The development and application of methods for activity-based protein profiling
Nadim Jessani and Benjamin F Cravatt*

Recent advances in genomic and proteomic technologies have begun to address the challenge of assigning molecular and cellular functions to the numerous protein products encoded by prokaryotic and eukaryotic genomes. In particular, chemical strategies for proteome analysis have emerged that enable profiling of protein activity on a global scale. Herein, we highlight these chemical proteomic methods and their application to the discovery and characterization of disease-related enzyme activities.

Introduction
With the availability of complete genome sequences for numerous prokaryotic and eukaryotic organisms, biological researchers are confronted with an immense number of novel genes and gene products in need of functional assignment. The field of proteomics aims to accelerate this process by developing and applying methods for the parallel analysis of large numbers of proteins [1,2]. Proteomic strategies include efforts to characterize both protein expression and protein function on a global scale. The most mature method for analyzing protein expression patterns utilizes two-dimensional electrophoresis (2DE) for the separation of proteins coupled with protein staining and mass spectrometry (MS) for protein detection and identification, respectively [3,4]. Although 2DE-MS methods permit the consolidated analysis of the relative levels of many proteins across multiple proteomic samples, these approaches suffer from an inability to resolve several important protein classes, including low abundance and membrane-associated proteins [5,6]. To address these shortcomings, powerful strategies for the gel-free analysis of proteomes have emerged, including isotope-coded affinity tagging (ICAT) for quantitative proteomics [7] and multidimensional protein identification technology (MudPIT) for comprehensive proteomics [8], both of which utilize liquid chromatography (LC) and MS for protein separation and detection, respectively. Nonetheless, these methods, like 2DE-MS, still focus on measuring changes in protein abundance and, therefore, provide only an indirect estimate of dynamics in protein function. Indeed, several important forms of post-translational regulation, including protein–protein and protein–small-molecule interactions [9], may elude detection by abundance-based proteomic methods.

To facilitate the analysis of protein function, several proteomic methods have been introduced to characterize the activity of proteins on a global scale. These include large-scale yeast two-hybrid screens [10,11] and epitope-tagging immunoprecipitation experiments [12,13], which aim to construct comprehensive maps of protein–protein interactions, and protein microarrays [14,15], which aim to provide an assay platform for the rapid assessment of protein activities. Although these methods have the advantage of assigning specific molecular functions to individual protein products, they typically rely on the recombinant expression of proteins in artificial environments and, therefore, do not directly assess the functional state of these biomolecules in their native settings. To address this issue, a chemical proteomic strategy referred to as activity-based protein profiling (ABPP) has emerged that utilizes active site-directed probes to profile the functional state of enzyme families directly in complex proteomes [16,17].

In contrast to conventional proteomic strategies, which aim to catalogue the entire complement of protein products in a given sample, ABPP is designed to address the proteome at the level of discrete enzyme families, providing a way to distinguish, for example, active enzymes from their inactive zymogen [18] and/or inhibitor-bound forms [18–21]. Here, we review recent examples of the biological application of ABPP as it relates to: 1) comparative profiling, for the discovery of enzyme activities associated with discrete physiological and pathological states, and 2) competitive profiling, for the discovery of potent and selective inhibitors of enzymes. These studies highlight the benefits that accompany parsing the proteome into tractable, functional units (activity states of...
given enzyme classes) to provide information not readily achieved with abundance-based techniques. Additionally, we discuss the future challenges that face ABPP and related functional proteomic endeavors.

**ABPP: a chemical strategy for profiling enzyme activities in complex proteomes**

An activity-based chemical probe consists of three general elements: first, a reactive group (RG), which binds and covalently modifies the active sites of a broad range of enzymes from a particular enzyme class (or classes); second, a linker region; and third, a chemical tag for the consolidated detection and isolation of probe-labeled enzymes from complex proteomes (Figure 1). Two general strategies have been introduced for the design of activity-based probes. The first approach incorporates well-characterized affinity labels as the RG to direct probe reactivity toward enzymes sharing a similar catalytic mechanism and/or substrate specificity. These directed ABPP efforts have generated probes that profile several enzyme classes in complex proteomes, including serine hydrolases (SHs) [18,19], cysteine proteases [20–23] and proteasomal subunits [24]. Directed efforts have also succeeded in creating initial probes for tyrosine phosphatases [25] and glycosidases [26]; however, whether these probes can detect their protein targets in whole proteomes remains unknown. In a second strategy for ABPP, libraries of candidate probes are synthesized with variable binding groups appended to a common RG, and these reagents are screened against complex proteomes to identify activity-based labeling events. This non-directed version of ABPP has yielded sulfonate ester probes that label members of several mechanistically distinct enzyme classes in complex proteomes, including aldehyde dehydrogenases [27], enoyl coenzyme A hydratases, epoxide hydrolases, glutathione S-transferases, phosphofructokinases and transglutaminases [28,29]. Because the design, synthesis and initial evaluation of ABPP probes has been the subject of several recent reviews [16,17,30], we focus instead on the biological application of these reagents for the discovery of disease-related enzyme activities, as well as lead inhibitors that target these enzymes.

**Comparative ABPP of human cancer cells**

Cancer remains one of the most prevalent and life-threatening diseases for which effective treatments and cures are lacking. The identification of enzyme activities selectively expressed by tumor cells or tissues may provide a rich source of new biomarkers and targets for the diagnosis and treatment of cancer. With this objective in mind, Jessani and colleagues [31] set out to quantitatively profile the activity, subcellular distribution and glycosylation state of members of the SH superfamily expressed across a panel of human cancer cell lines. Because SHs represent one of the largest and most diverse enzyme classes in higher eukaryotic proteomes, collectively constituting approximately 1% of the predicted protein products encoded by the human genome [32,33], this study also offered an opportunity to test whether the proteomic information garnered by ABPP was of sufficient quantity and quality to depict the higher-order cellular properties of cancer cells. Profiling of the

*Figure 1*

Comparative ABPP of human cancer cells [31]. Complex proteome before and after treatment with a SH-directed ABPP probe, FP-Rhodamine (for structures of activity-based probes targeting other enzyme classes see [16,17,30]). Active enzymes are denoted by open/unshaded active sites, with their inactive counterparts shaded in black. FP-rhodamine labeled activity profile: shown is a representative in-gel fluorescence analysis of an FP-rhodamine-labeled secreted proteome derived from invasive and non-invasive human melanoma cancer cells. Native and deglycosylated (PNGaseF treated) enzyme activities selectively associated with either invasive cells (e.g. urokinase and esterase D) or non-invasive cells (e.g. sialic acid 9-O-acetylesterase [SAE] and cathepsin A) are noted.
secreted, membrane-associated, and soluble fractions derived from human breast carcinoma and melanoma cell lines resulted in the identification of a cluster of activities that distinguished cancer lines according to their respective tissue of origin. Interestingly, nearly all of these activities were down-regulated in the most invasive cancer lines analyzed, which instead up-regulated a distinct set of secreted and membrane-associated SH activities. These invasiveness-associated enzyme activities included urokinase, a secreted serine protease with a recognized role in tumor progression, as well as a novel membrane-associated hydrolase, KIAA1363, for which no previous link to cancer has been made. Interestingly, high levels of KIAA1363 activity were also found to correlate with invasive behavior in a panel of ovarian cancer cell lines that, despite forming a discrete cluster based on global gene expression profiles [34], were otherwise relatively uncharacterized in terms of cellular phenotypes.

Notably, in contrast to the diverse patterns of enzyme activity observed in the secreted and membrane proteomes of cancer cells, their soluble proteomes appeared quite similar, with few enzyme activities exhibiting restricted patterns of distribution. This finding suggests that, at least for the SH superfamily, the membrane and secreted proteomes are especially enriched in enzyme activities that serve as markers of cellular phenotype, highlighting the value of methods, such as ABPP, that can analyze technically challenging proteomic fractions (e.g. secreted, membrane, glycosylated and low-abundance proteins).

Comparative ABPP of the Plasmodium falciparum life cycle

Findings from another ABPP study also highlight the benefit of addressing the proteome at the level of distinct enzyme classes. In an elegant study by Greenbaum and colleagues [35], activity-based probes specific for the papain subclass of cysteine proteases were applied to characterize the function of these enzymes in the Plasmodium falciparum life cycle. Whereas cysteine proteases are known to be essential for the survival of several human parasites, the specific roles played by these enzymes during the complex life cycle of P. falciparum have remained ill-defined. ABPP of P. falciparum proteomes isolated at various stages of the parasite life cycle identified a specific cysteine protease, falcipain 1, that was upregulated during the invasive merozoite stage of growth. Falcipain 1-selective inhibitors were then identified by screening epoxide-based chemical libraries for compounds that blocked probe labeling of this enzyme in complex proteomes. These inhibitors were subsequently demonstrated to inhibit parasite invasion of host erythrocytes, with no detectable effect on other parasite processes (as opposed to the general papain family protease inhibitor E-64, which produced multiple aberrations and ultimately developmental arrest). Notably, this ABPP analysis of falcipain 1 function and inhibition was carried out directly in whole parasite lysates, circumventing the need for technically difficult gene ablation experiments and/or recombinant enzyme expression.

Competitive ABPP for discovering potent and selective reversible enzyme inhibitors

While activity-based probes serve as powerful tools for the discovery of enzyme activities associated with discrete physiological and pathological processes, the target promiscuity displayed by these profiling agents limits their utility for defining the biological function of individual enzymes. As described above [35], ABPP has been applied to identify irreversible inhibitors that, for certain enzyme classes such as cysteine proteases, appear to achieve sufficient selectivity to serve as useful pharmacological agents in vivo. However, for most enzyme classes, irreversible inhibitors, by virtue of their inherent reactivity, display poor target selectivity and are therefore less desirable than reversible inhibitors as lead compounds for pharmacological studies and drug development.

To adapt ABPP for the discovery of reversible enzyme inhibitors, Leung and colleagues devised a competitive screen to evaluate the activity of a library of candidate reversible inhibitors against SH activities expressed in mouse tissue proteomes [36]. In this study, proteomes were incubated with a library of candidate inhibitors and an SH-directed probe for a restricted period of time during which the majority of enzymes had not yet reacted to completion with the probe (Figure 2). Under such kinetically controlled conditions, the binding of competitive reversible inhibitors to specific enzymes was detected as a reduction in probe labeling. By performing this screen in mouse brain and heart proteomes using varying inhibitor concentrations, both the potencies (IC_{50} values) and selectivities of inhibitors were determined concurrently. Importantly, calculated IC_{50} values, as measured by ABPP, matched closely with K_{i} values determined by standard substrate assays. Hierarchical cluster analysis of resulting datasets demonstrated that inhibitors selective for individual SHs could be readily distinguished from compounds that displayed comparable or greater activity toward multiple enzymes. Notably, inhibitors were discovered for both known enzymes of therapeutic interest (e.g. fatty acid amide hydrolase) and novel enzymes that lack known substrates. Interestingly, one such enzyme was the mouse orthologue of KIAA1363, the novel invasiveness-related membrane SH identified in the cancer cell line profiling described above. Thus, these studies demonstrate that lead inhibitors for testing the (patho)physiological function of novel enzymes such as KIAA1363 can be identified by competitive ABPP without ever needing to recombinantly express these proteins or develop a specific substrate assay. In summary, competitive ABPP constitutes a versatile, ‘substrate-free’ platform for identifying reversible or irreversible inhibitors of
individual members of large enzyme families. Moreover, because such inhibitors are tested against numerous enzymes in parallel within the context of their native proteomes, promiscuous agents can be readily triaged in favor of equally potent compounds that display high target selectivity.

Conclusions and future challenges
Considering that proteins are the mediators of nearly all biochemical events that underlie cell and organismal (patho)physiology, the need to develop general methods to measure the levels and activities of these biomolecules directly in cell, tissue, and fluid proteomes is apparent. Still, the field of proteomics faces myriad technical challenges that far exceed those encountered by its sister discipline, genomics. While nucleic-acid-based profiling techniques exploit the exquisite specificity of antisense hybridization for detection, most proteins do not possess such readily available high-affinity binding partners. Additionally, unlike nucleic acids, proteins cannot be amplified, thus hindering their detection in samples of limited quantity. Finally, proteins display more diverse biochemical properties than oligonucleotides, meaning that no single experimental protocol may be suitable for the analysis of all proteins. These issues have engendered a new breed of strategies for comparative proteomics, in which the complex array of proteins expressed by a given biological sample are parcelled into manageable fractions before analysis. Toward this end, ABPP (and related chemical efforts) have proven particularly effective, as these strategies focus proteomic experiments on specific classes of proteins. Because activity-based probes, by design, target enzymes based on shared active site features (mechanistic and/or structural), rather than mere expression level, this method provides exceptional access to low-abundance proteins in samples of high complexity.

Moreover, by labeling the active sites of their enzyme targets, activity-based probes afford a level of functional information not attainable with abundance-based proteomic methods, permitting active enzymes to be distinguished from their inactive counterparts. Although the development and application of ABPP has proceeded at a rapid pace, many important challenges remain. First, ABPP is still only applicable to a modest portion of the proteome. In considering future targets for ABPP, enzymes such as kinases [9] and GTPases [37] are notable, as their activities are regulated by a complex variety of post-translational mechanisms in vivo. The successful generation of activity-based probes for additional enzyme (and protein) classes will probably require a combination of both directed and non-directed efforts guided by an ever-increasing understanding of active site structure and reactivity. Second, most ABPP experiments have been conducted with cell and tissue extracts, which can only, at best, approximate the native environment experienced by proteins in vivo. To address this shortcoming, tag-free versions of ABPP have recently been introduced that permit the detection of enzyme activities in living cells [38,39] and animals [38]. Finally, selecting the most appropriate method for the resolution and detection of probe-labeled proteomes remains an open question for which there may be no single correct answer. Comparative ABPP experiments have, in general, employed 1D-SDS–PAGE for protein separation, sacrificing a degree of resolution in exchange for much greater throughput. The value of 1D-SDS–PAGE in these studies is apparent when one considers that, in several cases [31**,36**], to achieve similar data with higher resolution methods would have required the implementation of more than 400 2DE gels and/or LC runs. On the other hand, the analysis of probe-labeled proteomes by higher

**Figure 2**

Competitive ABPP for discovering potent and selective reversible inhibitors [36**]. Complex proteome before and after treatment with a reversible inhibitor library and an activity-based probe. Active enzymes are denoted by open/unshaded active sites, with their inhibitor-bound counterparts shaded in color. FP-rhodamine labeled activity profile: competitive proteomic profiling of a library of candidate SH inhibitors. Shown is a representative in-gel fluorescence analysis of inhibitor-treated mouse brain membrane proteomes labeled with the SH-directed probe FP-rhodamine. Inhibitor-sensitive targets (FAAH and KIAA1363, single and double arrowheads, respectively) are highlighted.
resolution techniques such as 2DE and/or LC may be preferred in cases where detailed information is sought on a limited number of samples. Striking a suitable balance between the breadth and depth of proteomic analysis will probably depend on several variables, including the experimental question being addressed, the number and type of samples under examination, and the amount of available sample. Regardless of whether one aims to rapidly compare the protein profiles of hundreds of cells/tissues or seeks to obtain in-depth information on a few samples, ABPP offers researchers a powerful strategy for the functional analysis of the proteome and its individual constituents.

Acknowledgements
We thank the Cravatt laboratory for helpful discussions. This work was supported by the National Institutes of Health (Grant no. CA87660), the California Breast Cancer Research Program supported by the National Institutes of Health (Grant no. CA87660), the California Breast Cancer Research Program Dissertation Fellowship (NJ).

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest


17. Adam GC, Sorensen EJ, Cravatt BF: Chemical strategies for functional proteomics. Mol Cell Proteomics 2002, 1:781-790. This review highlights chemical approaches for both abundance-based and activity-based proteomics with an emphasis on methods that permit the quantitative comparison of proteins. Concepts underlying directed and non-directed approaches to activity-based probe design are discussed.


22. Faleiro L, Kobayashi R, Fearnhead H, Lazebnik Y: Multiple species of CPP32 and Mch2 are the major active caspases present in apoptotic cells. EMBO J 1997, 16:2227-2228.


30. Jeffrey DA, Bogyo M: Chemical proteomics and its application to drug discovery. Curr Opin Biotechnol 2003, 14:87-95. This review highlights the utility of chemical proteomics for target identification and validation as related to drug discovery and development. A thorough overview of probe design, including structures and mechanisms for various class selective activity-based probes, is also presented.


This paper reports the use of SH-directed activity-based probes to profile a panel of human breast and melanoma cancer cells resulting...
in comparative profiles that integrate activity, subcellular distribution, and glycosylation to distinguish cell lines based upon tissue of origin. A unique set of hydrolases was also identified that correlated with cancer cell invasiveness.


This paper describes the use of comparative ABPP to identify and characterize the role of a specific papain cysteine protease in the malaria parasite life cycle. Irreversible inhibitors selective for this protease were identified by screening compound libraries for their ability to block probe labeling and used to demonstrate a role for the enzyme in parasite invasion processes.


This paper describes the use of competitive ABPP to evaluate the potency and selectivity of reversible enzyme inhibitors directly in whole proteomes. A panel of candidate SH inhibitors was screened against multiple mouse tissue proteomes, and selective inhibitors were distinguished from structurally similar compounds that exhibited poor selectivity.

