncharacterized protein targets, but also to develop inhibitors against these targets [4–6]. Among these chemoproteomic strategies, activity-based protein profiling (ABPP) has been particularly useful for coupling the assessment of protein activities with inhibitor development (Figure 1). ABPP uses active site-directed or reactivity-based chemical probes to map protein functionality directly in complex proteomes. In addition to a chemical warfare that binds to functional sites within proteins, these probes also bear chemical handles, such as fluorophores, biotin, or alkynes for subsequent fluorescent visualization, avidin-enrichment and mass-spectrometry (MS)-based quantification, or biorthogonal conjugation of handles for subsequent analysis of protein activities, respectively [7]. Because these probes bind to functional sites within protein targets, small-molecule inhibitors can be competed against probe binding, enabling a universal assay strategy for identifying inhibitors against any protein that can be assayed with ABPP methods [3]. ABPP is amenable to many assay formats including gel-based formats measuring in-gel fluorescence, quantitative proteomic methods, high-throughput fluorescence polarization-based technologies, or imaging probes in living systems [8–12]. Additionally, because activity-based probes not only assess the activity of the protein target of interest, but also activities across an entire enzyme class when performed in complex biological samples, selectivity of inhibitors can be assessed on a proteome-wide scale. Subsequently, lead selective inhibitors can also be turned into probes through the development of biorothogonal analogs to further assess inhibitor selectivity [6]. From bench science to pre-clinical and clinical perspectives, developing covalent inhibitors of targets using ABPP has also been particularly advantageous. Target engagement and selectivity of covalent inhibitors can be assessed ex vivo using ABPP methods, which is not only valuable for biological characterization, but also for confirmation of target engagement in pre-clinical and clinical testing paradigms in animal models and humans [3]. In this section, we will initially focus on the original application of ABPP using activity-based probes to map protein activity and develop active-site directed inhibitors. There have been many activity-based probes developed over the past two decades including fluorophosphonate (FP)-based probes, vinyl sulfone or epoxide-based probes, diphenyl phosphonates, β-lactams and β-lactones, 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF), carbamates and heterocyclic ureas for serine hydrolases and proteases [13*,14]. Additional probes have been developed for cysteine proteases, glycosidases, serine, threonine, and tyrosine phosphatases, glycosidases, cytochrome P450s, ubiquitin-modifying enzymes, proteasomes, oxidoreductases, lysine acetyltransferases, and ATP-binding enzymes [14–16] (Figure 2). There have been many previous reviews on ABPP which have described activity-based probes and their utility so we will focus here on recent examples using this technology for inhibitor development.

Among the most successful application of ABPP has been in pharmacological interrogation of the serine hydrolase superfamily of enzymes that constitutes activities such as esterases, lipases, thioesterases, proteases, and peptidases, many of which have been linked to important (patho)physiological roles [17]. There have been many prior successes with ABPP in developing small-molecule inhibitors include the development of highly selective inhibitors against serine hydrolases such as fatty acid amid hydrolase (FAAH), monoacylglycerol lipase (MAGL), and diacylglycerol lipase (DAGL) involved in endocannabinoid and cicosanoid biology, KIAA1363 involved in ether lipid metabolism and cancer pathogenicity, and many others [17–21].

There have been many more examples in the past few years of using ABPP to develop inhibitors against both characterized and uncharacterized serine hydrolases. A selective and reversible platelet activating factor acetyl hydrolase 1B2 and 1B3 (PAFAH1B2 and PAFAH1B3) inhibitor P11 was identified through fluorescence-polarization based ABPP methods and was used to show that PAFAH1B2 and 1B3 were important metabolic drivers of cancer pathogenicity in multiple types of aggressive human cancer cells [22,23]. A highly selective inhibitor KLH45 was developed against the uncharacterized serine hydrolase DDHD2 using ABPP methods, and was utilized in mouse models to aid in characterizing DDHD2 as an enzyme that hydrolyzes triacylglycerols in the brain. Inloes et al. showed that inactivation of DDHD2 led to a debilitating neurodegenerative disease known as hereditary spastic paraplegia which results in lower limb spasticity and weakness and intellectual disability, due to the accumulation of triacylglycerols and lipid droplets in the brain [24]. Carboxylesterase 3 (Ges3) inhibitors were also discovered through chemical genetic screening coupled with ABPP as agents that ameliorate multiple features of the metabolic syndrome [25]. Inhibitors were also developed against previously uncharacterized hydrolase ABHD16A, KC01, using ABPP methods. Using subsequent metabolic approaches, Kamat et al. discovered that ABHD16A is a phosphatidylerine hydrolase that releases lysophosphatidylserine which in turn is further hydrolyzed by ABHD12. Previous studies had shown that ABHD12 loss of function led to a neurodegenerative disease known as polyneuropathy, hearing loss, ataxia, retinitis pigmentosa and cataract (PHARC) due to an accumulation of pro-inflammatory lysophosphatidylerine levels. The authors showed that inhibition of ABHD16A could reduce lysophosphatidylerine levels and subsequently attenuate neuroinflammatory responses [26] (Figure 3). In another study, Xu et al. used a substrate-competitive ABPP approach to identify carboxylesterase 1 as an oseltamivir-activating enzyme in
ABPP platforms and their applications. (a) Design of activity-based and reactivity-based biorthogonal probes. (B) Using isoTOP-ABPP platforms for identifying disease-relevant targets. Normal and diseased proteomes can be labeled with activity or reactivity-based probes, followed by appendage of isotopically light or heavy analytical biotin handles bearing a TEV protease cleavage sequence, followed by mixing of proteomes in a 1:1 ratio, avidin enrichment of probe-labeled proteins, TEV digestion to release probe-modified tryptic peptides, and quantitative proteomic analysis of probe-modified peptides. (C) Using isoTOP-ABPP for mapping hyper-reactive and functional hotspots in the proteome. The procedures would mirror those described in (B), except that proteomes would be labeled with varying concentrations of probes (e.g. 10x versus 1x probe concentration) to quantitatively map hyper-reactive sites. (D) Using isoTOP-ABPP for pharmacological interrogation of ligandable hotspots. The procedures would mirror those described in (B), except that proteomes would be treated with vehicle or inhibitor to map not only the sites of probe labeling, but also sites where the inhibitor displaced probe labeling, facilitating both inhibitor discovery for targets of interest as well as an assessment of its proteome-wide selectivity.
intestinal cell homogenates [27]. In a study by Wolf et al., the authors used competitive ABPP methods to screen a library of 50 small-molecule inhibitors consisting of known serine hydrolase-inhibiting scaffolds such as iso-
coumarins, phosphonates, and β-lactones against 13 dif-
ferent rhomboid serine proteases and identified both pan
rhomboid inhibitors as well as several that displayed
selectivity for certain rhomboids over others [28].

Examples of using competitive ABPP platforms for in-
hibitor discovery outside of the serine hydrolase family
include a recent report of inhibitors against two atypical
integral membrane hydrolases AIG1 and ADTRP that
degrade bioactive fatty acid esters of hydroxyl fatty acids
(FAHFs). Interestingly, AIG1 was identified through
ABPP profiling using a serine hydrolase probe, but the
probe reacted instead with a conserved threonine, and
ADTRP was identified due to its homology to AIG1.
Nonetheless, the authors were able to develop inhibitors
KC01 and JJH260 for AIG1 using competitive ABPP [29].

Competitive ABPP methods using activity-based probes
have also been used to identify off-target effects of
environmental chemicals. Medina-Cleghorn et al. used
ABPP platforms to map off-target profiles of multiple
widely used organophosphorus pesticides and showed
that MAGL and FAAH as well as many other serine
hydrolases were inhibited by these agents in vivo in mice.

Examples of activity and reactivity-based probes. ‘TAG’
refers to an analytical handle for analysis of probe-labeled
proteins such as a biotin or fluorophore handle.
Examples of small-molecule inhibitors developed using competitive ABPP platforms.

[30]. The authors further showed that inhibition of these targets led to downstream biochemical changes in substrate and product levels of the inhibited enzymes.

Overall, ABPP using probes targeting specific enzyme classes have been used successfully to develop pharmacological inhibitors against both characterized and uncharacterized protein targets and has expanded the enzyme/protein space that can be studied. Enzyme classes that have been pharmacologically interrogated with ABPP include hydrolases, cysteine proteases, glycosidases, cytochrome P450s, ubiquitin-modifying enzymes, proteasomes, lysine acetyltransferases, and ATP-binding enzymes [14–16]. However, most enzyme classes remain devoid of active-site directed probes, thus hindering our ability to functionally assay and pharmacologically interrogate proteins in these classes.

**ABPP with reactivity-based probes to map new proteome-wide druggable hotspots**

ABPP approaches using active-site directed probes have led to the development of many inhibitors for both characterized and uncharacterized enzymes, leading to their biological characterization and the translation of these inhibitors into clinical development. While many activity-based probes have been developed by the chemical biology field over the last decade, there are still many different enzyme classes and more importantly a diversity of protein functionalities that are not detectable with the current arsenal of active-site directed probes. More recent efforts in using ABPP have employed generally reactive electrophilic probes to more globally map protein functionality through the quantitative peptide-level mapping of hyper-reactive hotspots in complex proteomes using isotopic tandem orthogonal proteolysis — activity-based protein profiling (isoTOP-ABPP) [31*]. These reactivity-based probes that more generally react with nucleophilic hotspots within proteomes, such as cysteines and lysines, are not only restricted to catalytic sites within enzymes, but also solvent-accessible binding pockets, post-translational modification sites, cysteine oxidation sites, protein-protein interaction sites, and other types of regulatory or functional domains across the entirety of the proteome, thus greatly expanding the ligandable sites that can be addressed.
both functionally characterized and pharmacologically targeted [15,16,32,33] (Figure 2).

Reactivity-based probes consist of three features which enables mining and pharmaceutical targeting of proteome-wide functional hotspots. Representative reactivity-based probes consist of: first, an electrophilic warhead; second, an alkyne handle for ‘click chemistry’ conjugation of an enrichment handle for probe-labeled proteins; and finally an azide functionalized TEV protease recognition peptide linker containing a biotin group for avidin enrichment of probe-labeled proteins as well as an isotopically light or heavy valine for quantitative ratiometric mass-spectrometry (MS)-based proteomic analysis of probe-labeled peptides in proteomes to identify and quantitatively compare the reactivity of probes against specific amino acid sites of probe-modification (Figure 2). When coupled with isoTOP-ABPP platforms to quantitatively map reactivity of specific sites of probe-modification, this overall approach enables a much broader and global mapping of protein functionality and more importantly, facilitates the identification of ligandable and eventually druggable hotspots within protein targets that may have previously been considered to be undruggable.

Among these reactivity-based probes, the cysteine-reactive iodoacetamide (IA)-alkyne probe, has been particularly versatile in mapping hyper-reactive and functional cysteines in proteomes [31]. Additional cysteine reactive probes have also been developed for chemoproteomic profiling. IA-alkyne is a versatile cysteine-reactive probe, but cannot be used on living cells due to its high level of reactivity. To be able to profile cysteine reactivity in living cells, Abo and Weerapana developed a caged bromomethylketone electrophilic probe that can be spatially and temporally activated within living cells through irradiation to map cysteine reactivity in live cells [34]. Aegbog et al. developed an ethynyl benziodoxolone reagent to further expand the accessible cysteine landscape [35]. Different states of cysteine oxidation have also been targeted for isoTOP-ABPP profiling. Yang et al. used the dimeredone-based probe for selective labeling of S-sulfenic acids in intact cells and isoTOP-ABPP to map proteome-wide specific sites of S-sulfenylation [36]. In another study, Majmudar et al. showed that sulfinate-linked probes could be used to enrich and annotate hundreds of endogenous S-nitrosated proteins [37]. Conversely, these authors also showed that S-nitrosothiol-linked probes could be used to enable enrichment and detection of endogenous S-sulfurated proteins. Using these probes, the authors demonstrated that hydrogen peroxide increased S-sulfination of human DJ-1 at Cys106, but that Cys46 and Cys53 underwent full oxidation to sulfonic acids [37].

There have also been considerable efforts in developing reactivity-based probes that target other important nucleophilic amino acid residues in biology. Irreversible acetylation of serines on cyclooxygenase enzymes has long been understood as the mechanism for acetylsalicylic acid’s (aspirin) anti-inflammatory and antipyretic effects. Both Bateman et al. and Wang et al. developed an aspirin-alkyne probe to map the targets of aspirin acylation and surprisingly revealed that this probe could be used to broadly label lysine, serine, arginine, histidine, threonine, tyrosine, tryptophan and cysteine, thus revealing the potential diverse mechanisms underlying the anti-inflammatory actions of aspirin, but also showing the potentially versatility of this scaffold in expanding our future mapping of hyper-reactive and functional hotspots using isoTOP-ABPP platforms [38,39]. In another study, Lewallen et al. developed a biotin-conjugated phenylglyoxal probe to map global protein citrullination, a post-translational modification that is particularly heightened in rheumatoid arthritis, and identified >50 intracellular citrullinated proteins, of which more than 20 of these were involved in RNA splicing, suggesting a role for citrullination in linking RNA biology to inflammatory disorders [40]. Shannon et al. investigated the differences in reactivity of aryl halides by developing alkyne-tagged probes and using isoTOP-ABPP and showed that p-chloronitrobenzene was highly cysteine reactive, whereas the dichlorotriazine scaffold favored reactivity with lysines [13*].

**ABPP with reactivity-based probes to map the proteome-wide targets of endogenous and exogenous electrophiles**

Beyond global mapping of hyper-reactive and functional sites, reactivity-based probes have also been used in a competitive manner to map specific cysteines on particular protein targets or off-targets that are susceptible to modification by endogenously-derived electrophiles as well as pesticides and pharmaceutical agents. The IA-alkyne probe has been used to map cysteine-reactivity of endogenously-derived lipid aldehydes such as hydroxy-nonanal and S-nitrosoglutathione [41,42*]. Wang et al. showed that 4-hydroxy-2-nonenal shows selective labeling of certain cysteines on specific protein targets, such as an active-site proximal cysteine on ZAK kinase which results in enzyme inhibition and creates a negative feedback mechanisms that suppresses the activation of JNK pathways normally induced by oxidation stress [42*]. Zhou et al. recently performed competitive isoTOP-ABPP profiling of S-nitrosoglutathione against IA-alkyne labeling to quantitatively map particularly sensitive transnitrosation sites, and discovered that S-nitrosation of a cysteine residue distal to the HADH2 active site and a catalytic cysteine in the lysosomal aspartyl protease cathepsin D (CTSD) impaired catalytic activity or proteolytic activation of these proteins, respectively [41].

Reactivity-based probes have also been used to map the cysteine reactivity of both agrochemical and pharmaceutical agents as well. Medina-Cleghorn et al. used the
IA-alkyne probe to competitively profile the *in vitro* and *in vivo* reactivity of various cysteine-reactive pesticides and showed in-particular that the widely used fungicide chlorothalonil targets the catalytic cysteines of several enzymes involved in fatty acid oxidation and metabolism, thus identifying a novel mechanisms of toxicity for these pesticides [43]. Abegg *et al.* used their ethynyl benziodoxolone cysteine-reactive probe to reveal that the anticancer agent curcumin covalently modified cysteines on several key players of cellular signaling and metabolism, including casein kinase 1 gamma, showing potential mechanisms of action of this drug [35].

**ABPP with reactivity-based probes or reactive chemical scaffolds to develop small-molecule inhibitors against ligandable hotspots in proteins**

The coupled usage of reactivity-based chemoproteomic probes and isoTOP-ABPP platforms has led to the discovery of hundreds to thousands of potentially novel functional sites within proteins that may represent not only catalytic sites, but also regulatory or functional sites of post-translational modification, sites of oxidation, protein–protein interactions, metal binding, and allosteric regulation [13*,31*,35,39]. These methods have mostly been used for identification of these sites, or competitive studies to map reactivity of endogenous and exogenous electrophilic agents. However, the isoTOP-ABPP technology has much broader potential applications in the area of drug discovery, in which the sites of probe-labeling by reactivity-based probes each constitute potential ligandable sites within these proteins that can be pharmacologically interrogated for drug discovery efforts (Figure 3).

Crawford and Weerapana, for example, recently developed a library of functionalized dichlorotriazines and found that the L-leucine methyl ester functionalized DCT variant (LAS-17) proved to be a highly selective covalent modifier of GSTP1 through reacting with a tyrosine 108 which in-turn inhibits GSTP1 activity [44]. This inhibitor was subsequently used in another study where the dichlorotriazine-alkyne probe was used to discover that GSTP1 was heightened selectively in triple-negative breast cancer cells and that LAS17 treatment and GSTP1 blockade attenuated triple-negative breast cancer pathogenicity through impairing breast cancer metabolism and signaling [45]. In another example, Wang *et al.* showed that chloromethyltriazoles were a promising cysteine-reactive scaffold and that its reactivity could be tempered to confer selectivity for certain cysteines on others. The authors showcase a chloromethyltriazole AA-CW236 as the first potent, selective, non-pseudosubstrate inhibitor of the O(6)-alkylguanine DNA methyltransferase (MGMT) [46].

More recently, this method was used in concert with chemoproteomic screening of a cysteine-reactive chloroacetamide and acrylamide fragment library to identify pharmacological ligands against over 700 cysteines across 637 unique protein targets or sites within protein targets classically considered to be undruggable [47*]. The authors able to identify ligands against many protein-targets in parallel through the isoTOP-ABPP screening of cysteine-reactive fragments directly in complex proteomes or living cells. More importantly 82% of these protein did not have a small molecule modulator or probe annotated within the DrugBank database, indicating the promise of this type of approach for massively expanding upon the druggability of the proteome. Among these interactions, Backus *et al.* focused on one particular example where they identified a cysteine-reactive fragment that preferentially reacts with pro-caspases [47*].

**Conclusions**

Over the past nearly two decades since ABPP was introduced, ABPP platforms have shown great versatility in their applications, including identifying new disease-relevant targets, developing small-molecule inhibitors against these targets, assessing proteome-wide selectivity of drugs or revealing off-targets of pharmaceutical or environmental chemicals, and revealing novel functional sites in the proteome. Perhaps the most exciting application moving forward is coupling ABPP with modern quantitative proteomic methods and newer active-site directed, reactivity-based, or other types of probes to comprehensively and globally map all ligandable hotspots or binding pockets in the proteome towards greatly expanding our scope of pharmacological tool-chest and drugs for the largely pharmacologically-devoid and undruggable proteome. This review primarily covers how this can be potentially achieved with active-site targeted or reactivity-based probes, but these types of approaches can be even further advanced with reversibly-acting molecules that are made to act covalently with photoaffinity tags to even further broaden our scope of ligandable sites. For example, this has been recently accomplished with a fatty acid probe bearing a biorthogonal handle and photoaffinity tag to reveal the lipid interaction proteome. Niphakis *et al.* used this approach to not only show that the lipid interaction proteome is rich in potential drug targets, but also identified a lipid binding protein nucleobindin-1 that perturbs the metabolism of endocannabinoids and used the lipid probe to develop a selective chemical modulator for this protein [48].

Furthermore, while covalent inhibitors have been historically frowned upon due to their potential non-selective reactivity, coupling the development of inhibitors with chemoproteomic platforms such as ABPP enables the synthesis of potent and selective covalent inhibitors against protein targets through ABPP driving the medicinal chemistry and vice versa [7,47*,49]. The future focus of using chemoproteomics and ABPP towards expanding our scope of the druggable proteome will not only lead to
potential new cures for diseases but will be an engine for discovery that will rival the promise seen with emerging gene editing methods.

Acknowledgements

We thank the members of the Nomura Research Group for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health (R01CA173667), American Cancer Society Research Scholar Award (RSG14-242-01-TBE), and DOE Breakthrough Award (CDMRP W81XWH-15-1-0050).

References and recommended reading

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

- of special interest


This paper reveals new reactive probe scaffolds that can be used for chemoproteomics. The authors show that various aryl halides can be designed to confer selectivity towards certain amino acids over others.


This paper was the first studies to use the isoTOP-ABPP technology to quantitatively map cysteine reactivity in complex proteomes to identify hyper-reactive and functional hotspots.


42. Wang C, Weerapana E, Blewett MM, Cravatt BF: A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. Nat Methods 2014, 11:79-85. This study used the isoTOP-ABPP platform in a competitive manner to map the proteome-wide reactivity of lipid-derived electrophiles, showing that even seemingly non-specific lipid aldehydes show selective reactivity that results in specific functional outputs.


47. Backus KM, Correia BE, Lum KM, Forl S, Horning BD, González-Páez GE, Chatterjee S, Lanning BR, Teijaro JR, Olson AJ et al.: Proteome-wide covalent ligand discovery in native biological systems. Nature 2016, 534:570-574. This study showed the feasibility of massively expanding our scope of pharmacologically interrogating large numbers of cysteines (both known and previously unknown and undruggable) in the proteome through competitively profiling cysteine-reactive fragment libraries against the binding of cysteine-reactive probes directly in complex proteomes.
