Activity-based probes for functional proteomics

Matthew P. Patricelli

Abstract

Achieving an understanding of the functional role of all proteins expressed by a complex organism will require the development of methods that enable the rapid monitoring of protein function on a global scale. Current genomics and proteomics technologies fall short of this goal since they measure only the relative abundance of transcripts and proteins, respectively. Recent efforts in several laboratories have led to the development of tagged chemical probes that selectively react with families of active enzymes based on shared mechanistic features. These activity-based probes (ABPs) permit the quantitation and comparison of multiple protein activities simultaneously in complex proteomes. In this paper, the general properties and design features of ABPs will be discussed with an emphasis on the use of ABPs for activity-based proteome analysis.

INTRODUCTION

Recent advances in genomic information and technologies have redefined the scale of biological experimentation. Results from genome sequencing projects now suggest that the human genome contains more than 30,000 unique genes which could give rise to as many as 100,000 distinct protein products. These staggering numbers have shifted the focus of new biotechnologies from those that facilitate the characterisation of single genes or proteins, to methods that enable the rapid monitoring of all possible transcripts (genomics)\(^1,2\) or proteins (proteomics)\(^3,4\) contained in a living organism. The fields of genomics and proteomics have quickly advanced to meet the demands of post-genome biology. DNA microarrays, or gene chips, now allow genomics researchers to compare rapidly the relative abundance levels of all the predicted transcripts produced by the human genome. Proteomics strategies have built on advances in mass spectrometry and two-dimensional electrophoresis (2DE) to enable the separation, quantitation and identification of hundreds to thousands of proteins from a single 2DE gel\(^5,6\).

Despite the advances in genomics and proteomics, the ultimate goal of gaining an understanding of the function of all proteins in the proteome is not yet in reach. The use of DNA-based microarrays provides a convenient and comprehensive measure of transcript abundance changes; however, recent evidence suggests that transcript levels do not correlate well with the levels of their corresponding proteins\(^7,8\). The direct analysis of proteins circumvents some of the potential pitfalls of the more indirect genomics approaches, although the most widespread proteomic method of 2DE is relatively slow and cumbersome. Further, while standard genomic and proteomics techniques can accurately measure the relative abundance levels of transcripts and proteins, respectively, they do not succeed in providing a measure of protein activity. The activities of many classes of proteins, including proteases, kinases and phosphatases, are tightly regulated through post-translational forms of regulation. The most common forms of activity altering post-translational regulation are phosphorylation\(^9\) and complexation with endogenous inhibitors or activators\(^10\). These modifications typically occur in the absence of any...
detectable changes in transcript or protein abundance and are thus invisible to traditional genomics and proteomics methods. In order to provide useful information regarding the functional state of proteins from many biologically important protein families, analysis methods ideally would extend beyond measures of transcript or protein abundance and provide a readout of protein activity.

Recently, several groups have developed chemical strategies that permit the direct monitoring of multiple enzyme activities in complex biological samples. A common element of these strategies is the use of tagged chemical probes (activity-based probes, ABPs), to assess the reactivity of enzyme families in crude mixtures. The use of ABPs for functional proteomics provides a promising approach to increasing our ability to assess protein activities on a global scale.

**FROM AFFINITY LABELLING TO FUNCTIONAL PROTEOMICS**

All ABPs developed to date are based to a large extent on the decades-old method of affinity labelling. In its broadest definition, an affinity label is a chemical reagent that reacts covalently and irreversibly with a protein (or other cellular component such as a DNA, RNA or carbohydrate) at a site(s) that is dependent on a specific interaction between the target and the affinity label. While ABPs are affinity labels in the classical sense, they possess several specific properties that permit their use as probes of enzyme activity. First, ABPs by definition react only with proteins, namely enzymes, that have a catalytic activity. Further, the reactivity of an ABP with its target(s) must depend on the catalytic activity or active site accessibility of the enzyme. For these reasons, all ABPs developed to date target enzyme active sites. Finally, all ABPs have a tag (typically biotin) that permits visualisation and selective capture of the target(s) for subsequent identification. Thus, ABPs represent a specific subgroup of affinity labels with carefully selected properties that enable an assessment of enzyme activity. It should be noted that affinity labelling probes that target non-enzymatic proteins such as receptors could also be utilised for functional proteomics. Due to the ambiguity in defining what constitutes the activity of such proteins, however, only probes directed at enzymes will be considered as ABPs for the purposes of this paper.

The use of ABPs for functional proteomics requires a departure from several traditional elements of affinity labelling approaches. In the past, affinity labelling has most often been performed on purified or partially purified protein preparations where the protein of interest is a major species. In the rare cases where affinity labels have been used in crude biological samples, most researchers have focused on designing molecules that specifically target one protein. In contrast, the ultimate goal of any functional proteomics method is the capacity to monitor simultaneously the activities of every protein present in a complex proteome. Thus, a useful ABP must react specifically with its intended targets in an activity-dependent manner, even when the targets are minor components of a complex mixture. This stringent requirement for specificity must be carefully weighed against the fact that the practical value of a probe for functional proteomics is directly related to the number of enzymes that it reacts with in an activity-dependent manner. In principle, provided an ABP reacted only with active protein species, the value of the probe would decrease only when the number of proteins it targets exceeds the limits of resolution for subsequent quantitation and analysis. With the resolution afforded by 2DE, this number would surely be in the thousands.

The desired probe characteristics defined above highlight an important balance that must be achieved between the selectivity of a probe for reaction with...
Activity-based probes for functional proteomics

ABPs have been developed for profiling serine hydrolases and cysteine proteases

ABPs share at three core structural elements:
- Reactive group
- Binding or affinity group
- Tag

Two functional proteomics applications of ABPs:
- Comparative profiling of enzyme activities between biological samples
- Assessment of small molecule-protein interactions in complex biological mixtures

Ideal ABP reactive groups impart ‘chemical selectivity’ for catalytically important residues

enzyme-active sites and its ability to react with numerous targets in an activity-based manner. While these characteristics may seem incompatible, several groups have developed ABPs that fit these criteria and have proven useful for functional proteomics profiling. ABPs have been developed that target serine hydrolases, as well as the caspase and cathepsin groups of the cysteine hydrolases. Each of these three types of probes has been shown to react selectively with active species of multiple, mechanistically related enzymes. Additionally, biotin tags on the probes were used to enrich and identify targets unambiguously from biological samples by mass spectrometry. ABP-like molecules have been described in the literature in addition to the three classes discussed above; however, in these cases, the probes either did not contain a capturable tag for protein identification, and/or they were shown to react only with one enzyme. Some of these examples will be discussed in the final section of this paper.

The published literature reveals two primary proteomics applications for ABPs. First, ABPs can be used in a manner analogous to traditional proteomics experiments to compare the activity levels of proteins between either multiple proteomes or a proteome subjected to multiple varied conditions. The advantage inherent in the use of ABPs for these types of studies is that protein activity changes can be observed that occur in the absence of abundance changes. The second common application of ABPs is the assessment of small molecule inhibitor targets in complex environments. In the presence of an active site-directed small molecule inhibitor, the labelling rates of ABPs are decreased due to competition. The affinity of a small molecule for various enzymes can be attained by measuring the decrease in ABP labelling rates in the presence of the inhibitor. When drug effects are monitored using standard proteomic or genomic methods, only the secondary effects of drug interactions can be observed. The use of ABPs permits the direct visualisation of primary drug targets as well as downstream effectors.

Secondary effects of drug treatment can be distinguished from primary targets by removing the drug prior to ABP labelling. Thus, the use of ABPs in functional proteomics offers significant advantages over abundance-based methods of proteomics and genomics.

STRUCTURE AND DESIGN OF ABPS

To date, the most successful strategy for ABP development has been to design probes that react with enzyme families based on mechanistic features and binding selectivities that are shared among many members. These ABPs share several common structural elements, including a reactive group, a binding group, a linker and a tag (Figure 1). The arrangement and composition of these features vary between different probe classes. The probes in Figure 1 highlight several different strategies for the design of ABPs that result in probes with different structural features. The presence and extent of binding groups and linkers, the nature of the traceable tag and the mechanism of covalent labelling diverge considerably between these three probes.

One common design feature among these three ABPs is the use of well-studied reactive groups that are known to target specifically catalytic nucleophiles in the enzyme-active sites. The fluorophosphonate reactive group used on the FP-Peg-TMR probe (Figure 1C) is a classical affinity label for serine hydrolases that selectively reacts with the activated serine nucleophile of catalytic triad and other serine hydrolase classes. Likewise, the epoxide (Figure 1A) and acyloxy-methyl ketone (Figure 1B) reactive groups used on the DCG-004 and biotin–AOMK probes are known to react very slowly with most cysteine residues, but react rapidly with the nucleophilic cysteine present in the active site of cysteine proteases. The availability...
of such well-characterised chemical groups provides a convenient first step for the generation of ABPs.

When a reactive group is used that imparts an element of 'chemical specificity', the degree of binding elements required to achieve adequate probe selectivity can be minimised. In the most extreme case, the reactive fluorophosphonate moiety in the serine hydrolase probe FP-Peg-TMR is attached directly to a tetra-ethylene glycol linker that is terminated with a fluorescent tag. Despite the absence of any distinct substrate-like binding motifs, FP-Peg-TMR and related ABPs have been shown to target selectively numerous serine hydrolases, including serine proteases, lipases, esterases and amidases. The common functional elements of the probes are annotated. These three probes utilise mild electrophiles to react selectively with nucleophilic active site residues.
Activity-based probes for functional proteomics

employed for quantitation and analysis. Until recently, the most common tracers utilised for affinity labelling reagents were radioisotopes. While radioactive tags offer the advantage of being highly sensitive and direct methods for detection, they provide no means for the selective enrichment of labelled proteins. In order to identify the protein targets of an ABP, it is often necessary to remove unlabelled material from the sample in order to avoid interference from abundant, unlabelled proteins in subsequent mass spectrometry (MS)-based identification methods. When radiolabelled probes are used in complex samples, there is no straightforward or direct method for identifying the specific protein targets of the probe. Since the ability to identify the proteins being monitored by an ABP is of central importance to the concept of functional proteomics, tags must be used that allow for the selective capture and identification of the probe targets.

Several ABPs have been developed that use a biotin label for screening and enrichment.12-15,19 The use of biotin as a tag has the distinct advantage that reagents and procedures are available for the selective enrichment of biotinylated–ABP-labelled proteins. Multiple groups have used avidin beads to capture and subsequently identify ABP-labelled proteins by gel electrophoresis and MS.13,14,19 For screening applications where protein identification is not required, labelled proteins can be visualised at the 10–50 femtomole level using a western blot probed with avidin–HRP conjugates.13 Unfortunately, these indirect detection methods are time consuming and not very quantitative.

A recently developed strategy that incorporates many advantages of radioactive and biotin-based detection methods is the use of fluorescent tags for monitoring ABP reactions in proteomic samples.20 When fluorescent ABPs are used, labelled proteins are visualised directly in polyacrylamide gels using commercially available laser-based scanning instruments. This detection method has been shown to provide approximately 100-fold superior sensitivity to biotin–avidin-based detection, and requires less sample handling than both biotin–avidin and radioactive visualisation procedures. Selective enrichment of fluorescent activity–based probes can be accomplished using immobilised antibodies directed at the fluorophore, yielding purified proteins for MS identification (M. Patricelli, unpublished results).

The creation of ABPs has now been accomplished independently in several laboratories, using somewhat different strategies in each case. Despite the stringent requirements for the use of ABPs in complex proteomes, the structural diversity of these ABPs provides hope that more ABPs will be developed in future efforts to provide even greater coverage of the enzyme proteome.

INCREASING ACCESSIBLE PROTEOMIC SPACE WITH ABPs

As detailed above, most ABPs developed to date have been built on classical affinity labelling moieties which have long been known to target selectively active sites of large enzyme families. Unfortunately, for many important enzyme families, such well-established chemical affinity reagents do not yet exist. The future success of functional proteomics platforms based on ABPs will surely depend on the ability to extend this approach to other important enzyme classes — including large families like phosphatases, kinases and metalloproteases among others. In assessing the potential difficulties in targeting these and other enzyme families, two distinct problems can be predicted. First, the targeting of enzyme families that utilise covalent catalysis but for which there are no known chemical affinity reagents will require the identification of new functional groups that can yield selective ABPs. The second difficulty will be the design of ABPs for enzyme classes that do not have active sites primed for covalent catalysis such as kinases and
Metallohydrolases. For these enzymes, the specific targeting of catalytic residues may prove more difficult and an alternative approach may be necessary.

Recently, Adam et al. have taken a combinatorial approach to the goal of finding ABPs that target new enzyme families. In this study, a library of moderately reactive sulphonate-based electrophiles was attached through an alkyl linker to biotin. These compounds were then screened in proteomic samples for activity-dependent protein labelling. Several reagents were identified that appeared to react covalently with enzymes in native proteomes but which exhibited minimal reactivity in denatured proteomes — a hallmark of activity-based labelling. In particular, a pyridyl-sulphonate-based probe (Figure 2A) was identified as a novel affinity reagent for functionally probing class I aldehyde dehydrogenase. Importantly, this probe also reacted with several other proteins that have not yet been identified. Perhaps most importantly, the reactivity of other probes from the sulphonate library exhibited a high degree of variability, with certain compounds reacting with entirely non-overlapping sets of proteins. Thus, from a single study of one class of reactive group, multiple ABPs were generated that appear to target a number of distinct enzymes. This type of approach holds promise for the development of novel affinity reagents to be used for targeting new enzyme families with ABPs.

Conceptually, the most difficult issue in the design of new ABPs may lie in the targeting of enzyme classes that do not utilise covalent catalysis and thus do not contain unusually nucleophilic active site residues. In order to target enzymes such as kinases and metalloproteases, a fundamentally different approach may be required. One possible solution to this problem is to design ABPs that contain relatively non-selective reactive groups attached to high affinity, reversible binding motifs. Several natural product enzyme inhibitors, including microcystin, fumagillin and wortmannin, act in this manner, suggesting that such an approach is feasible. Ideally, for ABP design, the binding group selected would be a general high affinity inhibitor for numerous enzymes. Additionally, for this method to succeed, the reactive functionality would need to be positioned by the binding group so that it could potentially react with a nearby residue that may or may not be mechanistically important but which is conserved among many enzymes. One potential advantage of this strategy is that it could conceivably be applied even to non-enzymatic proteins such as receptors that contain small molecule- or peptide-binding sites.

Several affinity reagents have been generated and are commercially available.

**Figure 2**: Alternative approaches to ABP design. (A) This pyridyl sulphonate-based probe was a member of a library of potential ABPs that all contained sulphonate-reactive groups. The probe shown was found to target class I aldehyde dehydrogenase as well as several other unidentified enzymes. (B) Reactive co-factor analogues such as γ,δ-P~8-N~7-ATP may prove useful for profiling the activities of enzyme families that do not have unusually nucleophilic active site residues.
that follow the general design concepts discussed above.\textsuperscript{34} In particular, a significant body of research has focused on the use of co-factor analogues containing photo-activatable crosslinking groups and radioactive tags for monitoring co-factor–dependent enzymes in biological samples.\textsuperscript{24,25,35,36} The photo-activatable reactive groups on these probes react with anything in proximity once activated and rely on reversible binding interactions with target active sites to promote specific active site labelling. Using co-factor–like structures as binding groups is particularly attractive in that a large number of enzymes use common co-factors such as ATP and NAD and bind these co-factors with reasonably high affinity. Due to the extreme reactivity of the photo-activated reactive groups, these probes can react covalently with a broad range of amino acid side chains provided they are in close proximity at the time of irradiation. Affinity probes based on photo-activatable co-factor analogues have been shown to react with enzymes such as creatine kinase,\textsuperscript{25} various ATPases\textsuperscript{36} and steroid 5 α-reductase,\textsuperscript{24} among others. In one particularly relevant study, the photo-labelling of creatine kinase by γ-\textsuperscript{32}P-8-N\textsubscript{3}-ATP (Figure 2B) was found to be reduced in the brains from Alzheimer’s disease patients relative to normal brains.\textsuperscript{25} Decrease in labelling correlated with a decrease in enzymatic activity despite the presence of equal levels of the enzyme in both tissues. Thus, the γ-\textsuperscript{32}P-8-N\textsubscript{3}-ATP probe would likely be a useful ABP for functional proteomics if it utilised biotin instead of a radioisotopic tag.

CONCLUSIONS

The potential applications of activity-based probes to proteomics experiments have only begun to be explored. While several ABPs have been described that function effectively in complex proteomes, the direct application to biological problems has, to date, been limited. Creating a general activity-based proteomics platform will require a significant expansion of the proteomic space covered by ABPs. Despite the challenges inherent in designing suitable ABPs, the innovative chemical and analytical strategies detailed here will hopefully enable functional profiling of the entire enzyme proteome and beyond.

References


