Activity Based Probes for Proteases: Applications to Biomarker Discovery, Molecular Imaging and Drug Screening

Marko Fonović[1,2] and Matthew Bogyo[1,*]

[1]*Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, California 94305, USA and* [2]*Department of Biochemistry and Molecular Biology, Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

Abstract: Recent advances in global genomic and proteomic methods have lead to a greater understanding of how genes and proteins function in complex networks within a cell. One of the major limitations in these methodologies is their inability to provide information on the dynamic, post-translational regulation of enzymatic proteins. In particular proteases are often synthesized as inactive zymogens that need to be activated in order to carry out specific biological processes. Thus, methods that allow direct monitoring of protease activity in the context of a living cell or whole animal will be required to begin to understand the systems-wide functional roles of proteases. In this review, we discuss the development and applications of activity based probes (ABPs) to study proteases and their role in pathological processes. Specifically we focus on application of this technique for biomarker discovery, *in vivo* imaging and drug screening.

Key Words: Proteases, activity based probes, proteomics, biomarkers.

1. INTRODUCTION

Almost 300 complete genome sequences have been published in the last five years and more than 1000 eukaryotic and prokaryotic sequencing projects are currently underway [1]. While raw sequence data is highly informative it often fails to provide insight into the functional roles of specific gene products. Furthermore, differences in splicing and post-translational modifications of a single gene can give rise to a variety of products leading to a high degree of complexity. For example it is thought that the 40,000 human genes can generate greater than 1 million distinguishable functional protein entities. To begin to address this complexity the field of proteomics was established with the important goal of analyzing the functional regulation of all proteins in a given proteome.

Over the past 20 years the field of proteomics has seen many advances in technology, specifically this area of research has flourished with the advent of user-friendly and affordable mass spectrometers. Interestingly, one of the most commonly used proteomics methods remains one that was developed in the early days of the field, namely two dimensional gel chromatography coupled to mass spectrometry (2D/MS). Historically, 2D gel chromatography has been the separation method of choice for resolving complex protein mixtures as it can be performed easily without the need for extensive expertise or special equipment. However, this analytical method has its limitations. It is not suitable for determination of extremely large or small proteins, very basic or acidic proteins and membrane proteins. It also has limited ability to detect low abundance proteins and has limited quantification capabilities. To resolve some of these critical issues, additional approaches such as SELDI (solid phase laser induced ionization) [2] and MudPIT (multidimensional protein identification technique) [3] have been developed. Both of these methods utilize protein separation with various chromatographic resins prior to MS analysis. Separation can be performed either on a chip (SELDI) or on a multi-dimensional liquid chromatographic (LC) column (MudPIT). Chromatographic separations significantly increase resolving power and allow detailed analysis of sub-populations of proteins within a complex proteome. In addition quantification can be achieved by coupling isotope labeling of the proteome to direct LC analysis (ICAT and SILAC) [4, 5].

While the proteomic methods described above allow for global analysis of protein abundance they do not provide information about the regulation of enzyme activity. Most enzymes including proteases are often tightly regulated on a post-translational level leading to a potentially significant divergence of abundance and activity. To begin to address some of these limitations chemical or activity based proteomics has been established to characterize protein activity and provide a means to directly monitor functional regulation in complex proteomes (Fig. **1**). This technique uses small-molecule activity based probes (ABPs) that covalently modify key active site residues through a highly specific chemical reaction that depends on enzyme activity. These probes, due to their high degree of selectivity, can be used in complex samples such as cell lysates, intact cells and even whole organisms. In this review we describe recent advances and strategies for activity based proteomics of proteases and their potential impact on target and drug discovery.

2. DESIGN AND SPECIFICITY OF ACTIVITY BASED PROBES THAT TARGET PROTEASES

Structurally, all ABPs have a shared basic design that makes use of elements for targeting, subsequent chemical modification and detection of labeled products. Perhaps the

*Address correspondence to this author at the Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA; Tel: (650)-725-4132; Fax: (650)-725-7424;
E-mail: mbogyo@stanford.edu

1381-6128/07 $50.00+.00

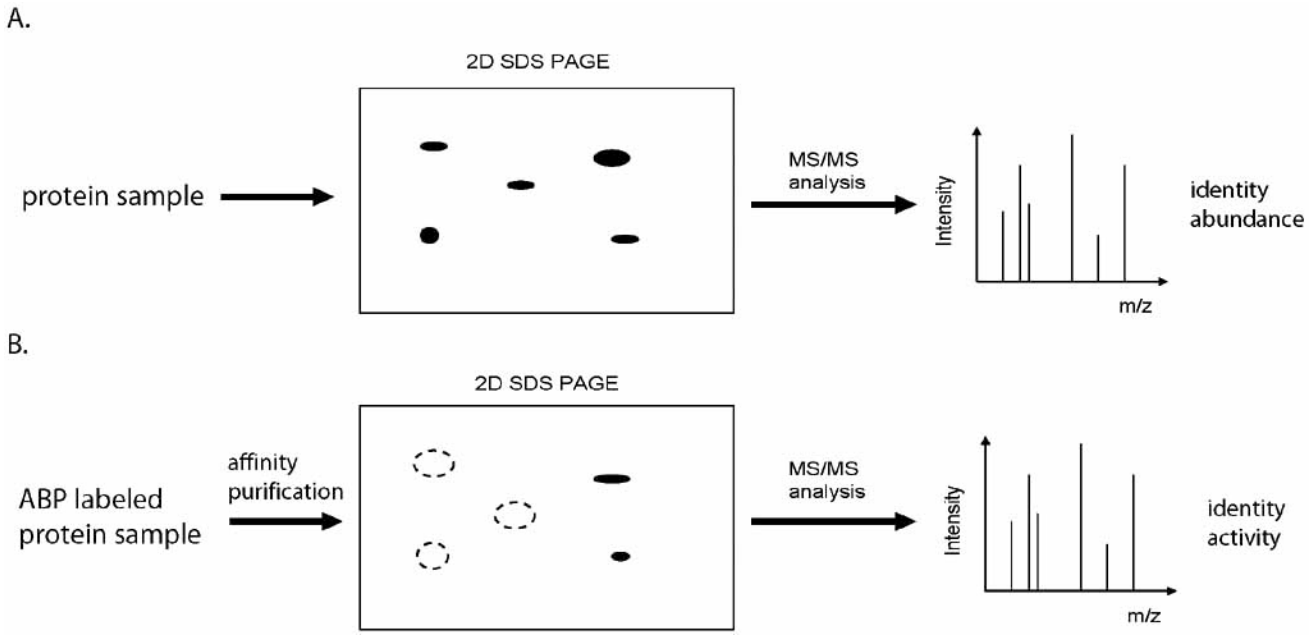

Fig. (1). Comparison of abundance based proteomics and activity based proteomics.

A. Standard proteomic analysis using 2D SDS-PAGE coupled to a mass spectrometry. In this method total proteins are separated by 2D SDS-PAGE, proteins are visualized by silver or Commassie stain and individual protein spots are excised, digested with trypsin and identified by mass spectrometry. This methodology provide identification and quantification of overall protein abundance. B. The activity based proteomics approach. A complex proteome is labeled with an active site probe allowing specific isolation of active targets by affinity chromatography. Unlabeled proteins are removed from the sample (dashed circles) and intensity of labeled protein bands correlates to their activity rather than abundance. Only active protein species are excised and analysed by mass spectrometry.

most crucial component of an ABP is the reactive functional group, sometimes referred to as a "warhead". This reactive group covalently binds to amino acid residues of the target enzyme (Fig. **2**). The majority of activity based probes use electrophilic groups that are able to form a covalent bond with key catalytic nucleophiles located in the active site of the enzyme. These warhead groups must be chemically reactive towards an activated nucleophile of the target enzyme but not reactive enough to modify other free neucleophiles (i.e. cysteine or glutathione) in the proteome. Targeting of an ABP to a specific subset of enzymes is often accomplished by using a variable region that mimics a true substrate. In the case of proteases, this region is almost always a peptide which can be tailored to bind distinct sets of targets by varying the sequence of natural or non-natural amino acids. In addition this recognition region is often used to separate the reactive functional group from the tag to allow accessibility and reduce steric hindrance from the often bulky labeling group. A variety of tagging groups have been used in the design of activity based probes. These reporters usually allow either direct detection or isolation of the labeled enzyme or both. Biotin, I^{125} and various types of fluorophores are most commonly used as tags. However, some types of tags are not cell permeable and therefore cannot be used for ''*in vivo*'' labeling. Furthermore many of the cell permeable fluorescent tags are bulky and may interfere with binding of probes to target enzymes. To address these issues, an approach was recently developed where tags are chemically linked to the probe after labeling and lysis of cells. This strategy relies on the bio-orthogonal ''click'' chemistry in which copper catalyzes a cycloaddition reaction between an alkyne group on the probe and an azide group on the tag [6]. Using this method, smaller, cell permeable probes can be used allowing proteins to be labeled in their native environment in intact cells [7].

In general ABPs have been applied to profile enzyme activities using two different approaches. The directed approach involves the design and synthesis of probes that target a specific enzyme class based on knowledge of catalytic mechanism. The second method is focused on the design of generally reactive probes that can be used to search for target classes that have interesting regulatory patterns (i.e. in disease progression). Both methods provide a means to simplify the proteome to a set of relevant target enzymes that can then be studied in more detail.

2.1. Directed Approach

Activity based probes have been specifically designed to target enzyme families with well-established catalytic mechanisms, including proteases [8-11], kinases [12, 13], and phosphatases [14]. In most cases probe design has taken advantage of mechanism based inhibitors that have been developed by biochemist and medicinal chemists. In particular the abundance of structural and kinetic data on many classes of proteases has enabled the design of small molecules with specific irreversible inhibitory activities [15]. By using known specific synthetic inhibitors as a template, activity based probes that target different types of protease classes have been designed (Fig. **2**).

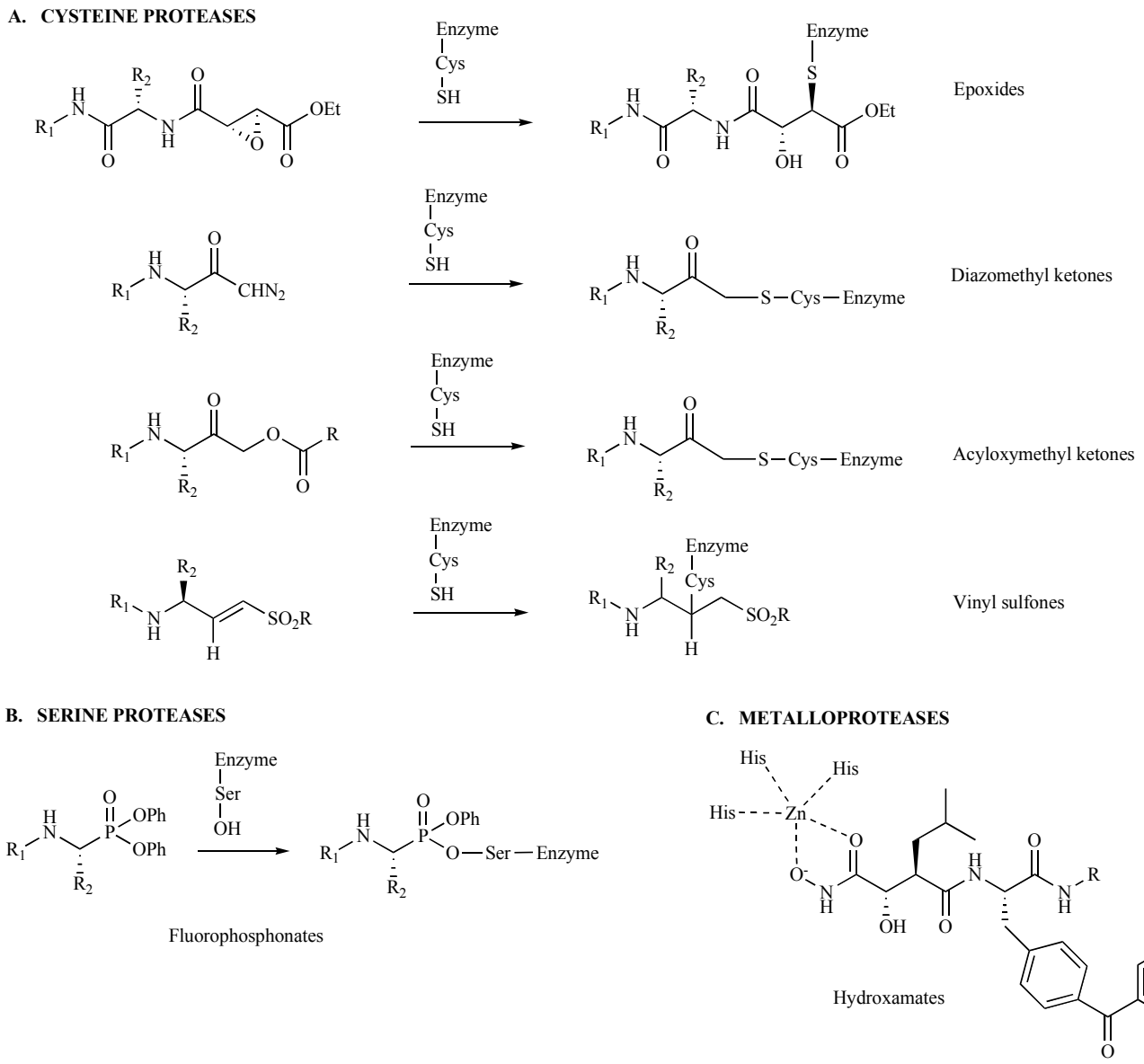

Fig. (2). Structures of protease ABPs and their chemical reaction mechanisms leading to activity dependant target modification.

A number of diverse protease directed probes have been developed. **A**. Covalent labeling mechanism of ABPs degined to specifically target cysteine proteases. The reactive functional groups shown display highly specific reactivity towards the cysteine nucleophile. **B**. Covalent labeling of a serine protease by a peptide diphenyl phosphonate ester. The phosphonate reactive group is highly selective for the serine nucleophile of a serine protease. **C**. Affinity labeling of metalloproteases using a tight binding peptide hydroxamate probe. Since metallo proteases do not use a covalent attack mechanism by an active site amino acid, probes must carry an alternatively reactive functional group that can target residues within the active site. In this example a photocrosslinking group is used to covalently link the probe to the target enzyme.

Probes for Cysteine Proteases

Cysteine proteases represent one of the four primary mechanistic classes of proteases. Their common feature is their catalytic mechanism: they all utilize a Cys residue as the primary nucleophile and a His residue as the general base for proton transfer. They have been further divided into six clans based on the structure of the active site. The papain family of cysteine proteases is a member of clan CA and has been extensively studied due to its involvement in various important physiological and pathological processes [16-18]. E-64 is a natural product inhibitor containing a reactive epoxide warhead that is highly specific for papain family cysteine proteases [19]. Since its discovery in 1976, it has been widely used for studies of cysteine proteases and has been the central structural element in several classes of activity based probes [8, 20-22]. The E-64 derived probe, DCG-04 has been used to determine functional roles of papain-like proteases in processes such as: tumor progression [23], cataract formation [24], prohormone processing [25], parasitic invasion [26] and cell cycle regulation [27]. In addition to E-64-derived epoxides, other reactive groups have also been used for papain like protease profiling: diazomethyl ketones [28], acyloxymethyl ketones [29-31] and vinyl sulfones [8].

Clan CD is the second most abundant clan of cysteine proteases. Its members include caspases, gingipains, legumains, clostripains and separases. Among this family, caspases are of particular interest due to their crucial role in apoptosis. Not surprisingly the first generation of ABPs that target CD clan proteases were directed towards caspases. In fact an activity based probe was used to identify the first caspase (caspase 1, β-interluken converting enzyme -ICE) [32]. Initial caspase probes used acyloxymethyl ketone (AOMK) and aldehyde reactive groups coupled to a specific peptide sequence and a biotin tag. Acyloxymethyl and chloromethly ketone based probes have also been used to study separase and its role in cell division [33, 34]. Recently, AOMK based probes that target diverse CD clan of cysteine proteases have been reported [30].

Probes for Serine Proteases

The serine hydrolase family represents approximately 1% of human genome and includes numerous proteases, lipases, esterases, amidases and transacylases. Like cysteine proteases, serine proteases participate in many important cellular processes, such as blood coagulation [35], apoptosis [36], neurotransmitter catabolism [37, 38], cancer [39] and protein maturation [40]. The majority of serine hydrolases are irreversibly inhibited by fluorophosphonates (FP) thus prompting its use as a general activity based probe for this family. A ^{3}H-labeled version of diisopropyl fluorophosphonate (DFP) is commercially available however, this probe lacks sensitivity and also cannot be used to isolate and identify target enzymes. To circumvent these problems, a series of analogs of DFP containing aliphatic or PEG based linkers and a variety of tags have been developed [9, 10, 41]. These FP probes have been used in a range of profiling experiments, most notably in studies of serine hydrolase biomarkers for cancer [see below 10, 42, 43]. Additionally, peptides modified with a c-terminal phosphonate have been developed as selective inhibitors of serine proteases [15] and probes containing di phenyl phosphonate esters were shown to function as effective probes of typsin family proteases [44] as well as for a viral serine protease [45]. These probes contain a peptide scaffold and therefore react only with serine proteases and not other members of the serine hydrolyase family.

Metalloproteases

In contrast to cysteine and serine proteases, metalloproteases do not use an amino acid nucleophile for direct covalent attack on a substrate but rather use a zinc activated water molecule for peptide bond hydrolysis [46]. Thus, metallo proteases do not form acyl-enzyme intermediates and cannot be targeted using simple electrophiles that mimic a protein substrate. Although some examples of electrophile-based covalent inhibitors for metalloproteases were reported, they were not suitable for probe design because of their weak potency and lack of selectivity [47-49]. Alternatively, activity based probes that are capable of labeling metalloproteases were designed using a tight binding peptide hydroxamate scaffold carrying a label and photocrosslinking group [11]. These probes selectively label metalloproteases after irradiation with UV light. The primary disadvantage of this type of ABPs is that they are not suitable for use in living cells and whole animals.

2.2. Non-Directed Approach

For many enzyme classes, specific reactive functional groups have not been developed thus limiting the ability to use a directed approach. In order to expand the range of enzyme families that can be targeted using ABPs a non-directed or combinatorial approach was introduced [50, 51]. In this approach, a library of candidate probes was synthesized using a simple alkyl scaffold carrying a series of generally reactive electrophiles. Probes were composed of a variable alkyl/aryl binding group, a sulfonate ester reactive group, an aliphatic linker and a rhodamine or biotine tag. A carbon electrophile (sulfonate ester) was selected as a reactive group because many carbon electrophile bearing natural products are known to covalently bind to multiple enzyme classes [52-54]. Libraries of probes were screened in direct labeling assays using a diverse set of tissue and cell proteomes. Probe labeled protein bands that were sensitive to heat denaturation prior to labeling were scored as positive hits. Members of at least nine distinct mechanistic enzyme classes were detected by this approach. Among them were dehydrogenases, kinases, hydratases and transferases. Interestingly, none of the enzymes targeted by the sulfonate library was targeted by previously described classes of specific activity probes. In addition, sulfonate probes labeled various active site residues (aspartate, cysteine, glutamate, tyrosine) and in some enzymes, reacted with non active site amino acids [55].

A library of ABPs containing the α-chloroacetamide (α-CA) reactive group coupled to a variable dipeptide binding group was also developed for non-directed activity based profiling [56]. The α-CA electrophile was selected for two reasons: first, it is relatively small resulting in minimal influence on non-covalent probe-protein interactions and second, like other carbon electrophiles (sulfonate esters, epoxides) the α-CA can label a variety of active site residues. Complex proteomes from tissue samples and cell lines were labeled with the α-CA probe library resulting in the identification of more than 10 different enzyme classes as targets. Interestingly, the overlap of targets with the sulfonate ester library was low, suggesting that both libraries target unique subsets of the proteome.

3. APPLICATION OF ACTIVITY BASED PROFILING

3.1. Identification of Biomarkers and Target Discovery

Completion of the human genome sequencing project has intensified the search for biomarkers of human disease. Genomic technologies such as DNA microarrays and single nucleotide polymorphism (SNP) analysis have been leading the way in the discovery of gene sequences linked to human diseases. DNA microarrays are an efficient method that provides comparative analysis of the entire complement of mRNA populations within a biological sample in a single experiment. Unfortunately, many pathological conditions result from interactions and processes that take place on the protein level. In contrast to mRNA, proteins cannot be amplified and they exhibit a high level of heterogeneity due to various post-translational modifications. Activity based profiling coupled with mass spectrometry has proved to be an excellent tool for identification of protein-based disease markers in complex proteomes (Fig. **3**).

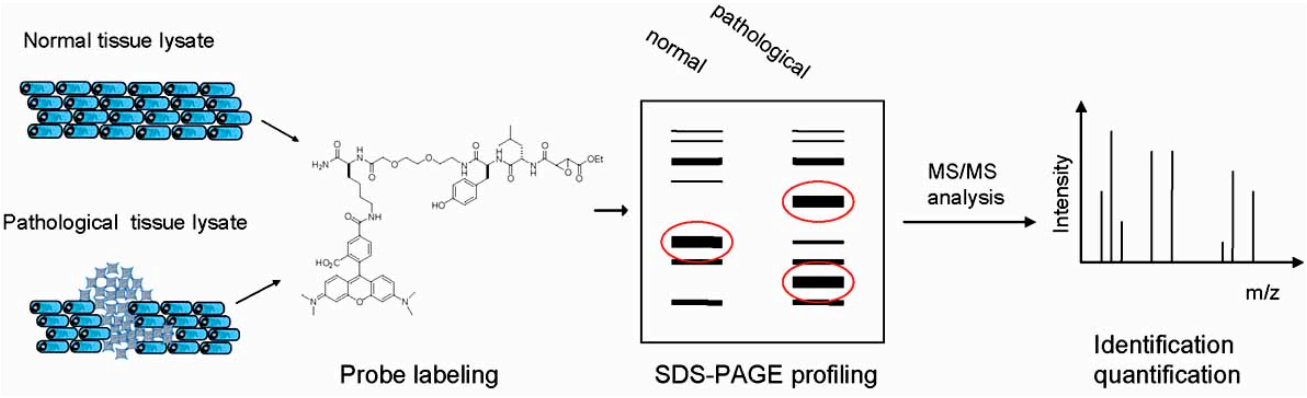

Fig. (3). Application of activity based profiling to biomarker discovery.

Activity based probes can be used to profile complex proteome to identify enzymatic targets that are regulated in disease progression. In this example normal and pathological tissue proteomes are labeled by a fluorescently tagged ABP followed by analysis by SDS-PAGE. Probe labeled proteins are visualized by laser scanning of the resultant gel and proteins whose activities are differentially regulated in pathological and normal proteome can be identified. These targets can then be identified by mass spectrometry and may serve as potential biomarkers for disease.

Some of the first efforts to use ABPs for biomarker discovery have made use of a fluorescent serine hydrolase reactive probe FP-Rhodamine to profile activities in a series of cancer cell lines [42]. Probe labeled samples were separated by SDS-PAGE and fluorescently labeled proteins visualized by laser scanning of the resulting gel. In a parallel experiment, proteomes were labeled by the biotinylated version of the probe and target proteins isolated by affinity chromatography and identified by mass spectrometry. Using the resulting activity profiles of a series of serine hydrolase targets, cancer cell lines were classified into functional subtypes based on tissue of origin and state of invasiveness. Different populations of serine hydrolases were found to be expressed in cell lines with different cellular invasiveness properties. A similar approach was also used to profile hydrolases found in different stages of breast cancer [39]. Serine hydrolase specific probes (fluorophosphonate reactive group) and non-directed probes (sulfonate ester reactive group) were used to study the differences in activity profiles of MDA-MB-231 human breast cancer cells, when grown in cell culture and after tumor formation in mouse mammary fat pads. These studies showed that cancer cells exhibited different activity profiles when grown *in vitro* (in culture) and *in vivo* (xenograft tumors). More than seven types of enzyme activities with distinct expression patterns were identified. These findings suggest that studies performed with human cancer cell lines in culture, may not be predictive of the behaviour of these lines *in vivo*. It was also noted that many dramatic alterations in enzyme activities occurred as a result of post-transcriptional events, again confirming the value of activity based profiling methods.

In another example of an application of ABPs to search for diagnostic markers hydroxamate based probes for metalloproteases were used for profiling of an invasive melanoma cell line [11]. Neprilysin, a membrane associated metalloprotease was found to be highly upregulated in melanoma cell lines even though it is known to degrade several mitogenic peptides and is considered to be a negative regulator of tumorigenesis [57]. This finding suggested that in some tumor types, neprilysin may also contribute to the progression of cancer. Neprylisin was also shown to be a good target for the matrix metalloprotease inhibitor GM6001 (ilomostat) which is currently in clinical trials for cancer [58]. Many MMP inhibitors failed in clinical trials due to toxicity that may be related to their broad reactivity towards different classes of metalloproteases whose functions are poorly understood. These studies highlight the application of ABPs for identification of drug 'off targets' in complex cellular and animal models.

Biomarkers for additional pathological states, such as obesity, have also been studied using ABPs. The generally reactive α-chloroacetamide based probe was used to study differentially expressed proteins in lean and obese mice [56]. Several distinct enzyme activities were identified, among them hydroxypyruvate reductase implicated in the biosynthesis of glucose from serine was found to be six fold upregulated in obese mice. This suggest, that nonclassical pathways of glucose biosynthesis may play a part in obesity related syndromes such as type II diabetes [59, 60].

Activity based profiling has also been used to identify proteases involved in the life cycle of malaria parasite *Plasmodium falciparum*. The papain family-specific probe DCG-04 was used to identify falcipain 1 as a protease that is upregulated in the invasive merozoite stage of *P. falciparum* growth [26]. This finding suggests that falcipain 1 may play an important role in host cell invasion, making it an interesting target for antimalarial drugs.

For most of the activity based probes developed so far only a limited group of enzymes in a specific enzyme class can be targeted at once. This is problematic when attempting to monitor large families of enzymes such as the metalloproteases. To overcome this problem, Cravatt and co-workers have a developed a strategy in which small libraries of structurally diverse photoreactive hydroxamate probes with complementary metalloprotease selectivity are used to profile

tissue proteomes. Following SDS-PAGE analysis, a set of optimal probes can be selected that provides the greatest coverage of the family. This optimal set of probes can then be used for more extensive LC/MS based analysis in diverse biological samples. In the initial application of this approach more than 20 metalloproteases were identified, including members from nearly all the main branches of this enzyme class [61].

Taken together these examples of applications of ABPs strongly support the utility of activity based profiling in the process of drug discovery by providing a means to identify and validate targets in the initial stages of the process. ABPs allow direct labelling of target enzymes in crude proteomes thereby highlighting potential targets for selection based on the correlation of activities within a given disease pathology (Fig. **3**). ABPs can then be used for detailed biochemical and biological assays to validate their roles in disease progression.

3.2. Inhibitor Discovery

High throughput inhibitor screens are important tools of the pharmaceutical industry for discovery of novel therapeutic agents. Generally, screens are substrate-based assays which rely on either purified native enzymes or enzymes from a recombinant source. They are therefore dependant on the ability to express, isolate and purify an enzyme target in a form that is capable of processing the reporter substrate. However, such '*in vitro*' assays provide only limited information about *in vivo* potency and selectivity of a compound for a series of related enzymes. Furthermore, enzymes that cannot be expressed in recombinant form or isolated from tissues in sufficient quantities and enzymes with unknown substrates and reaction conditions, can not be screened for drug leads in such assays.

ABPs have been used to develop small molecule inhibitor screens that solve many of the shortcomings of standard *in vitro* kinetic assays. Since ABPs bind directly to a catalytic active site residue of a target enzyme, they can be used to measure small molecule inhibitor binding using a competition assay. Furthermore, the screen can be performed in complex cellular mixtures that contain multiple related target enzymes. Both potency and selectivity of inhibitor binding can be quantified by monitoring loss of probe labeling as a result of titration of a small molecule inhibitor into the sample (Fig. **4**). Using this method enzymes can be screened in their native environment, thereby eliminating the need for recombinant expression, purification and development of specific substrate assays. Furthermore, activity based profiling allows inhibitor screening of more than one target within a specific proteome in a single experiment.

Important information about target enzyme specificity can also be obtained by screening of positional scanning libraries (PSLs) of inhibitors. This approach has been used to profile specificities of multiple papain family proteases as well as the multiple active sites of the proteasome [62, 63]. The library screening approach allows the contribution of each amino acid residue in the inhibitor to be determined individually. By screening PSLs against multiple related targets it is possible to generate an 'affinity fingerprint' for each enzyme. This information can then be used to create a

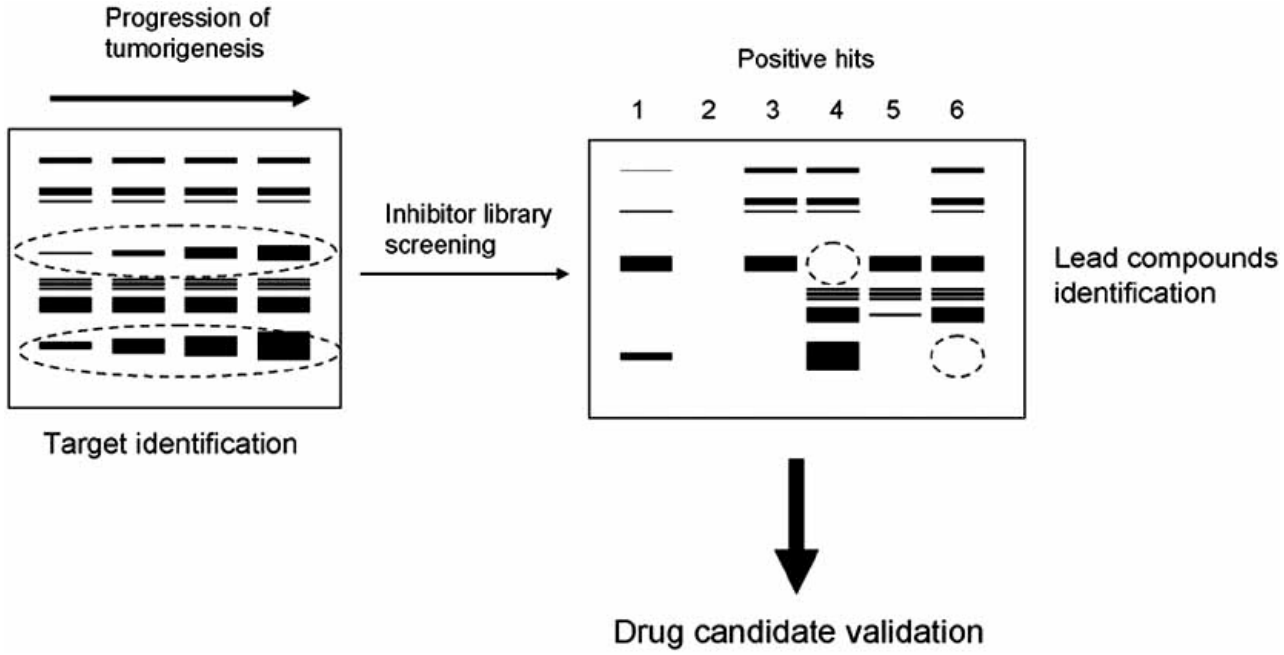

Fig. (4). Application of activity based probes in drug discovery.

ABPs can be used to profile protease activities during disease progression to identify enzymes that are likely to contribute to disease phenotype. Initially staged sample proteomes are directly labeled using an ABPs. Labeled target proteins whose activities are correlated with disease progression can then be identified (left panel dashed circles). Subsequently, libraries of lead compounds (often based on the probe structure) can be screened to identify compounds that target specific enzyme activities linked to disease progression. Compounds that bind to the active site of a target prevent labeling by the ABP and are seen as a loss of labeling (inhibitors 4 and 6 right panel dashed circles). Lead compounds can then be used to validate the target enzyme and can serve as lead compounds for drug development.

database of reference affinity patterns for known enzymes. Such a database facilitates classification of unknown proteases from complex proteomes by comparison of their affinity fingerprints to the reference database. In addition, combining specificity information for each residue enables the design of optimized selective inhibitors. An example of this approach is the recent development of selective caspase inhibitors which were used for the study of caspase activation kinetics during apoptosis [64]. A significant technical advance in the high-throughput screening of chemical libraries "*in vitro*" was also made with the development of an enzyme microarray technology. This method allows rapid analysis of inhibition kinetics of enzyme targets arrayed on a chip and uses activity based probes for the final readout of activity [65]. Using this technique it is possible to measure inhibition kinetics of multiple combinations of targets and inhibitors using extremely small quantities of each.

So far, activity based profiling has been applied mostly for identification of irreversible enzyme inhibitors. Although irreversible inhibitors are useful experimental tools, they are less desirable as compounds for drug design because of their inherent reactivity. While ABPs are covalent inhibitors and therefore best suited for analysis of irreversible inhibitors, it is possible to optimize competition assays to allow assessment of reversible inhibitor binding. Fluorophosphonate-based probes that target serine hydrolases have been used for screening of mouse proteomes with a reversible inhibitor library [66], demonstrating the utility of this approach.

While activity based probes have value in the early stages of drug development, they can also be utilized in the costly late stages of drug development. This is the stage when drug candidates are selected for initial animal studies for determination of general toxicity and pharmacodynamic properties. Since ABPs can be used for *in vivo* competition and imaging studies, they can provide important information regarding specificity and potency of the drug candidate in a whole animal. Candidates with unfavorable side effects can than be screened for cross-reactivity with related enzyme targets. Such approaches enable quick identification and elimination of drug candidates likely to fail later in development thereby potentially limiting losses associated with compound attrition.

3.3. *In Vivo* Imaging of Enzyme Activities

Protein activity inside the living cell is regulated on many levels by factors such as spatial and temporal expression, small molecule or cofactor binding and posttranslational modification. *In vitro* approaches are usually not capable of reproducing complex intracelullar conditions. Therefore, methods that allow *in vivo* visualization of enzyme activity in intact cells or whole organisms would provide a much better understanding of biochemical and physiological processes. Two main approaches have been taken toward imaging of proteolytic activity *in vivo*. The first makes use of reporter substrates that produce a fluorescent signal when processed by a protease while the second utilizes fluorescently labeled ABPs that directly label active proteases [67].

In the simplest of methods, protease imaging substrates are composed of synthetic peptides attached to a fluorogenic or colorimetric molecule that is released only after proteolytic processing by a protease. Such substrates were used for detection of caspase activity in apoptotic cells [68]. Activity was visualized by detecting the fluorochrome or dye molecules that were released after substrate cleavage. The main limitation of this approach is high background signal which results from illumination of cells with UV light. An improved method makes use of quenched fluorogenic substrates, utilizing fluorescence resonance energy transfer (FRET). FRET-based probes contain two fluorochromes situated less than 100 Å apart. The emission wavelength of the donor fluorochrome overlaps with the exitation wavelength of acceptor, allowing transfer of the donor emission to the acceptor without radiation. Proteolytic cleavage of the linker between two fluorochromes changes the fluorescence intensity, which can be monitored and used as an indirect readout of protease activity. FRET-based reporters eliminate many drawbacks of the direct fluorogenic substrates by eliminating artifacts associates with variations in probe concentrations and cell thickness. The main disadvantage of all substrate based imaging methods is lack of specificity toward enzyme targets. Peptide scaffolds of the probes can usually serve as a substrate for more than one protease. Furthermore, a lack of cell permeability and rapid diffusion of the reporter limits their use in high resolution localization studies.

ABP based approaches have the potential to provide better selectivity since the reactive group tends to be highly specific for specific protease classes and can be varied easily to target distinct subsets of proteases. Although cell permeability can often be an issue, many ABPs can freely penetrate cell membranes due to their hydrophobic nature. Furthermore, since ABPs directly modify an enzyme target through formation of covalent bond, any signal observed in a whole cell or organism can be traced back to the enzyme by direct biochemical analysis.

Multiple examples of applications of ABPs for visualization of proteases in intact cells have been reported. Fluorescent ABPs have been used for monitoring of caspase activity after induction of apoptosis in variety of cell lines [69, 70]. Fluorescent ABPs that target papain family cysteine protease were also used for *in vivo* imaging studies in a mouse model for pancreatic cancer [23]. In this study a fluorescent ABP was intravenously injected into a mouse. After several hours, tissue samples were collected and protease activity visualized by fluorescence microscopy. After imaging, SDS-PAGE analysis of lysed tissue revealed the activity profiles and identities of labeled proteases. In addition, treatment of animals with broad-spectrum inhibitors could be directly monitored by visualization of loss of protease labeling in tissues. This application has the potential to significantly improve the process of drug testing in animal models by allowing direct assessment of pharmacodynamic properties of lead compounds.

One of the major limitation to the application of fluorescent ABPs to cell-based imaging applications is that they produce a fluorescencene signal both when bound to a target enzyme and when they are free in solution. To address this limitation and enable direct real time studies of protease activity in live cells, quenched activity based probes (qABPs) have been designed [31]. In this approach, a quencher group was placed in close proximity to the fluorophore, thereby

preventing fluorescence emission. When the qABP covalently binds to the target enzyme, the quencher group is lost as part of the leaving group resulting from the nucleophilic substitution reaction with the active site residue. qABPs therefore only emit fluorescence when they are bound to their target enzyme. These probes have been successfully applied for real time imaging of cysteine protease activity in living cells and have great potential for whole-body imaging applications.

4. CONCLUSIONS AND FUTURE DIRECTIONS

With the techniques of activity based proteomics now firmly established, future efforts in this field will shift towards broadening the technology. The available spectrum of activity based probes will be expanded to allow profiling of additional enzyme classes as well as individual targets within a specific class. In addition, development of new classes of structurally diverse probe libraries will inevitably lead to identification of new probe scaffolds. Applications of activity based profiling to direct mass spectrometry methods is likely to lead to methods that will improve the limit of detection and omit the time consuming 'in gel' analytical methods of current standard proteomic techniques. The first steps in this direction have already been reported using biotinylated serine hydrolase probes for enrichment of target proteins that could be identified by MudPit LC-MS/MS [43]. The advent of click chemistry applications to activity based probes will enable even more accurate target profiling and this approach had already been used for development of small-molecule cell based screens [71]. Sensitive fluorescent reporter groups make ABPs potentially useful tools for '*in vivo*' analysis and quantification of protease activity using highly sensitive techniques such as CE-LIF (capillary electrophoresis with laser induced fluorescence). This platform was recently used for profiling of serine hydrolyases and cysteine proteases in various murine proteomes [72]. The high degree of reproducibility and low detection limits, makes this approach an important tool for pharmacokinetic and pharmacodynamic studies of drug candidates.

ACKNOWLEDGEMENTS

The authors would like to thank members of the Bogyo Lab for critical evaluation of the manuscript. This work was supported by a NIH National Technology Centers for Networks and Pathways grant as part of the Center on Proteolytic Pathways grant #U54 RR020843.

REFERENCES

References 73-75 are related articles recently published in Current Pharmaceutical Design.

[1] Medina M. Genomes, phylogeny, and evolutionary systems biology. Proc Natl Acad Sci USA 2005; 102(Suppl 1): 6630-5.

[2] Seibert V, Ebert MP, Buschmann T. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. Brief Funct Genomic Proteomic 2005; 4(1): 16-26.

[3] Chen EI, Hewel J, Brunhilde Felding-Habermann B, Yates JR 3rd. Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). Mol Cell Proteomics 2006; 5(1): 53-6.

[4] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999; 17(10): 994-9.

[5] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, *et al*. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 2002; 1(5): 376-386.

[6] Speers AE, Adam GC, Cravatt BF. Activity based protein profiling *in vivo* using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. J Am Chem Soc 2003; 125(16): 4686-7.

[7] Speers AE, Cravatt BF. Profiling enzyme activities *in vivo* using click chemistry methods. Chem Biol 2004; 11(4): 535-46.

[8] Bogyo M, Verhelst S, Bellingard-Dubouchaud V, Toba S, Greenbaum D. Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. Chem Biol 2000; 7: 27-38.

[9] Liu Y, Patricelli MP, Cravatt BF. Activity-based protein profiling: the serine hydrolases. Proc Natl Acad Sci USA 1999; 96(26): 14694-9.

[10] Kidd D, Liu Y, Cravatt BF. Profiling serine hydrolase activities in complex proteomes. Biochemistry 2001; 40(13): 4005-15.

[11] Saghatelian A, Jessani N, Joseph A, Humphrey M, Cravatt BF. Activity-based probes for the proteomic profiling of metalloproteases. Proc Natl Acad Sci USA 2004; 101(27): 10000-5.

[12] Bishop AC, Buzko O, Shokat KM. Magic bullets for protein kinases. Trends Cell Biol 2001; 11(4): 167-72.

[13] Cohen MS, Zhang C, Shokat KM, Taunton J. Structural Bioinformatics-Based Design of Selective, Irreversible Kinase Inhibitors. Science 2005; 308: 1318-1321.

[14] Kumar S, Zhou B, Liang F, Wang WQ, Huang Z, Zhang ZY. Activity-based probes for protein tyrosine phosphatases.Proc Natl Acad Sci USA 2004; 101(21): 7943-8.

[15] Powers JC, Asgian JL, Ekici OD, James KE. Irreversible inhibitors of serine, cysteine and threonine proteases. Chem Rev 2002; 102: 4639-4750.

[16] Turk V, Turk B, Turk D. Lysosomal cysteine proteases: facts and opportunities. EMBO J 2001; 20(17): 4629-33.

[17] Yasuda Y, Kaleta J, Bromme D. The role of cathepsins in osteoporosis and arthritis: rationale for the design of new therapeutics. Adv Drug Deliv Rev 2005; 57(7): 973-93.

[18] Fehrenbacher N, Jaattela M. Lysosomes as targets for cancer therapy. Cancer Res 2005; 65(8): 2993-5.

[19] Barrett AJ, Kembhavi AA, Brown MA, Kirschke H, Knight CG, Tamai M, *et al*. L-trans-Epoxysuccinyl-leucylamido(4-guanidino) butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. Biochem J 1982; 201(1): 189-98.

[20] Greenbaum D, Medzihradsky KF, Burlingame A, Bogyo M. Epoxide electrophiles as activity-dependant cysteine protease profiling and discovery tools. Chem Biol 2000; 7(8): 569-81.

[21] Greenbaum D, Baruch A, Hayrapetian L, Darula Z, Burlingame A, Medzihradsky KF, Bogyo M. Chemical approaches for functionally probing the proteome. Mol Cell Proteomics 2002; 1(1): 60-8.

[22] Verhelst SH, Bogyo M. Solid-phase synthesis of double-headed epoxysuccinyl activity-based probes for selective targeting of papain family cysteine proteases. Chembiochem 2005; 6(5): 824-7.

[23] Joyce JA, Baruch A, Chehade K, Meyer-Morse N, Giraudo E, Tsai FY, *et al*. Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. Cancer Cell 2004; 5(5): 443-53.

[24] Baruch A, Greenbaum D, Levy ET, Nielsen PA, Gilula NB, Kumar NM, *et al*. Defining a link between gap junction communication, proteolysis, and cataract formation. J Biol Chem 2001; 276(31): 28999-9006.

[25] Yasothornsrikul S, Greenbaum D, Medzihradszky KF, Toneff T, Bundey R, Miller R, *et al*. Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. Proc Natl Acad Sci USA 2003; 100(16): 9590-5.

[26] Greenbaum DC, Baruch A, Grainger M, Bozdech Z, Medzihradszky KF, Engel J, *et al*. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. Science 2002; 298(5600): 2002-6.

[27] Goulet B, Baruch A, Moon NS, Poirier M, Sansregret LL, Erickson A, *et al*. A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. Mol Cell 2004; 14(2): 207-19.

[28] Mason RW, Wilcox D, Wikstrom P, Shaw EN. The identification of active forms of cysteine proteinases in Kirsten-virus-transformed

mouse fibroblasts by use of a specific radiolabelled inhibitor. Biochem J 1989; 257(1): 125-9.
[29] Bromme D, Smith RA, Coles PJ, Kirschke H, Storer AC, Krantz A. Potent inactivation of cathepsins S and L by peptidyl (acyloxy)methyl ketones. Biol Chem Hoppe Seyler 1994; 375(5): 343-7.
[30] Kato D, Boatright KM, Berger AB, Nazif T, Blum G, Ryan C, *et al*. Activity-based probes that target diverse cysteine protease families. Nat Chem Biol 2005; 1(1): 33-8.
[31] Blum G, Mullins SR, Keren K, Fonovič M, Jedeszko C, Rice MJ, *et al*. Dynamic imaging of protease activity with fluorescently quenched activity-based probes. Nat Chem Biol 2005; 1(4): 203-9.
[32] Thornberry NA, Bull HG, Calaycay JR, Chapman KT, Howard AD, Kostura MJ, *et al*. A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. Nature 1992; 356(6372): 768-74.
[33] Thornberry NA, Peterson EP, Zhao JJ, Howard AD, Griffin PR, Chapman KT. Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones. Biochemistry 1994; 33(13): 3934-40.
[34] Uhlmann F, Wernic D, Poupart MA, Koonin EV, Nasmyth K. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell 2000; 103(3): 375-86.
[35] Kalafatis M, Egan JO, van 't Veer C, Cawthern KM, Mann KG. The regulation of clotting factors. Crit Rev Eukaryot Gene Expr 1997; 7(3): 241-80.
[36] Vandennabeele P, Orrenious S, Zhivotovsky P. Serine proteases and calpains fulfill important supporting roles in the apoptotic tragedy of cellular opera. Cell Death Differ 2005; 12(9): 1219-24.
[37] Cravatt BF, Giang DK, Mayfield SP, Boger DL, Lerner RA, Gilula NB. Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. Nature 1996; 384(6604): 83-7.
[38] McKinney MK, Cravatt BF. Structure and function of fatty acid amide hydrolase. Annu Rev Biochem 2005; 74: 411-32.
[39] Jessani N, Humphrey M, McDonald WH, Niessen S, Masuda K, Gangadharan B, *et al*. Carcinoma and stromal enzyme activity profiles associated with breast tumor growth *in vivo*. Proc Natl Acad Sci USA 2004; 101(38): 13756-61.
[40] Steiner DF. The proprotein convertases. Curr Opin Chem Biol 1998; 2(1): 31-9.
[41] Patricelli MP, Giang DK, Stamp LM, Burbaum JJ. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. Proteomics 2001; 1(9): 1067-71.
[42] Jessani N, Liu Y, Humphrey M, Cravatt BF. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. Proc Natl Acad Sci USA 2002; 99(16): 10335-40.
[43] Jessani N, Niessen S, Wei BQ, Nicolau M, Humphrey M, Ji Y, *et al*. A streamlined platform for high-content functional proteomics of primary human specimens. Nat Methods 2005; 2(9): 691-697.
[44] Hawthorne S, Hamilton R, Walker BJ, Walker B. Utilization of biotinylated diphenyl phosphonates for disclosure of serine proteases. Anal Biochem 2004; 326(2): 273-5.
[45] Marnett AB, Nomura AM, Shimba N, Ortiz de Montellano PR, Craik CS. Communication between the active sites and dimer interface of a herpesvirus protease revealed by a transition-state inhibitor. Proc Natl Acad Sci USA 2004; 101(18): 6870-5.
[46] Coleman JE. Zinc enzymes. Curr Opin Chem Biol 1998; 2(2): 222-34.
[47] Rasnick D, Powers JC. Active site directed irreversible inhibition of thermolysin. Biochemistry 1978; 17(21): 4363-9.
[48] Brown S, Bernardo M, Li ZH, Kotra LP, Tanaka Y, Fridman R, *et al*. Potent and Selective Mechanism-Based Inhibition of Gelatinases. J Am Chem Soc 2000; 122(28): 6799-800.
[49] Ikejiri M, Bernardo MM, Meroueh SO, Brown S, Chang M, Fridman R, *et al*. Design, synthesis, and evaluation of a mechanism-based inhibitor for gelatinase A. J Org Chem 2005; 70(14): 5709-12.
[50] Adam GC, Cravatt BF, Sorensen EJ. Profiling the specific reactivity of the proteome with non-directed activity-based probes. Chem Biol 2001; 8(1): 81-95.
[51] Adam GC, Sorensen EJ, Cravatt BF. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. Nat Biotechnol 2002; (8): 805-9.
[52] Wymann MP, Bulgarelli-Leva G, Zvelebil MJ, Pirola L, Vanhaesebroeck B, Waterfield MD, *et al*. Wortmannin inactivates phosphoinositide 3-kinase by covalent modification of Lys-802, a residue involved in the phosphate transfer reaction. Mol Cell Biol 1996; 16(4): 1722-33.
[53] Runnegar M, Berndt N, Kong SM, Lee EY, Zhang L. *In vivo* and *in vitro* binding of microcystin to protein phosphatases 1 and 2A. Biochem Biophys Res Commun 1995; 216(1): 162-9.
[54] Liu S, Widom J, Kemp CW, Crews CM, Clardy J. Structure of human methionine aminopeptidase-2 complexed with fumagillin. Science 1998; 282(5392): 1324-7.
[55] Adam GC, Burbaum J, Kozarich JW, Patricelli MP, Cravatt BF. Mapping enzyme active sites in complex proteomes. J Am Chem Soc 2004; 126(5): 1363-8.
[56] Barglow KT, Cravatt BF. Discovering disease-associated enzymes by proteome reactivity profiling. Chem Biol 2004; 11(11): 1523-31.
[57] Turner AJ, Isaac RE, Coates D. The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. Bioessays 2001; 23(3): 261-9.
[58] Overall CM, Lopez-Otin C. Strategies for MMP inhibition in cancer: innovations for the post-trial era. Nat Rev Cancer 2002; 2(9): 657-72.
[59] Kurukulasuriya R, Link JT, Madar DJ, Pei Z, Richards SJ, Rohde JJ, *et al*. Potential drug targets and progress towards pharmacologic inhibition of hepatic glucose production. Curr Med Chem 2003; 10(2): 123-53.
[60] Ross SA, Gulve EA, Wang M. Chemistry and biochemistry of type 2 diabetes. Chem Rev 2004; 104(3): 1255-82.
[61] Sieber SA, Niessen S, Hoover HS, Cravatt BF. Proteomic profiling of metalloprotease activities with cocktails of active-site probes. Nat Chem Biol 2006; 2(5): 274-81.
[62] Nazif T, Bogyo M. Global analysis of proteasomal substrate specificity using positional-scanning libraries of covalent inhibitors. Proc Natl Acad Sci USA 2001; 98(6): 2967-72.
[63] Greenbaum DC, Arnold WD, Lu F, Hayrapetian L, Baruch A, Krumrine J, *et al*. Small molecule affinity fingerprinting. A tool for enzyme family subclassification, target identification, and inhibitor design. Chem Biol 2002; 9(10): 1085-94.
[64] Berger AB, Witte MD, Denault JB, Sadaghiani AM, Sexton KM, Salvesen GS, *et al*. Identification of early intermediates of caspase activation using selective inhibitors and activity based probes. Mol Cell 2006; 23(4): 509-21.
[65] Funeriu DP, Eppinger J, Denizot L, Miyake M, Miyake J. Enzyme family-specific and activity-based screening of chemical libraries using enzyme microarrays. Nat Biotechnol 2005; 23(5): 622-7.
[66] Leung D, Hardouin C, Boger DL, Cravatt BF. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. Nat Biotechnol 2003; 21(6): 687-91.
[67] Baruch A, Jeffery DA, Bogyo M. Enzyme activity--it's all about image. Trends Cell Biol 2004; 14(1): 29-35.
[68] Gurtu V, Kain SR, Zhang G. Fluorometric and colorimetric detection of caspase activity associated with apoptosis. Anal Biochem 1997; 251(1): 98-102.
[69] Bedner E, Smolewski P, Amstad P, Darzynkiewicz Z. Activation of caspases measured in situ by binding of fluorochrome-labeled inhibitors of caspases (FLICA): correlation with DNA fragmentation. Exp Cell Res 2000; 259(1): 308-13.
[70] Amstad PA, Yu G, Johnson GL, Lee BW, Dhawan S, Phelps DJ. Detection of caspase activation in situ by fluorochrome-labeled caspase inhibitors. Biotechniques 2001; 31(3): 608-10, 612, 614.
[71] Evans MJ, Saghatelian A, Sorensen EJ, Cravatt BF. Target discovery in small-molecule cell-based screens by *in situ* proteome reactivity profiling. Nat Biotechnol 2005; 23(10): 1303-7.
[72] Okerberg ES, Wu J, Zhang B, Samii B, Blackford K, Winn DT, *et al*. High-resolution functional proteomics by active-site peptide profiling. Proc Natl Acad Sci USA 2005; 102(14): 4996-5001.
[73] Pusch W, Kostrzewa M. Application of MALDI-TOF mass spectrometry in screening and diagnostic research. Curr Pharm Des 2005; 11(20): 2577-91.
[74] Riedel W. J, Mehta M. A, Unema P. J. Human cognition assessment in drug research. Curr Pharm Des 2006; 12(20): 2525-39.
[75] Takahashi H, Sano H, Chiba H, Kuroki Y. Pulmonary surfactant proteins A and D: innate immune functions and biomarkers for lung diseases. Curr Pharm Des 2006; 12(5): 589-598.