Am J Pharmacogenomics 2004; 4 (6): 371-381
1175-2203/04/0006-0371/$31.00/0

Activity-Based Protein Profiling

Applications to Biomarker Discovery, *In Vivo* Imaging and Drug Discovery

Alicia B. Berger, *Phillip M. Vitorino* and *Matthew Bogyo*

Department of Pathology, Stanford University School of Medicine, Stanford, California, USA

Contents

Abstract . . . 371
1. Proteomics – What Are the Tools? . . . 372
2. Activity-Based Probes and Their Applications . . . 372
3. What is an Activity-Based Probe? . . . 372
   3.1 The Reactive Group . . . 373
   3.2 The Linker Region . . . 374
   3.3 The Tag Region . . . 374
4. What Are the Targets of Activity-Based Probes? . . . 374
5. Applications of Activity-Based Probes . . . 376
   5.1 Identification of Biomarkers for Human Disease . . . 376
   5.2 *In Vivo* Imaging of Enzyme Activities . . . 378
   5.3 Small Molecule Screening and Target Discovery . . . 378
6. Future Directions . . . 379

Abstract

The genomic revolution has created a wealth of information regarding the fundamental genetic code that defines the inner workings of a cell. However, it has become clear that analyzing genome sequences alone will not lead to new therapies to fight human disease. Rather, an understanding of protein function within the context of complex cellular networks will be required to facilitate the discovery of novel drug targets and, subsequently, new therapies directed against them.

The past ten years has seen a dramatic increase in technologies that allow large-scale, systems-based methods for analysis of global biological processes and disease states. In the field of proteomics, several well-established methods persist as a means to resolve and analyze complex mixtures of proteins derived from cells and tissues. However, the resolving power of these methods is often challenged by the diverse and dynamic nature of the proteome. The field of activity-based proteomics, or chemical proteomics, has been established in an attempt to focus proteomic efforts on subsets of physiologically important protein targets. This new approach to proteomics is centered around the use of small molecules termed activity-based probes (ABPs) as a means to tag, enrich, and isolate, distinct sets of proteins based on their enzymatic activity.

Chemical probes can be 'tuned' to react with defined enzymatic targets through the use of chemically reactive warhead groups, fused to selective binding elements that control their overall specificity. As a result, ABPs function as highly specific, mechanism-based reagents that provide a direct readout of enzymatic activity within complex proteomes. Modification of protein targets by an ABP facilitates their purification and isolation, thereby eliminating many of the confounding issues of dynamic range in protein abundance. In this review, we outline recent advances in the field of chemical proteomics. Specifically, we highlight how this technology can be

applied to advance the fields of biomarker discovery, *in vivo* imaging, and small molecule screening and drug target discovery.

The completion of the sequencing of the human genome[1] generated much excitement because of the possibility that this information would lead to the discovery of novel drug targets and, subsequently, new therapies to fight human disease. While access to complete genome sequences may directly result in the identification of disease markers in the form of polymorphisms in critical gene sequences, an understanding of the importance of a given gene in disease pathology requires a complete understanding of the function and regulation of its gene product. This simple fact significantly complicates our search for important new drug targets, as the number of protein species comprising any given proteome is estimated to be ten to a hundred times the number of genes in the genome. Furthermore, this number increases exponentially once post-translational modifications are considered. This overwhelming complexity is the 'Achilles heel' of the field of proteomics, which strives to isolate, identify, and quantify all proteins in complex mixtures or 'proteomes'. While impressive advances have been made in basic analytical techniques, the presently available proteomic methods have many limitations; new technologies are currently being developed to address these limitations.

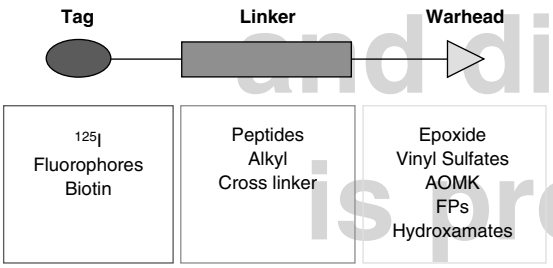

Fig. 1. General structure of an activity-based probe (ABP). ABPs are made up of three primary components that can be classified as the tag, linker, and warhead (reactive) elements. The key to developing an ABP relies on the selection of an optimal warhead group that can form a covalent bond through a mechanism-based reaction with an enzyme's catalytic machinery. The linker region serves as both a spacer between the tag and reactive warhead but is also used to drive selectivity of probe binding. The tag can take the form of a number of different reporters that can be analyzed using a range of detection methods. Some examples of each of these elements are listed (for a more complete description of these elements see the review by Jeffery and Bogyo[10]). **AOMK** = acyloxymethyl ketones (peptidal); **FP** = fluorophosphonate.

1. Proteomics – What Are the Tools?

Perhaps the most commonly used proteomic method is two-dimensional gel electrophoresis coupled to mass spectrometry (2DE/MS).[2,3] This method relies on two-dimensional isoelectric focusing/SDS- polyacrylamide gel electrophoresis (PAGE) gels to separate complex protein mixtures. The greatest limitation for 2DE/MS methods is their lack of resolving power that prevents the resolution and identification of low-abundance and membrane proteins.[4,5] Additionally, 2DE/MS methods are difficult and expensive to adapt to large-scale, high-throughput analyses. Two recent methods, isotope-coded affinity tagging (ICAT) and surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF), have been developed to address the issues of protein abundance, dynamic range, and resolution.[6-9] These methods have received much attention in the field of proteomics due to their ability to reduce complex proteomes to smaller, more tractable mixtures that can be analyzed in great detail. However, both of these methods make use of purification methods that are not dependent on protein function or enzyme activity, but rather on overall protein abundance. Consequently, they provide large data sets that can only be used to infer relevance of a protein to a given process based on changes in its abundance.

2. Activity-Based Probes and Their Applications

The field of chemical- or activity-based proteomics has recently been established as a means to enhance our understanding of global protein function. This technique uses small-molecule activity-based probes (ABPs) to profile enzymatic activities of a range of related enzymes within a complex cellular environment. ABPs are small molecules that covalently modify active enzymes through reaction of a suitably designed 'warhead' group with key catalytic residues of a target enzyme (figure 1). They can be used to detect, isolate, and identify active enzymes of interest in a range of samples including cells, tissues, and whole animals.[10-13] A range of ABPs have been developed that target a number of key enzyme families. A semi-comprehensive list is shown in table I.

3. What is an Activity-Based Probe?

Structurally, ABPs contain three elements (figure 1): (i) a reactive group also referred to as the 'warhead'; (ii) a linker region

Table I. The most commonly used activity-based probes grouped by target class. References with details of each probe or commercial sources from which probes are available are provided

	Biotin tags	Fluorescent tags	Radiolabels	Hemmagglutanin (HA) tags
Cysteine proteases				
Cathepsins/calpains	DCG-04[14]	BODIPY-DCG-04[15]	Cbz-Try-Ala-N2[16]	
			JPM-565, [17] JPM-OEt,[18] LHVS-PhOH[18]	
Cathepsin B	NS-196[19]		MB-074[18]	
Caspases	Biotin-X-VAD(OMe)-fmk[a]	SR-VAD-fmk[b]		
	Z-VK-X-(biotin)-D(OMe)-fmk[a]	FAM-VAD-fmk[b]		
De-ubiquitinating enzymes				HA-Ub-X, X = VS, Cl, Br2, Br3, VME, VSPH, VCN[20]
Arg-Gingipain	BiRK[21]			
Serine proteases				
All hydrolyases	FP-Biotin[22]	FP-PEG-TRM, -fluoroscein[23]	^{3}H-DFP[c]	
	FP-PEG-Biotin[24]	FP-fluoroscein[22]		
Threonine proteases				
The proteasome	Epoxomicin biotin[25]		NP-LLL-VS[a]	
	AdaLys(Bio)Ahx$_3$L$_3$-VS[a]			
Metalloproteases				
Matrix metalloproteases		HxBP-Rh[26]		
Phosphatases				
Protein tyrosine phosphatases	α-bromobenzylphosphonate (BPP) probes[27,28]			
Multiple families				
Aldehyde dehydrogenases, thiolase, NAD/NADP-dependent oxidoreductase, Enoyl CoA hydratase, epoxide hydrolyase glutathione *S*-transferase		Various sulfonate ester probes[12,29]		

a Available from Calbiochem (San Diego, CA, USA).

b Available from Immunochemistry Technologies, LLC. (Bloomington, MN, USA).

c Available from New England Nuclear (Torrance, CA, USA).

CoA = coenzyme A; **NAD** = nicotinamide adenine dinucleotide; **NADP** = NAD phosphate.

that can be used to control specificity of the probe; and (iii) a tag used for identification or purification.

3.1 The Reactive Group

The reactive group is perhaps the most significant and difficult piece of the probe to design. It functions to covalently link the ABP to an amino acid residue in an enzyme's active site. Furthermore, it must react only when the target enzyme is active. One of the biggest challenges in developing activity-based chemical probes is choosing a functional group that reacts preferably with the enzyme of interest while exhibiting minimal indiscriminate bio-reactivity. As a result, warheads are designed to have intermediate reactivity that can be controlled through the binding affinity of the probe for the target enzyme. Sulfonate esters,[29] fluorophosphonates,[22] peptide acyloxymethyl ketones

(AOMKs)[30] and epoxides[14] are examples of warheads that have been successfully incorporated into ABPs.

All probes differ in their lower limits of detection. AOMK's and epoxides have been used to label the caspase and papain families of the cysteine proteases while the fluorophosphonates and diphenylphosphonates have been used to target serine hydrolases.[22,31] Recently, ABPs have been designed to target metalloproteases by making use of the strong zinc-chelating hydroxamate moiety attached to a photo-activated cross-linker warhead.[26] Finally, there have been several examples of masked warheads that become activated by a target enzyme leading to their nonspecific modification of nearby residues in the enzyme active site. Such strategies have been employed to target purified proteases[32] and phosphatases,[27,28] but have yet to be used in complex proteomes.

3.2 The Linker Region

The linker region can be viewed as a bridge between the reactive group and the labeling tag. This probe element serves to prevent steric hindrance by the tag that could inhibit the reactivity of the probe. In its most basic form, a linker can take the form of an extended alkyl or polyethylene glycol (PEG) spacer. Additionally, the linker can serve as a specificity factor enabling targeting of the probe to a specific enzyme or class of enzymes. For example, to target proteases, this specificity region can be engineered to contain peptide sequences. Such sequences can be easily modified to drive specificity of the probe towards distinct sub-families of proteases.[33]

3.3 The Tag Region

The tag allows the identification or purification of modified enzymes. Biotin, fluorescent small molecules, and radioactive isotopes are most commonly incorporated into ABPs as tags, and all three can be used to visualize labeled proteins after SDS-PAGE. Biotin tags are used for affinity purification of modified enzymes and their subsequent identification through mass spectrometry. For direct visualization of labeled targets, fluorescent and radiolabeled tags are often used. Fluorescent or radiolabeled tags have an advantage over biotin as they have a higher dynamic range and require less time and handling to generate data.[10] Additionally, multiple fluorescent tags with non-overlapping excitation/emission spectra can be utilized to multiplex sample analysis using gel-based methods.[15,23]

4. What Are the Targets of Activity-Based Probes?

Not surprisingly, activity-based protein profiling efforts have focused largely on serine and cysteine proteases and hydrolases. These two enzyme classes have been the focus of extensive medicinal chemistry efforts and have lead to an abundance of potential mechanism-based reactive functional groups. Furthermore, these enzymes use a covalent attack mechanism resulting in the formation of a transient acyl-enzyme intermediate between the proteases and its protein substrate. Thus, simple tagging of a covalent suicide substrate enables the conversion of an inhibitor into an ABP.

The cysteine protease family is diverse and members of this family play an important role in the basic biological process, such as cell division,[34] cell death,[35] and antigen presentation.[36] Several classes of ABPs have been designed to target a range of cysteine protease targets.[14,16,18,37,38] A general peptide epoxide probe based on the natural product E-64 has found widespread use in studies of papain family proteases.[14] This probe has been use to define the functional roles of papain family proteases in processes such as tumor progression,[39] angiogenesis,[39] cataract formation,[40] pro-hormone processing,[41] malarial infections,[42] bacterial growth,[43] and plant response to pathogens.[44] Furthermore, probes based on the acyloxymethyl and fluoromethyl ketone have been used to target other non-papain family proteases. These targets include the caspases[37,45] (involved in cell death) and separase[34] (a key player in the process of cell division). Extensive use of these ABPs has lead to a greater understanding of the function of these proteases in basic biological processes.

The serine hydrolase family is a large and functionally diverse family comprising approximately 1% of the human genome and includes numerous proteases, lipases, esterases, and amidases.[1,46] Like the cysteine proteases, this family has been the focus of numerous studies using ABPs. In particular, a general fluorophosphonate (FP)-based probe has been designed that shows broad reactivity with this family of enzymes.[22] FP-biotin and related analogs have been applied to study the function of a number of serine hydrolases including fatty acid amide hydrolase (FAAH). Furthermore, due to their broad class-specific reactivity, these probes have been used to profile serine hydrolases in a range of cancer cell lines (see section 5).[47]

Metalloproteases, like serine and cysteine protease, play a key role in peptide hormone processing, tissue remodeling, and cancer.[48-51] However, this protease family uses a tightly bound water molecule to initiate attack of substrate, thereby circumventing the acyl enzyme intermediate. As a result, the design of ABPs for this

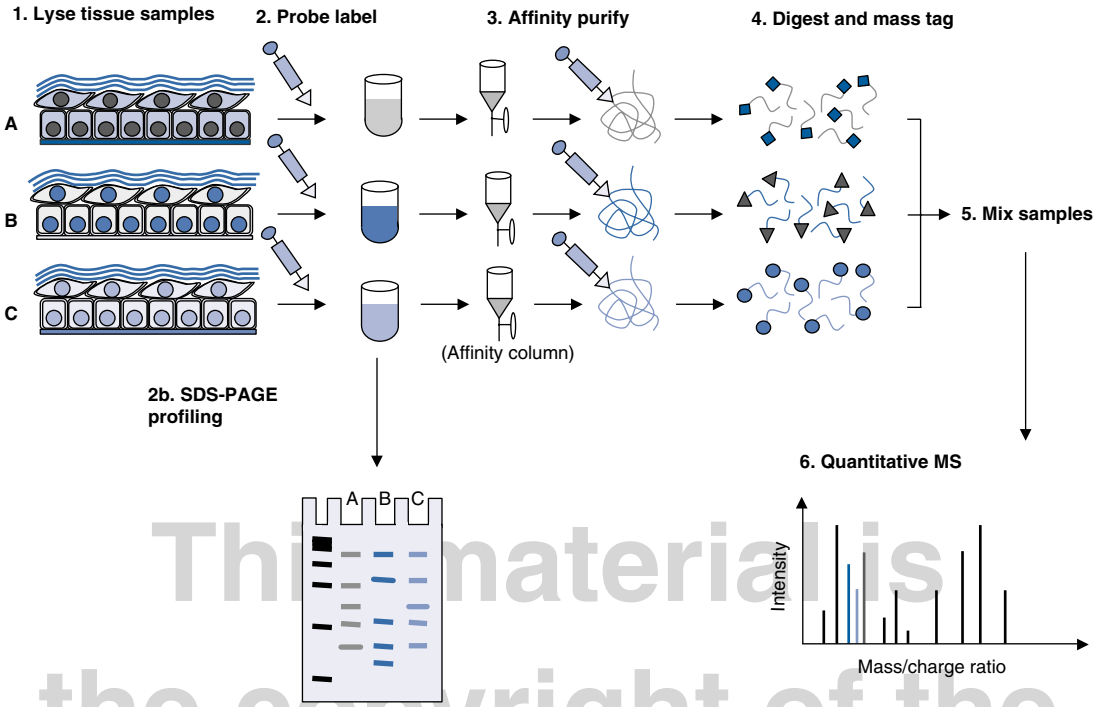

Fig. 2. Application of activity-based proteomics to biomarker discovery. Schematic of a workflow used to generate activity-based proteomic profiles. In this example three tissue samples (A, B, and C) are used to generate three distinct proteomes. Labeling with an activity-based probe (ABP) followed by direct analysis by SDS-PAGE provides general activity profiles. In this example the ABP reveals profiles of active enzymes in each sample that are unique, with some enzymes showing identical activity in all three samples while others show increased or decreased activity relative to the other samples. If the labeled probes contain an affinity tag, the labeled targets can be isolated by affinity chromatography. The resulting purified enzymes can then be digested to produce peptide fragments, followed by labeling of the free amino termini with an isotopically labeled tag. Mixing of the samples followed by high-resolution liquid chromatography/mass spectrometry (MS)/MS allows both the identity and relative abundance of the targets to be determined. When coupled together, these methods allow rapid profiling of multiple proteomes by SDS-polyacrylamide gel electrophoresis (PAGE) to determine samples that contain biomarkers of interest.

class of proteases is substantially more challenging. Recently, the first demonstration of an ABP targeting metalloproteases was reported.[26] This new class of ABPs contains a zinc-chelating hydroxamate coupled to a peptide backbone containing a photo-cross-linking group. These probes can be used to selectively label metalloproteases after irradiation with UV light. Since the hydroxamic acid group binds with high affinity, low probe concentrations can be used allowing the photo-reactive group to modify exclusively proteases targets. While this is a big leap forward for metalloprotease-specific probes, it still remains difficult to use these probes in the context of live cells or whole animals.

The human genome is thought to contain at least 500 protein kinases, all with a distinct function.[1,46] These enzymes play a critical role in almost all aspects of cell function and are attractive targets for functional proteomic strategies. The difficulty in the design of ABPs to target kinases lies in the lack of a sufficiently reactive nucleophile in the active site. Furthermore, most small molecule inhibitors of kinases are designed to bind in the conserved ATP-binding pocket of the enzymes.[52] This pocket is highly conserved among all ATP-binding enzymes, thereby making it difficult to tune the selectivity of a probe. As a result, the only reported ABP developed to target kinases is 5′-fluorosulfonylbenzoyl adenosine[53] and the applications of this reagent as an ABP have been limited. However, several covalent inhibitors that target the epidermal growth factor receptor (EGFR)[52] and phosphatidylinositol-3 (PI3) kinase families[54] have been developed. It is likely that these structures will be, or already have been, used to generate effective kinase ABPs.

The opposing effect of protein phosphorylation by kinases is controlled by protein phosphatases. The study of the phosphatase family of enzymes would also benefit from development of ABPs. Phosphatases make ideal targets for ABP design as their catalytic site includes a cysteine residue that acts as a nucleophilic thiolate for attack of a substrate.[55] Previously, suicide substrates that contain a masked electrophile have been used as probes for phosphatases.[27,28] These reagents, while potentially exciting, have yet to be confirmed as effective in the labeling of endogenous phosphatases in biologically relevant samples. Additionally, the design

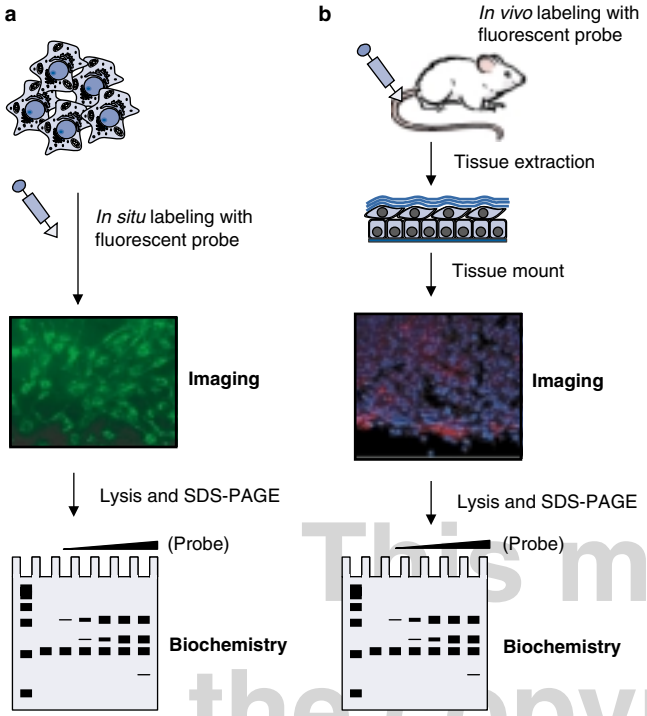

Fig. 3. Methods for *in situ* and *in vivo* imaging of enzyme activity using an activity-based probe (ABP). ABPs that freely penetrate cell membranes can be used to profile enzyme activities in live cells as well as in whole animals. (**a**) Intact cells are labeled by addition of a fluorescent ABP to the culture media. Following incubation, the probe is removed and the cell washed to reduce signal from unbound probe. Live cells can be imaged using standard fluorescence microscopy to visualize the subcellular localization of active enzymes (in this example using a green fluorophore). The same cells can be collected, lysed, and analyzed by reducing SDS-polyacrylamide gel electrophoresis (PAGE). Probe-labeled enzymes are visualized by scanning of the gel using a flatbed laser scanner. (**b**) A fluorescent ABP is injected into the tail vein of a mouse. After a short incubation period, tissue is collected from desired organs, mounted, and analyzed by fluorescence microscopy. Active proteases within the given tissue can be visualized by the fluorescent signal (in this example using a red fluorescence). Similarly to *in situ* experiments, tissues can be homogenized and analyzed biochemically using reducing SDS-PAGE. The labeled enzyme can then be visualized by laser scanning of the fluorescent gel. In both cases, ABPs provide a facile method for microscopic imaging of active enzymes followed by direct biochemical analysis of probe labeled targets.

of cell-permeable probes for phosphatases is likely to be difficult due to the need for a highly charged phosphate mimetic.

In addition to all the directed proteomic profiling strategies outlined above, nondirected probes have been developed to target multiple enzyme classes concurrently.[12] Typically, nondirected probes are identified by screening libraries of small-molecule scaffolds linked to a reactive warhead. These initial screening efforts identify probes with optimal labeling properties in complex proteomes. Inhibitor libraries containing mixed sulfonate ester reactive groups have been used to screen both insoluble and soluble fractions of several tissues as well as several invasive breast cancer lines.[12] In addition to labeling class I aldehyde dehydrogenase (ALDH-1), these probe libraries labeled six mechanistically different enzyme classes that had not been previously identified or characterized in activity-based studies. Such broad-spectrum approaches are likely to increase the repertoire of available profiling tools and broaden the applications for activity-based proteomics.

5. Applications of Activity-Based Probes

The development of ABPs that target diverse enzyme families has resulted in a wide range of applications for activity-based proteomics. Highly class-specific probes can be applied to complex proteomes from large collections of biological samples in order to profile global changes in enzyme activity. Analysis of these profiles from healthy and disease samples provides a means to identify novel biomarkers (figure 2). Furthermore, *in situ* and *in vivo* application of ABPs allows direct imaging of protease activity within the context of whole cells and even whole animals (figure 3). Finally, direct tagging of enzyme targets provides a means to both isolate novel targets and screen small molecules and drug leads for selectivity and potency in biologically relevant samples (figure 4). Each of these applications will be outlined below.

5.1 Identification of Biomarkers for Human Disease

The completion of the human genome sequence and large-scale sequencing projects aimed at the analysis of polymorphisms within our genome has led to an intensification of interest in the discovery of biomarkers for human disease. Clearly, analysis of single nucleotide polymorphisms (SNPs) will provide a rapid means of identifying many new disease markers. However, in some cases alterations in cellular physiology leading to disease will be reflected only at the level of changes in regulation of enzyme function. In these cases, tools such as ABPs will be required to identify the relevant players.

Not surprisingly, some of the first attempts at global profiling of enzyme activity in disease-associated tissues and cells have focused on cancer. The general serine hydrolase reactive probe FP-Biotin has been used to profile the activity of serine hydrolases in part of the set of NCI-60 melanoma and breast cancer cell lines.[47] This fluorescently labeled ABP was used to rapidly generate profiles of serine hydrolases by simple SDS-PAGE analysis followed by isolation and identification of labeled enzymes. The resulting labeling profiles allowed cancer cell lines to be

classified into melanoma, breast, or invasive cancer cell clusters based on their levels of active serine hydrolases. This study demonstrates the potential power of activity profiles and shows that a limited dataset based on activity levels can provide enough information to allow correct classification of cell lines. Interestingly, the estrogen receptor-negative (ER−) breast cancer line MDA-MB-435 was the only cell line that failed to cluster based on its classification, showing greater similarity to melanoma lines.[47] This finding was in agreement with previous gene expression-based profiling experiments, further suggesting that the cell line has been misclassified.[56] Intriguingly, in the majority of melanoma cell lines, several serine hydrolases showed increased activity measured by ABP labeling, but the same hydrolases showed decreased activity in the highly invasive melanoma cell line MUM-2B. In addition, the similarity in urokinase activity (an established serine protease biomarker for cancer progression *in vivo*[57]) in the cell lines NCI/ADR, MDA-MD-231, and MUM-2B, despite drastic differences in transcript levels for the urokinase gene in these cell lines, was also unexpected.

This result highlights a key point that most enzymes are controlled at some level by post-translational events that would not be reflected in their expression profiles. Finally, activity-based profiling results identified several enzyme targets including the membrane-associated serine hydrolase, KIAA1363, that were strongly up-regulated in MUM-2B and the invasive breast carcinoma cell line MDA-MB-231.[47] The identification of these potentially promising enzymes targets suggest that activity-based profiling is an effective method for biomarker discovery.

In another example of activity-based profiling, a recently developed ABP for metalloproteases, HxBP-Rh, was used to profile metalloprotease activities in invasive melanoma cell lines.[26] Neprilysin, a membrane-associated metalloprotease previously shown to be highly up-regulated in melanoma cell lines,[58] was also identified using activity-based profiling. Importantly, the labeling of neprilysin by HxBP-Rh was potently inhibited by ilomostat (GM6001), a matrix metalloprotease (MMP) inhibitor currently in clinical development.[48,59-61] These data highlight the utility of ABPs for identification of drug 'off-targets' and suggest that ilomostat has potent inhibitory effects on metalloproteases outside the MMP family that could lead to unanticipated toxicity *in vivo*.

In addition to the use of direct probes to profile defined enzyme classes in complex proteomes, nondirected probes also show promise for the identification of biomarkers of human disease. Broad spectrum probes based on the reactive sulfonate ester warhead have been applied to profile a number of tissues and cells lines.[12,29] These probes modify a wide range of mechanistically distinct enzymes, thereby providing a larger dataset when used for profiling complex proteomes. Initial efforts using the sulfonate ester probes yielded several enzymes whose activity appeared to be altered in cancer cells. Of particular interest was the identification of omega glutathione S-transferase (GST0-1) as being up-regulated in its activity in ER− breast cancer cell lines. This protein was not previously associated with invasive breast cancers, suggesting that it may serve as a novel biomarker and further efforts to characterize its function should be considered. Taken together, the initial forays into activity-based profiling using both directed and nondirected probes may lead to the identification of useful biomarkers for diagnostic purposes and provide preliminary validation of potential new drug targets.

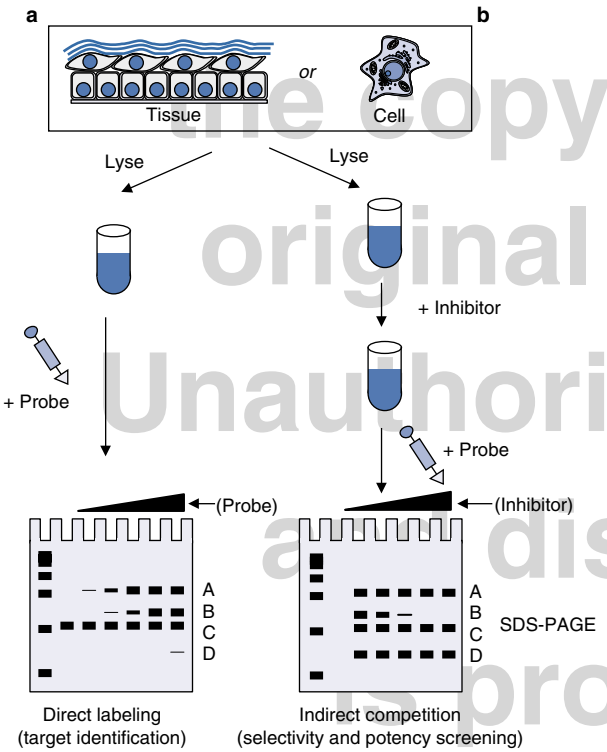

Fig. 4. Applications for activity-based probes (ABP) in drug discovery. (**a**) ABPs can be used to directly label total cellular lysates from cells or tissues from disease and normal samples. The resulting profiles can be used to select potentially interesting enzyme targets for further validation. (**b**) Samples can likewise be treated first with an inhibitor or drug lead and then labeled with an ABP. When the inhibitor binds to its enzyme target it blocks labeling of the active site by the ABP. This is visualized by probe labeling followed by SDS-polyacrylamide gel electrophoresis (PAGE) analysis. The resulting competition profile provides information regarding both selectivity towards a set of targets and also potency of the lead compound. While this example is based on *in vitro* labeling of cells and tissue extracts, a similar workflow can be used for *in vivo* drug selectivity profiling.

5.2 *In Vivo* Imaging of Enzyme Activities

Perhaps the greatest challenge for those attempting to define protein function is the need to assess how a given enzyme is regulated in the context of its native cellular environment. In addition to post-translational modifications and interactions with co-factors or inhibitors, most enzymes are regulated by their localization and temporal expression within a cell or tissue. Thus, cellular and whole-body imaging methods will be required to assess where and when an enzyme functions. Several approaches have been taken to develop cell-based and non-invasive imaging methods. The first steps towards this goal have focused on reporter substrates that produce a fluorescent signal when processed by an enzyme (for review see Baruch et al.[62]). The second method makes use of fluorescently labeled ABPs that directly tag the active enzymes *in vivo*.

Several classes of reporters have been developed to image enzyme activity. These imaging tools rely on fluorescent, near infrared (NIR) fluorescent, and bioluminescent tags to produce a signal. Several types of fluorescent reporter substrates can be used to image a range of enzyme activities *in vivo*.[62] All fluorescent substrate-based probes must be designed to produce a signal only after being processed by an enzyme target. Strategies to produce activity-dependent signals include peptide-bound fluorophores that are fluorescent only after release from the peptide,[63] and fluorphores in which the signal results from the loss of proximity to a quenching group as a result of enzymatic processing.[64-66] A third type of reporter strategy uses fluorescence resonance energy transfer (FRET); FRET can be disrupted by the processing of the substrate by the target enzyme. Such FRET-based reporters have been designed for several classes of enzymes.[62] While all of these methods have shown promise in the *in vivo* and *in situ* imaging of enzymes, they are limited by the selectivity of the substrate for the target enzyme. Furthermore, it remains difficult to link the signal from a reporter back to the enzyme responsible for producing it, thereby confounding interpretation of data.

ABPs have the potential to circumvent some of the problems associated with substrate-based imaging. Since ABPs directly modify an enzyme target through formation of a covalent bond, any signal observed in a whole cell or organism can be traced back to the enzyme by direct biochemical analysis (figure 3). ABPs that target cysteine proteases have been applied to visualize protease activity in intact cells.[15]

Taking this approach one step further, these probes have been used for *in vivo* imaging studies in a mouse model for pancreatic cancer.[39] In this study a fluorescent ABP was intravenously injected into a mouse and allowed to circulate for several hours. Subsequently, tissue was collected from the pancreas and protease activity visualized by fluorescence microscopy. After imaging the tissue, tissue lysis followed by SDS-PAGE analysis provided an activity profile that revealed the identity and relative activity levels of probe modified proteases that produced the fluorescent images. Furthermore, a similar *in vivo* labeling experiment was performed in animals that had been treated with an inhibitor of cysteine proteases.[39] The imaging and biochemical analysis provided a direct readout of the efficacy, potency, and biodistribution of the inhibitor. Clearly, this application of ABP-mediated *in vivo* imaging has the potential to dramatically improve the process through which drug leads are evaluated in animal model studies.

While an ABP-directed approach to *in vivo* imaging has its advantages, there are several limitations. First and foremost, it is unclear how much enzyme activity will be required for high-contrast non-invasive imaging. The difficulty lies in the fact that, unlike a substrate-based approach, there is no amplification of signal as a result of enzyme turnover. This means that a single enzyme molecule is modified by a single probe molecule and the signal is directly dependent on the amount of active enzyme present. In many cases this level may be below the limits of detection of optical imaging detectors. In addition, some ABPs are not ideal for *in vivo* imaging because they are large molecules that contain bulky tags unable to cross the cell membrane. To address this issue, ABP's that lack a bulky tag have been developed.[67-69] These reagents make use of a specific chemical 'handle' that can be covalently tagged using chemistry that takes place after the lysis of a tissue or cell. Thus, the surrogate tag is small and does not block the probes cellular access, yet it still provides a means to tag and isolate targets. Such probes have been used to generate activity profiles in whole cells as well as after *in vivo* labeling in whole animals. As yet, there have been no reported efforts to link ABPs to image contrasting reagents for use in molecular resonance imaging (MRI), most likely due to the necessity of high reagent concentrations for these imaging techniques.

5.3 Small Molecule Screening and Target Discovery

Currently, nearly 500 proteins have been validated as drug targets, yet only 122 have been targeted by orally bioavailable small molecule inhibitors.[70,71] About half of all known drug target families are enzymes that are amenable to chemical probe development; this number will grow as the number of targets continue to increase.[10] The protease family, specifically, makes up about 1–2% of the human genome. The cysteine family of proteases

likely contains additional drug targets yet to be identified.[11] Likewise, the serine hydrolase family comprises approximately 1% of the human genome and contains an extensive number of poorly validated targets that could be explored with chemical probes.[1,46] Additionally, activity-based proteomics technology has recently demonstrated success in targeting kinases, phosphatases, and metalloproteases.[26-28,72]

The discovery of new drugs is a lengthy process that starts with the validation of a protein target and requires optimization of the drug lead for properties including potency, selectivity, toxicity, adsorption, biodistribution, and metabolism. There are several points in this process where ABPs are likely to have a significant impact. At the earliest stage of the discovery process, a target must be identified and then validated before any costly chemistry efforts are expended. ABPs provide a means of identifying enzyme targets through direct labeling in crude proteomes (figure 4a). Using techniques described earlier in this review, potential targets can be selected based on the correlation of activities with a given disease pathology. ABPs can then be used to perform detailed biochemical and cell biological assays to validate their roles in disease progression. ABPs could also replace DNA sequencing as the preferred method for identifying constitutive activation. For example, several heterozygous activating mutations in EGFR have been identified in patients with non-small cell lung cancers (NSCLCs).[72] The mutations were identified through sequencing the *EGFR* gene in patients with a dramatic clinical response to the drug gefitinib. An effective ABP targeting EGFR would provide a direct readout of kinase activity and, therefore, could replace sequencing methods for identifying patients with constitutive activation of this enzyme.

On validation of a target, the first rounds of inhibitor screening can begin. Typically, screens are established using substrate-based assays that rely on purified enzymes (usually from a recombinant protein source). These high throughput screens require the ability to express or purify the enzyme target in a form that is capable of processing an optimal reporter substrate. These assays, while robust, often only poorly recapitulate the biological process that the enzyme catalyzes. Furthermore, *in vitro* assays provide only limited information regarding *in vivo* potency and selectivity of a given inhibitor. To address these problems, ABPs have been used to develop small molecule inhibitor screens.[33,73] These screens make use of a labeled probe to compete with small molecule inhibitors for binding in the enzyme active site (figure 4b). This type of assay can be performed on large sets of enzymes without the need to optimize the assay for each enzyme. Furthermore, and perhaps most importantly, this type of screen can be carried out in crude proteomes without the need to express and purify all related enzyme family members. As a result, the screen provides information on both the potency and selectivity of a small molecule inhibitor within the true cellular environment.

Using this approach, ABPs have been used to identify a number of novel protease inhibitors for use in biological studies of protease function.[15,42] Furthermore, competition screens of diverse libraries of inhibitors provide so-called 'affinity fingerprints' that can be used to catalog closely related enzymes based on their inhibitor binding properties.[33] Similar screens have been performed using a general serine hydrolase probe.[73] These studies show that competition-based assays can be used to identify both reversible and irreversible inhibitors from large and small libraries of compounds.

The late stages of the drug discovery process are often the most difficult and costly. This is the stage of the process when lead compounds are selected for initial animal studies to determine their general toxicity and pharmacodynamic properties. ABPs also afford the possibility to advance late-stage studies in animals as they can be used for *in vivo* competition and *in vivo* imaging studies that provide vital information regarding the specificity and potency of the drug lead in the context of a whole animal. Lead candidates that show unfavorable side effects can then be screened for cross-reactivity with related enzyme targets that may have been absent from initial *in vitro* screening efforts. Such information provides a method of rapid triage of compounds that are likely to perform poorly in costly, late-stage clinical trials. Thus, at multiple points, activity-based profiling can significantly streamline the drug discovery process.

6. Future Directions

With the potential applications for activity-based proteomics now firmly established, the focus must now shift towards further improving the technology. The next major advance will be to couple activity-based profiling methods to gel-free analytical techniques that make use of mass spectrometry for the direct identification of targets and quantification of their activities. Furthermore, new analytical methods will be required to enable high-throughput competition assays to be performed on medium to large libraries of compounds using small quantities of biological samples. Such methods will also enable extensive activity-based profiling of large sets of clinically relevant tissue samples for identification of biomarkers. Finally, there is a need to further expand the set of available probes so that we can begin to analyze a larger complement of enzyme families. Currently, the number of enzyme classes

that can be targeted by ABPs is limited to a handful of enzyme families, such as cysteine and serine proteases. Perhaps the development of more structurally diverse libraries will yield new probe scaffolds. The field of proteomics will most certainly continue to experience groundbreaking technological innovation that will lead to an even greater understanding of the complex nature of protein activities. ABPs are tools that will facilitate this understanding.

Acknowledgments

This work was funded by a National Technology Centers for Networks and Pathways grant, NIH grant U54 RR020843 (awarded to Dr Bogyo). Dr Berger was funded by NHGRI trainiing grant 5T32HG00044.

References

1. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. Science 2001 Feb 16; 291: 1304-51
2. Corthals GL, Wasinger VC, Hochstrasser DF, et al. The dynamic range of protein expression: a challenge for proteomic research. Electrophoresis 2000 Apr; 21: 1104-15
3. Patton WF. A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. Electrophoresis 2000 Apr; 21: 1123-44
4. Santoni V, Molloy M, Rabilloud T. Membrane proteins and proteomics: un amour impossible? Electrophoresis 2000 Apr; 21: 1054-70
5. Gygi SP, Corthals GL, Zhang Y, et al. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc Natl Acad Sci U S A 2000 Aug 15; 97: 9390-5
6. Gygi SP, Rist B, Gerber SA, et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999 Oct; 17: 994-9
7. Zhou H, Ranish JA, Watts JD, et al. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. Nat Biotechnol 2002 May; 20: 512-5
8. Goshe MB, Smith RD. Stable isotope-coded proteomic mass spectrometry. Curr Opin Biotechnol 2003 Feb; 14: 101-9
9. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 2004 Feb; 15: 24-30
10. Jeffery DA, Bogyo M. Chemical proteomics and its application to drug discovery. Curr Opin Biotechnol 2003 Feb; 14: 87-95
11. Speers AE, Cravatt BF. Chemical strategies for activity-based proteomics. Chembiochem 2004 Jan 3; 5: 41-7
12. Adam GC, Cravatt BF, Sorensen EJ. Profiling the specific reactivity of the proteome with non-directed activity-based probes. Chem Biol 2001 Jan; 8: 81-95
13. Adam GC, Sorensen EJ, Cravatt BF. Chemical strategies for functional proteomics. Mol Cell Proteomics 2002 Oct; 1: 781-90
14. Greenbaum D, Medzihradszky KF, Burlingame A, et al. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. Chem Biol 2000 Aug; 7: 569-81
15. Greenbaum D, Baruch A, Hayrapetian L, et al. Chemical approaches for functionally probing the proteome. Mol Cell Proteomics 2002 Jan; 1: 60-8
16. Mason RW, Bartholomew LT, Hardwick BS. The use of benzyloxycarbonyl[^{125}I]iodotyrosylalanyldiazomethane as a probe for active cysteine proteinases in human tissues. Biochem J 1989 Nov 1; 263: 945-9
17. Shi GP, Munger JS, Meara JP, et al. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. J Biol Chem 1992; 267 (11): 7258-62
18. Bogyo M, Verhelst S, Bellingard-Dubouchaud V, et al. Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. Chem Biol 2000 Jan; 7: 27-38
19. Schaschke N, Assfalg-Machleidt I, Lassleben T, et al. Epoxysuccinyl peptide-derived affinity labels for cathepsin B. FEBS Lett 2000; 482 (1-2): 91-6
20. Borodovsky A, Kessler BM, Casagrande R, et al. A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. EMBO J 2001; 20 (18): 5187-96
21. Mikolajczyk J, Boatright KM, Stennicke HR, et al. Sequential autolytic processing activates the zymogen of Arg-gingipain. J Biol Chem 2003; 278 (12): 10458-64
22. Liu Y, Patricelli MP, Cravatt BF. Activity-based protein profiling: the serine hydrolases. Proc Natl Acad Sci U S A 1999 Dec 21; 96: 14694-9
23. Patricelli MP, Giang DK, Stamp LM, et al. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. Proteomics 2001 Sep; 1: 1067-71
24. Kidd D, Liu Y, Cravatt BF. Profiling serine hydrolase activities in complex proteomes. Biochemistry 2001; 40 (13): 4005-15
25. Meng L, Kwok BH, Sin N, et al. Eponemycin exerts its antitumor effect through the inhibition of proteasome function. Cancer Res 1999; 59 (12): 2798-801
26. Saghatelian A, Jessani N, Joseph A, et al. Activity-based probes for the proteomic profiling of metalloproteases. Proc Natl Acad Sci U S A 2004 Jul 6; 101: 10000-5
27. Kumar S, Zhou B, Liang F, et al. Activity-based probes for protein tyrosine phosphatases. Proc Natl Acad Sci U S A 2004 May 25; 101: 7943-8
28. Zhu Q, Huang X, Chen GYJ, et al. Activity based fluorescent probes that target phosphatases. Tetrahedron Lett 2003; 44: 2669-72
29. Adam GC, Sorensen EJ, Cravatt BF. Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. Nat Biotechnol 2002 Aug; 20: 805-9
30. Thornberry NA, Peterson EP, Zhao JJ, et al. Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones. Biochemistry 1994 Apr 5; (33): 3934-40
31. Hawthorne S, Hamilton R, Walker BJ, et al. Utilization of biotinylated diphenyl phosphonates for disclosure of serine proteases. Anal Biochem 2004 Mar 15; 326: 273-5
32. Zhu Q, Girish A, Chattopadhaya S, et al. Developing novel activity-based fluorescent probes that target different classes of proteases. Chem Commun (Camb) 2004 Jul; 7 (13): 1512-3
33. Greenbaum DC, Arnold WD, Lu F, et al. Small molecule affinity fingerprinting: a tool for enzyme family subclassification, target identification, and inhibitor design. Chem Biol 2002 Oct; 9: 1085-94
34. Uhlmann F, Wernic D, Poupart MA, et al. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell 2000 Oct 27; 103: 375-86
35. Salvesen GS. Caspases and apoptosis. Essays Biochem 2002; 38: 9-19
36. Lecaille F, Kaleta J, Bromme D. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem Rev 2002 Dec; 102: 4459-88
37. Faleiro L, Kobayashi R, Fearnhead H, et al. Multiple species of CPP32 and Mch2 are the major active caspases present in apoptotic cells. EMBO J 1997 May 1; 16: 2271-81
38. Mason RW, Wilcox D, Wikstrom P, et al. The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabelled inhibitor. Biochem J 1989 Jan 1; 257: 125-9
39. Joyce JA, Baruch A, Chehade K, et al. Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. Cancer Cell 2004 May; 5: 443-53
40. Baruch A, Greenbaum D, Levy ET, et al. Defining a link between gap junction communication, proteolysis, and cataract formation. J Biol Chem 2001 Aug 3; 276: 28999-9006
41. Yasothornsrikul S, Greenbaum D, Medzihradszky KF, et al. CathepsinL in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. Proc Natl Acad Sci U S A 2003 Aug 5; 100: 9590-5
42. Greenbaum DC, Baruch A, Grainger M, et al. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. Science 2002 Dec 6; 298: 2002-6

43. Oleksy A, Golonka E, Banbula A, et al. Growth phase-dependent production of a cell wall-associated elastinolytic cysteine proteinase by Staphylococcus epidermidis. Biol Chem 2004 Jun; 385: 525-35

44. van der Hoorn RA, Leeuwenburgh MA, Bogyo M, et al. Activity profiling of papain-like cysteine proteases in plants. Plant Physiol 2004 Jul; 135: 1170-8

45. Boatright KM, Renatus M, Scott FL, et al. A unified model for apical caspase activation. Mol Cell 2003 Feb; 11: 529-41

46. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001 Feb 15; 409: 860-921

47. Jessani N, Liu Y, Humphrey M, et al. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. Proc Natl Acad Sci U S A 2002 Aug 6; 99: 10335-40

48. Overall CM, Lopez-Otin C. Strategies for MMP inhibition in cancer: innovations for the post-trial era. Nat Rev Cancer 2002 Sep; 2: 657-72

49. Puente XS, Sanchez LM, Overall CM, et al. Human and mouse proteases: a comparative genomic approach. Nat Rev Genet 2003 Jul; 4: 544-58

50. Chang C, Werb Z. The many faces of metalloproteases: cell growth, invasion, angiogenesis and metastasis. Trends Cell Biol 2001 Nov; 11: S37-43

51. Turner AJ, Isaac RE, Coates D. The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. Bioessays 2001 Mar; 23: 261-9

52. Garcia-Echeverria C, Traxler P, Evans DB. ATP site-directed competitive and irreversible inhibitors of protein kinases. Med Res Rev 2000 Jan; 20: 28-57

53. Colman RF. Affinity labeling of purine nucleotide sites in proteins. Annu Rev Biochem 1983; 52: 67-91

54. Wymann MP, Bulgarelli-Leva G, Zvelebil MJ, 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 Apr; 16: 1722-33

55. Mustelin T, Feng GS, Bottini N, et al. Protein tyrosine phosphatases. Front Biosci 2002 Jan 1; 7: d85-142

56. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 2000 Mar; 24: 227-35

57. Konecny G, Untch M, Pihan A, et al. Association of urokinase-type plasminogen activator and its inhibitor with disease progression and prognosis in ovarian cancer. Clin Cancer Res 2001 Jun; 7: 1743-9

58. Kanitakis J, Narvaez D, Claudy A. Differential expression of the CD10 antigen (neutral endopeptidase) in primary versus metastatic malignant melanomas of the skin. Melanoma Res 2002 Jun; 12: 241-4

59. Galardy RE, Grobelny D, Foellmer HG, et al. Inhibition of angiogenesis by the matrix metalloprotease inhibitor N-[R-2-(hydroxamidocarbonymethyl)-4-methylpentanoyl)]L-tryptophan methylamide. Cancer Res 1994 Sep 1; 54: 4715-8

60. Grobelny D, Poncz L, Galardy RE. Inhibition of human skin fibroblast collagenase, thermolysin, and *Pseudomonas aeruginosa* elastase by peptide hydroxamic acids. Biochemistry 1992 Aug 11; 31: 7152-4

61. Solorzano CC, Ksontini R, Pruitt JH, et al. A matrix metalloproteinase inhibitor prevents processing of tumor necrosis factor alpha (TNF alpha) and abrogates endotoxin-induced lethality. Shock 1997 Jun; 7: 427-31

62. Baruch A, Jeffery DA, Bogyo M. Enzyme activity: it's all about image. Trends Cell Biol 2004 Jan; 14: 29-35

63. Gurtu V, Kain SR, Zhang G. Fluorometric and colorimetric detection of caspase activity associated with apoptosis. Anal Biochem 1997 Aug 15; 251: 98-102

64. Packard BZ, Toptygin DD, Komoriya A, et al. Design of profluorescent protease substrates guided by exciton theory. Methods Enzymol 1997; 278: 15-23

65. Komoriya A, Packard BZ, Brown MJ, et al. Assessment of caspase activities in intact apoptotic thymocytes using cell-permeable fluorogenic caspase substrates. J Exp Med 2000 Jun 5; 191: 1819-28

66. Hirata H, Takahashi A, Kobayashi S, et al. Caspases are activated in a branched protease cascade and control distinct downstream processes in Fas-induced apoptosis. J Exp Med 1998 Feb 16; 187: 587-600

67. 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 Apr 23; 125: 4686-7

68. Speers AE, Cravatt BF. Profiling enzyme activities *in vivo* using click chemistry methods. Chem Biol 2004 Apr; 11: 535-46

69. Ovaa H, van Swieten PF, Kessler BM, et al. Chemistry in living cells: detection of active proteasomes by a two-step labeling strategy. Angew Chem Int Ed Engl 2003 Aug 11; 42: 3626-9

70. Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002 Sep; 1: 727-30

71. Drews J. Drug discovery: a historical perspective. Science 2000 Mar 17; 287: 1960-4

72. Sordella R, Bell DW, Haber DA, et al. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. Science 2004; 305: 1163-7

73. Leung D, Hardouin C, Boger DL, et al. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. Nat Biotechnol 2003 Jun; 21: 687-91

Correspondence and offprints: Dr *Matthew Bogyo*, Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Mail Code 5324, Stanford, CA 94305-5324, USA.
E-mail: Mbogyo@stanford.edu