MINIREVIEW

Activity-based Proteomics of Enzyme Superfamilies: Serine Hydrolases as a Case Study*

Published, JBC Papers in Press, February 10, 2010, DOI 10.1074/jbc.R109.097600

Gabriel M. Simon and Benjamin F. Cravatt
From the Department of Chemical Physiology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037

Complete genome sequences have revolutionized our view of living systems. The number of genes possessed by organisms ranging from bacteria to yeast to humans is now more or less confidently assigned and has provided a framework for understanding complex cellular and physiological processes at a biochemical level. This framework is, however, plagued by huge knowledge gaps represented, perhaps most notably, by a daunting number of uncharacterized gene products. These include many predicted proteins that lack discernible sequence homology to other proteins of known function, as well as expansive protein families, of which only a modest subset of members have been assigned biochemical activities (1, 2).

Working under the Darwinian assumption that every protein has evolved to perform a unique function that ultimately benefits the host organism, it follows that significant gaps in our knowledge of the proteome imperil ongoing computational and experimental attempts to build molecular networks that explain higher order life processes. This problem is accentuated by the realization that among human protein families containing many poorly characterized members are the fundamental components of signal transduction (receptors, ion channels), gene regulation (transcription factors), and metabolic (enzymes, transporters) pathways. Ongoing and future “systems biology” endeavors would thus greatly benefit from new technologies that facilitate assignment of protein function on a global scale. These technologies can take the form of methods that map fundamental features of protein behavior, including intermolecular (e.g. gene-gene, protein-protein, protein-DNA, protein-metabolite) interactions (3, 4), cellular and subcellular localization (5), post-translational modification state (6), and biochemical activities (7–9). In this minireview, we will focus on the last category in our discussion of the chemoproteomic method activity-based protein profiling (ABPP),2 which aims to globally characterize the functional state of enzymes in native biological systems (8, 9). Our goal is to showcase how, in little more than a decade, ABPP has developed into a versatile platform for enzyme annotation in the genomic era.

Enzyme Analysis by ABPP: Serine Hydrolases as a Case Study

ABPP has been successfully applied to many enzyme classes (8). For the purposes of this minireview, however, we will focus our attention on its use to study serine hydrolases (SHs), which have served as a versatile testing group for ABPP and its various applications. SHs utilize a conserved serine nucleophile to hydrolyze amide, ester, and thioester bonds in both protein and small molecule (metabolite) substrates. They are one of the largest and most widely distributed enzyme classes in all three kingdoms of life, including humans, where SHs constitute ~1% of the proteome (10). These include >100 serine proteases (e.g. trypsin, elastase, thrombin) that hydrolyze principally peptide bonds in proteins, as well as another 110+ SHs that act on metabolites and peptides and are hereafter referred to as mSHs. The mSHs include esterases (e.g. acetylcholinesterase), lipases (e.g. cytosolic phospholipase A2), peptidases (e.g. dipeptidyl peptidase IV), and amides (e.g. FAAH) and can be organized into several distinct subfamilies based on three-dimensional structure and catalytic mechanism. The majority of mSHs (>60%) adopt an α,β-hydrolase fold and employ a Ser-His-Asp catalytic triad (11); however, multiple distinct and evolutionarily unrelated subclasses of mSHs also exist that utilize different folds and catalytic machinery, including the amidase signature enzymes (Ser-Ser-Lys triad) (12, 13) and patatin domain-containing lipases (Ser-Asp dyad) (14). Although many mSHs are well studied enzymes with established biochemical activities, it is striking to realize that nearly half of the human mSHs are completely uncharacterized with respect to physiological substrates and functions (Fig. 1A, blue enzymes). As will be described below, ABPP has proven especially well suited for investigating uncharacterized SHs.

Before proceeding further, however, it is first worth considering the potential evolutionary and functional significance of the large number of unannotated SHs found in the human proteome. Certain enzyme families, such as the cytochrome P450 class, contain many members that are subject to substantial

* This work was supported, in whole or in part, by National Institutes of Health Grants CA087660, CA132630, and DA025285. This work was also supported by the ARCS Foundation and a Koshland Graduate Fellowship in Enzyme Biochemistry (to G. M. S.) and by the Skaggs Institute for Chemical Biology. This is the fourth of six articles in the “Chemical Biology Meets Biological Chemistry Minireview Series.” This minireview will be reprinted in the 2010 Minireview Compendium, which will be available in January, 2011.

1 To whom correspondence should be addressed. E-mail: cravatt@scripps.edu.

2 The abbreviations used are: ABPP, activity-based protein profiling; SH, serine hydrolase; mSH, metabolic SH; FAAH, fatty acid amide hydrolase; FP, fluorophosphonate; MAGL, monoacylglycerol lipase; LC-MS, liquid chromatography-mass spectrometry; 2-AG, 2-arachidonoylglycerol; fluop, fluorescence polarization.
evolutionary drift to the extent that they do not display reciprocal orthology between humans and other mammals (15). These non-orthologous enzymes also tend to exhibit promiscuous and overlapping substrate specificities, suggestive of some degree of functional redundancy. On the basis of a similar evolutionary analysis, we believe that mSHs present a different case, where most members, including those that are uncharacterized, possess conserved non-redundant biochemical functions. A visual depiction of this evolutionary argument is shown in Fig. 1B, where each human mSH is distributed across a two-dimensional plane that compares sequence similarity between 1) the most closely related mouse enzyme (orthology; x axis) and 2) the most closely related (non-identical) human enzyme (homology; y axis). The vast majority of mSHs cluster in the lower right-hand corner of the plot, reflecting both very low homology to their nearest neighbors in humans (indicating functional uniqueness) and high sequence identity with their respective mouse orthologs (indicating strong evolutionary conservation). One striking feature of this plot is that well-characterized enzymes that perform specific biochemical functions in mammalian physiology, such as FAAH and acetylcholinesterase, appear alongside uncharacterized enzymes, such as ABHD13, LACTB, and BPHL. One can extrapolate from this analysis that the ~50% of mSHs that remain unannotated are likely to play just as critical roles in human biology as the 50% of enzymes from this class that are characterized, thus underscoring the need for new functional profiling methods, like ABPP, that can be broadly applied to SHs regardless of their degree of annotation.

Fluorophosphonates as Prototypical Activity-based Probes for SHs

The principal currency of ABPP is chemistry and, more specifically, chemical probes (Fig. 2A) (8, 16). These activity-based probes contain at least two elements: 1) a reactive group to label mechanistically related enzymes in an active site-directed manner and 2) a reporter tag to visualize, enrich, and identify probe-labeled enzymes. Much of the success of ABPP has hinged on the adaptation of classical affinity labels to serve as reactive groups for probe design. For SHs, their catalytic mechanism involves formation of a covalent acyl-enzyme intermediate that has enabled development of activity-based probes that incorporate the fluorophosphonate (FP) affinity label (Fig. 2B). Early studies showed that biotinylated and fluorescently tagged FPs could be used in combination with gel-based readouts to profile dozens of SH activities in complex proteomes (17–20). Contemporaneous studies using epoxide-based probes for cysteine proteases were also described (21, 22). These efforts laid the foundation for future technical advances and biological applications of ABPP that have converged on achieving two major goals: enzyme discovery and inhibitor discovery.

SH Discovery by ABPP

As with most proteomic methods (23), ABPP data are most straightforward to interpret when acquired in a comparative context, where differences in enzyme activities are measured across multiple biological samples. Comparative ABPP has proven particularly fruitful for discovering SH activities that are
MINIREVIEW: Enzyme Superfamily Activity-based Proteomics

SH Inhibitor Discovery by ABPP

As may be surmised from the aforementioned section, a major feature of ABPP is its capacity to acquire functional information on uncharacterized enzymes (e.g. a link between KIAA1363 and cancer, a potential role for ABHD6/ABHD12 in endocannabinoid metabolism). Still, the more complete annotation of these enzymes, including assignment of their endogenous substrates and products and (patho)physiological func-

3 G. M. Simon and B. F. Cravatt, unpublished data.
MINIREVIEW: Enzyme Superfamily Activity-based Proteomics

tions, requires selective inhibitors to perturb their activity in living systems. Indeed, one could make a similar argument for many other SHs that fall into the category of “characterized” enzymes because their functional analysis has often been restricted to in vitro (“test tube”) biochemistry experiments due to a lack of selective inhibitors for in vivo studies. Here, ABPP has found its second major application, viz. as a platform for the discovery and optimization of enzyme inhibitors (18, 22, 44).

Inhibitor discovery by ABPP is possible because small molecules can compete with activity-based probes for binding to enzyme active sites, thereby slowing the rate of probe labeling. These reductions in probe labeling can be detected by either gel-based (22, 44, 45) or LC-MS-based (45, 46) platforms (Fig. 24). Competitive ABPP has several advantages for inhibitor discovery over classical substrate assays. First, enzymes can be screened in their native proteomes without requiring recombinant expression or purification. Second, the activity-based probe serves as a “universal” assay for its enzyme targets regardless of their degree of functional annotation. Thus, inhibitors can be developed for totally uncharacterized enzymes. Third, inhibitors are screened against many enzymes in parallel (all of the protein targets of an activity-based probe present in a test proteome), enabling concurrent optimization of inhibitor potency and selectivity.

Competitive ABPP has played a central role in the development of inhibitors for several mSHs, including the endocannabinoid-degrading enzymes FAAH (44, 47) and MAGL (46, 48), as well as the cancer-associated enzyme KIAA1363 (49). These inhibitors have been used, in combination with phenotypic and metabolic profiling, to confirm the role of FAAH and MAGL in endocannabinoid metabolism (44, 46–48) and to demonstrate that KIAA1363 and MAGL regulate ether lipid (49) and fatty acid (24) metabolism in aggressive cancer cells, respectively. More recently, competitive ABPP has been combined with library screening to identify lead inhibitors for uncharacterized mSHs. These screens can take on multiple forms, including lower throughput gel-based profiling (45) and high-throughput solution assays that use fluorescence polarization (fluopol) for readout of probe-enzyme labeling (50). The latter format, termed fluopol-ABPP, can be used in a 384-well format to screen 100,000s of compounds against an enzyme of interest but does require purified protein.

Conclusions and Future Challenges

Approximately a decade since its inception (17, 21), ABPP has gained considerable use as a technology that bridges the fields of chemistry, enzymology, and proteomics toward the goal of characterizing proteins and pathways in complex biological systems. Here, we have attempted to showcase some of the key attributes of ABPP using the SH superfamily as a case study.

The size and diversity of SHs, combined with the availability of broad-spectrum activity-based probes that target these enzymes, have made them a valuable prototype for ABPP. Using ABPP, investigators have discovered SH activities that are dysregulated in important diseases, such as cancer, as well as selective inhibitors for these enzymes to test their metabolic and (patho)physiological functions. In some of these cases, the enzymes under investigation were previously uncharacterized, underscoring an important feature of activity-based probes, which are generally agnostic to the degree of annotation of their enzyme targets.

Studies to date with SHs have also highlighted many of the future challenges that face the characterization of enzymes by ABPP. First, when applied in isolation, ABPP offers only limited insights into enzyme function. Only through integrating ABPP with other methods, such as emerging proteomic (51–53), peptidomic (54), and metabolomic (55) strategies to discover enzyme substrates in native biological systems, can the actual physiological activities of enzymes be defined. In cases in which this integrated approach has been applied to uncharacterized mSHs, specific metabolic functions have been discovered (e.g. KIAA1363 and its role in ether lipid metabolism in cancer cells (49)). These initial findings provide support that uncharacterized mSHs are likely to possess unique biochemical functions in cells and tissues. In this regard, it is also worth noting that even enzymes that are considered well characterized can offer surprises when studied through the lens of such large-scale profiling methods. For instance, ABPP, in combination with metabolomics, revealed that MAGL regulates fatty acids levels in aggressive cancer cells (24) even though these lipids are not under the control of MAGL in most normal tissues (56). Thus, enzymes can exhibit context-dependent differences in their metabolic functions.

A second challenge relates to the broader use of competitive ABPP to develop enzyme inhibitors. Efforts to date with SHs have yielded considerable fruit, including useful pharmacological tools for multiple characterized and uncharacterized enzymes. Still, we estimate that selective inhibitors currently exist for <20% of the SH superfamily. Expanding our pharmacological coverage of this enzyme class (and others) will likely benefit from emerging ABPP platforms (such as fluopol-ABPP) that enable much larger libraries of small molecules to be screened against enzymes in a high-throughput manner. We should also mention that competitive ABPP examines only the selectivity of inhibitors within the target enzyme class; these compounds could still have additional unrelated targets in the proteome. In cases in which lead inhibitors are covalent, a more comprehensive survey of their proteomic targets can be achieved by offshoots of ABPP in which clickable analogs are used to directly identify inhibitor-modified proteins in proteomes (57, 58).

Finally, we should briefly discuss the extent to which findings on SHs can be applied to other enzyme classes. Of the many lessons learned from ABPP studies of SHs, perhaps none is more important than recognizing the value of high-quality, broad-spectrum, activity-based probes. The FPs (and arylphosphonates) are, in many ways, ideal activity-based probes in that they show broad reactivity across the SH class but little or no cross-reactivity with other enzymes. Similar success stories can be found for a limited number of other enzyme classes (such as the cysteine proteases (59)), but most activity-based probes show either more restricted reactivity within a given enzyme class (41) or labeling of enzymes from different mechanistic classes (60). These features are not necessarily drawbacks, especially with the advent of higher resolution LC-MS platforms for identifying targets of activity-based probes (i.e. there is no a priori reason why an activity-based probe needs to selectively label enzymes from one class as long as all of its targets in the proteome can be resolved). Nonetheless, they do emphasize that chemistry will continue to play a prominent role in advancing and refining new reactive chemotypes for use in ABPP (60).
MINIREVIEW: Enzyme Superfamily Activity-based Proteomics

From a biological perspective, SHs are certainly not special in terms of the large number of uncharacterized enzymes from this class that currently populate eukaryotic and prokaryotic proteomes. One could make a similar argument for virtually all major enzyme families. To the extent that one believes that the ~50% of the proteome that remains uncharacterized is as interesting as (or more than) the ~50% of the proteome that we know something about, ABPP should remain a forefront technology for ongoing and future research aimed at understanding the biochemical basis of life.

REFERENCES