Chemical Strategies for Functional Proteomics*

Gregory C. Adam‡, Erik J. Sorensen‡, and Benjamin F. Cravatt‡§¶

With complete genome sequences now available for several prokaryotic and eukaryotic organisms, biological researchers are charged with the task of assigning molecular and cellular functions to thousands of predicted gene products. To address this problem, the field of proteomics seeks to develop and apply methods for the global analysis of protein expression and protein function. Here we review a promising new class of proteomic strategies that utilizes synthetic chemistry to create tools and assays for the characterization of protein samples of high complexity. These approaches include the development of chemical affinity tags to measure the relative expression level and post-translational modification state of proteins in cell and tissue proteomes. Additionally, we discuss the emerging field of activity-based protein profiling, which aims to synthesize and apply small molecule probes that monitor dynamics in protein function in complex proteomes. ***Molecular & Cellular Proteomics 1:781–790, 2002.***

In response to the availability of complete genome sequences for numerous organisms, the field of proteomics has emerged with the goals of developing and applying methodologies that accelerate the functional analysis of proteins (1, 2). Strategies in proteomics can generally be divided into two categories that have complementary objectives: 1) the global characterization of protein expression, and 2) the global characterization of protein function. Large scale efforts to measure protein expression have typically relied on a combination of two-dimensional gel electrophoresis (2DE),[1] protein staining, and mass spectrometry (MS) for protein separation, detection, and identification, respectively (3). 2DE-MS methods are capable of simultaneously evaluating the relative abundance and modification state of numerous proteins from endogenous sources, thus permitting the identification of new proteins associated with discrete physiological and/or pathological states (*e.g.* nucleoside diphosphate kinase A as a marker for reduced metastatic potential in human prostate cancer cell lines (4)). However, several important protein classes, including low abundance and membrane-associated proteins, remain difficult to analyze by current 2DE approaches (3, 5, 6). Additionally, by focusing on measurements of protein abundance, 2DE-MS methods provide only an indirect assessment of protein function and may fail to detect important post-translational forms of protein regulation such as those mediated by protein-protein and/or protein-small molecule interactions (7).

To expedite the functional analysis of proteins, methods have also been introduced to examine protein activity on a global scale. These technologies include large scale yeast two-hybrid screens (8, 9), which aim to construct a comprehensive map of protein-protein interactions that occur in the cell, and protein microarrays (10, 11), which offer a platform to rapidly assess the function of recombinantly expressed proteins. Although capable of attributing specific molecular activities to individual protein products, these methods require that proteins be studied in artificial environments and therefore do not directly assess the functional state of these biomolecules in their native settings.

Recently a breed of chemical strategies has emerged that utilizes organic synthesis to create new tools and assays to advance the field of proteomics (12, 13). In this review, we describe chemical approaches for both abundance-based and activity-based proteomics with an emphasis on methods that permit the quantitative comparison of proteins, including low abundance and membrane-associated proteins, in samples of high complexity.

CHEMICAL APPROACHES FOR DETERMINING THE ABUNDANCE AND POST-TRANSLATIONAL MODIFICATION OF PROTEINS IN COMPLEX PROTEOMES

The most common method currently used by researchers in proteomics to monitor changes in protein abundance is 2DE-MS in which proteins are typically visualized and quantified by staining. Traditional staining methods, including Coomassie Blue and silver staining, are cost-effective but offer limited dynamic range and sensitivity (14). To improve these features, fluorescent dyes like SYPRO-ruby have recently been developed (15). Nonetheless, independent of the staining method employed, 2DE methods suffer from a lack of resolving power that hinders the detection of several important classes of proteins, including membrane-associated (5) and low abundance proteins (3).

From ‡The Skaggs Institute for Chemical Biology and the Departments of Chemistry and §Cell Biology, The Scripps Research Institute, La Jolla, California 92037

Received, September 16, 2002, and in revised form, October 3, 2002

Published, MCP Papers in Press, October 3, 2002, DOI 10.1074/mcp.R200006-MCP200

[1] The abbreviations used are: 2DE, two-dimensional electrophoresis; ABPP, activity-based protein profiling; ESI, electrospray ionization; FP, fluorophosphonate; ICAT, isotope-coded affinity tag; IMAC, immobilized metal affinity chromatography; LC, liquid chromatography; MS, mass spectrometry; PhIAT, phosphoprotein-specific isotope-coded affinity tag; PTP, protein tyrosine phosphatase; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl.

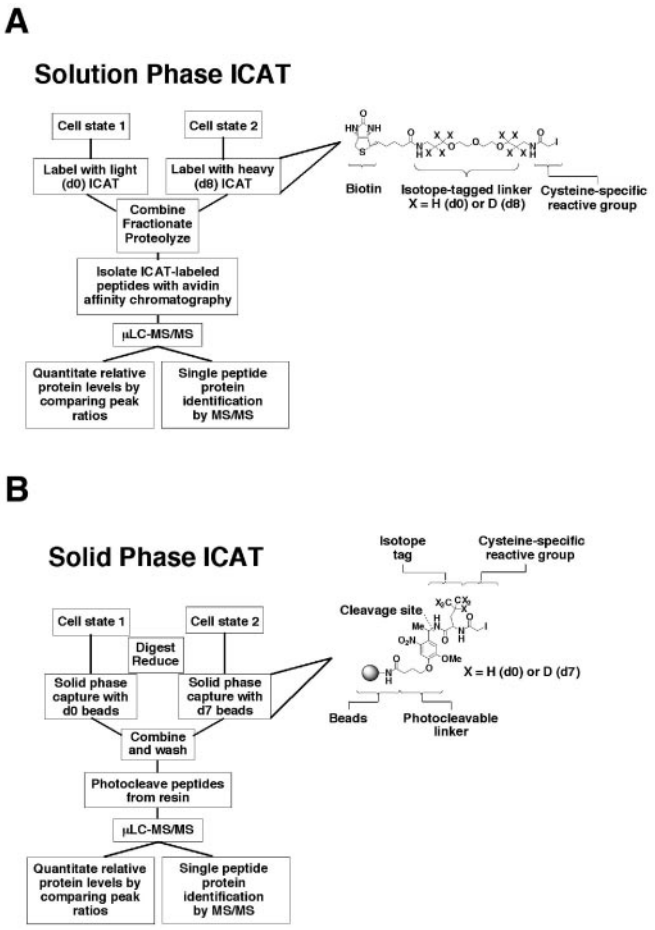

FIG. 1. **ICAT methods for quantitative proteomics.** Both solution phase (*A*; see Ref. 16) and solid phase (*B*; see Ref. 18) ICAT methods are outlined. See text for details.

The Isotope-coded Affinity Tag (ICAT) Method, a Chemical Approach to Quantify Protein Abundance in Complex Proteomes—As an alternative to 2DE-MS, a gel-free method for quantitative proteomics has been introduced that relies on chemical labeling reagents referred to as ICATs (16). These chemical probes consist of three general elements: a reactive group capable of labeling a defined amino acid side chain (*e.g.* iodoacetamide to modify cysteine residues), an isotopically coded linker, and a tag (*e.g.* biotin) for the affinity isolation of labeled proteins/peptides (Fig. 1*A*). For the quantitative comparison of two proteomes, one sample is labeled with the isotopically light (d_0) probe and the other with the isotopically heavy (d_8) version. To minimize error, both samples are then combined, digested with a protease (*i.e.* trypsin), and subjected to avidin affinity chromatography to isolate peptides labeled with isotope-coded tagging reagents. These peptides are then analyzed by liquid chromatography-mass spectrometry (LC-MS). The ratios of signal intensities of differentially mass-tagged peptide pairs are quantified to determine the relative levels of proteins in the two samples.

ICAT circumvents several of the previously described limitations of gel-based methods, providing improved access to important portions of the proteome like membrane-associated and low abundance proteins. For example, ICAT has been used to compare the microsomal fractions of normal and 12-phorbol 13-myristate acid-treated samples of the human HL-60 leukemia cell line (17). In this *in vitro* model of cellular differentiation, the ICAT method was capable of measuring the relative levels of 491 proteins, many of which were membrane-associated proteins and/or proteins of moderate to low abundance. Notably, this study identified previously unknown isoform-specific changes in protein kinase C that occurred during 12-phorbol 13-myristate acid-induced differentiation.

Recently the ICAT technology was converted to a format for the solid phase capture and release of chemically tagged peptides (18). In this study, the solid phase isotope-tagging reagent consisted of a thiol-specific reactive group attached via an isotopically modified amino acid (either d_0 or d_7 leucine) to an *o*-nitrobenzyl-based photocleavable linker bound to an aminopropyl-coated glass bead (Fig. 1*B*). Each of the two proteomes under comparison was digested with trypsin, and its cysteine-containing peptides were captured with either the light or heavy form of the solid phase reagent. The light and heavy beads were then combined, washed, and exposed to UV light to induce photocleavage of the linker. The isotopically labeled peptides, now present in solution, were then analyzed by LC-ESI-MS/MS. Compared with the original solution phase ICAT approach, the solid phase strategy required less sample handling and provided greater sensitivity for quantitative protein analysis. On the other hand, because solid phase ICAT involves proteolysis prior to probe labeling, the solution phase ICAT method may still be preferred in cases where the separation of labeled proteins is desired. For example, solution phase ICAT methods have been used in combination with 2DE to concurrently quantify changes in protein expression and modification state that occur in the yeast proteome in response to a metabolic shift (19).

Chemical Methods to Measure Protein Phosphorylation in Complex Proteomes—Building on the success of ICAT, related chemical proteomic strategies have been introduced to evaluate the post-translational modification state of proteins. In particular, several chemical reagents have been developed to measure the phosphorylation state of proteins in complex proteomes (20). Traditional methods for detecting protein phosphorylation include metabolic radiolabeling with inorganic [^{32}P]phosphate (21, 22) and affinity chromatography with either immobilized metal affinity chromatography (IMAC) (23) or phospho-specific antibodies (24). However, each of these techniques exhibits shortcomings for quantitative proteome analysis (Table I). For example, metabolic labeling with ^{32}P requires a viable cell source and therefore is not applicable for the proteomic analysis of human tissue specimens. Additionally, transitioning from the detection of ^{32}P-labeled proteins on 2DE gels to the molecular identification of these

TABLE I
A comparison of different proteomic strategies for measuring protein phosphorylation

Analytical method	Phosphorylated amino acids detected	Quantitative	Site of phosphorylation identified	General comments
Conventional methods				
^{32}P radiolabeling	Ser, Thr, Tyr	Semi	Difficult	Requires living samples
Phospho-specific antibodies	Ser, Thr, Tyr	No	Difficult	High nonspecific binding
IMAC	Ser, Thr, Tyr	Yes (d_0 or d_3 methanol)	Difficult	Methylation of carboxylic acids reduces background[a]
Chemical methods				
β-Elimination[b]	Ser, Thr	Yes (d_0 or d_4 ethanedithiol)	Yes (tag survives MS/MS)	Potential for β-elimination of *O*-linked carbohydrates
Phosphoramidate[c]	Ser, Thr, Tyr	Yes (d_0 or d_4 ethanolamine)	Difficult	Lengthy protocol

[a] See Ref. 25.
[b] See Refs. 26 and 27.
[c] See Ref. 28.

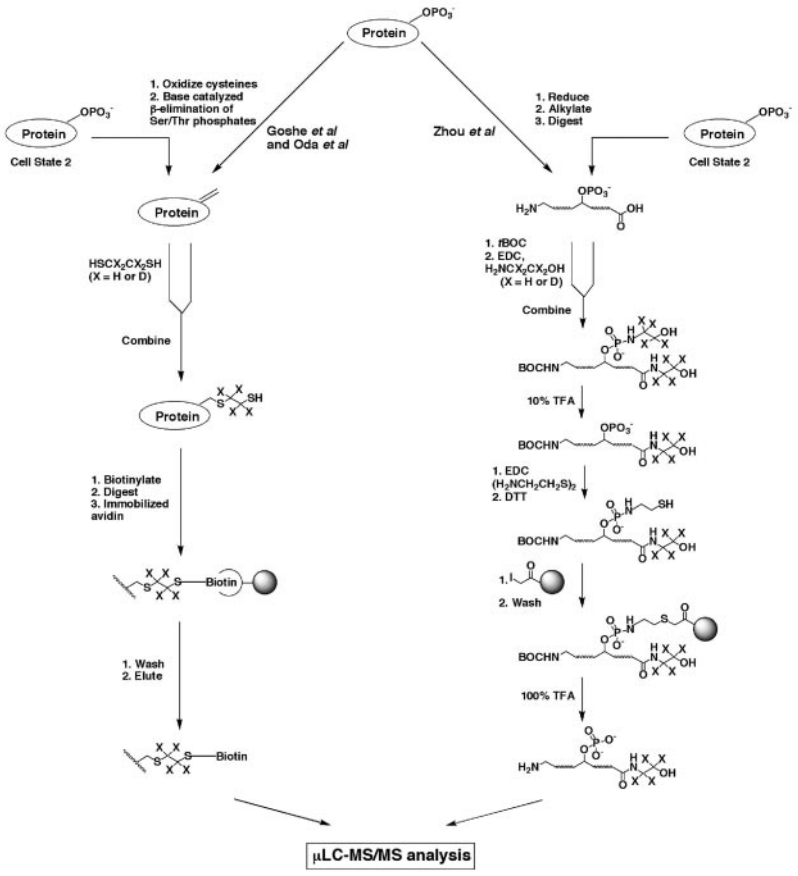

FIG. 2. **Chemical methods for measuring the phosphorylation state of proteins in complex proteomes.** Both the base-catalyzed phosphate elimination method (*left flow chart*; see Refs. 26 and 27) and the phosphoramidate modification method (*right flow chart*; see Ref. 28) are outlined. See text for details. *TFA*, trifluoroacetic acid; *tBOC*, *tert*-butoxycarbonyl; *DTT*, dithiothreitol; *EDC*, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride.

proteins can be challenging without the availability of target-specific enrichment reagents (*e.g.* antibodies). Affinity chromatography procedures like IMAC and phospho-specific antibodies have typically suffered from high levels of background binding by nonphosphorylated peptides and poor quantitation (notably, however, recent advances in IMAC methodologies may help to overcome these deficiencies (25)).

Two chemical tagging strategies for quantitative phosphoproteome analysis have recently been described. The first approach, concurrently put forth by two independent research groups (26, 27), involves the sequential base-catalyzed β-elimination of the phosphate group and nucleophilic addition of an affinity tag to the resulting dehydroalanine residue (Fig. 2). In both methods, cysteine residues on proteins were first oxidized with performic acid to prevent cross-reactivity in subsequent steps. Then treatment with base

transformed phosphoserine and phosphothreonine residues into Michael acceptors susceptible to nucleophilic attack with ethanedithiol. A reactive biotin reagent was then coupled to the free thiol end of ethanedithiol-modified sites, permitting the purification of the originally phosphorylated proteins by avidin affinity chromatography (enriched as either whole proteins or as peptides if preceded by digestion with trypsin). Affinity-isolated biotinylated peptides were then analyzed by ESI-LC-MS/MS, permitting the identification of the corresponding proteins as well as the specific sites of phosphorylation on these proteins. In the study by Goshe and colleagues (27), this method was adapted for quantitative analysis of phosphoproteomes by incorporating phosphoprotein-specific isotope-coded affinity tags (PhIATs). These PhIAT reagents consisted of two isotopic derivatives of ethanedithiol (a light (d_0) and heavy (d_4) version), each of which served as the nucleophile for one of the two proteomes under comparison (Fig. 2). The light and heavy PhIAT-modified proteomes were then combined, processed, and analyzed by ESI-LC-MS/MS as described previously for ICAT.

Because the strategy detailed above requires the β-elimination of a phosphate group to expose a site for affinity tagging, it is not capable of monitoring the phosphorylation state of tyrosine residues on proteins. In contrast, a second chemical method for phosphoproteome analysis developed by Zhou and colleagues (28) is applicable to phosphoseryl, -threonyl, and -tyrosyl residues. In this approach, proteins were first alkylated with iodoacetamide to block cysteine residues and then enzymatically digested with trypsin (Fig. 2). Following protection of the amino groups of the resulting peptide mixture with *tert*-butoxycarbonyl chemistry, the carboxyl/phosphoryl groups were modified with ethanolamine in a carbodiimide-catalyzed reaction. Treatment with acid promoted the hydrolysis of the less stable phosphoramidate bonds, thereby reforming the phosphates, which then underwent a reaction with a cystamine disulfide-bonded dimer. Reduction of the cystamine substituent resulted in the exposure of a free thiol group at each site of phosphorylation in the peptide sample. Thiol-modified phosphopeptides were then captured on the solid phase by reaction with iodoacetyl groups immobilized on glass beads. After stringent washing, phosphopeptides were released from the solid phase by phosphoramidate bond cleavage with trifluoroacetic acid and analyzed by ESI-LC-MS/MS.

A comparison of the phosphate elimination and phosphoramidate modification methods suggests that these approaches offer complementary advantages for phosphoproteome analysis. The phosphate elimination method requires fewer modification steps and results in chemically modified peptides suitable for tandem MS analysis to identify specific sites of phosphorylation (Table I). However, this strategy is only applicable to phosphoserine and phosphothreonine peptides. In contrast, the reversible phosphoramidate modification protocol can analyze any type of phosphorylated peptide but involves numerous derivatization steps and results in the recovery of unmodified phosphate groups, which typically disassociate during tandem MS analysis, confounding efforts to determine sites of phosphorylation. Importantly, however, both methods reduce sample complexity while at the same time enriching for phosphorylated proteins and therefore should provide access to low abundance constituents of the phosphoproteome. Additionally, because these chemical strategies offer a means to quantify changes in the phosphoproteome by isotope tagging, they should facilitate the discovery of molecular changes in signal transduction cascades associated with particular physiological and/or pathological processes.

CHEMICAL APPROACHES FOR DETERMINING THE ACTIVITY OF PROTEINS IN COMPLEX PROTEOMES

Conventional proteomics methods record variations in protein abundance and therefore provide only an indirect estimate of changes in protein activity. Accordingly, these approaches may fail to detect important post-translational forms of protein regulation such as those mediated by protein-protein and/or protein-small molecule interactions (7). To address these limitations, chemical strategies have been developed for activity-based protein profiling (ABPP) that utilize active site-directed probes to determine the functional state of enzymes in complex proteomes (12). Chemical probes for ABPP consist of at least two molecular elements: 1) a reactive group for binding to and covalently modifying the active sites of many members of a given enzyme class (or classes) and 2) a chemical tag for the rapid detection and isolation of reactive enzymes (Fig. 3*A*). Because these probes possess moderately reactive electrophilic groups, they are poised to selectively modify enzyme active sites, which are often enriched in nucleophilic amino acid residues important for catalysis.

To date, two general strategies for ABPP have been devised: 1) directed approaches that target specific classes of enzymes and 2) non-directed approaches that profile enzymes from several different classes. The chemical foundation for each of these methods as well as examples of their biological application are reviewed below.

Directed ABPP: the Design and Application of Activity-based Chemical Probes That Target Specific Classes of Enzymes

Directed ABPP approaches have capitalized on a rich history of mechanistic studies associated with particular classes of enzymes to create chemical probes with predictable proteome reactivities. By incorporating as the probe reactive group well known affinity labeling reagents, researchers have succeeded in creating ABPP probes that target, for example, serine hydrolases (29, 30) and subclasses of cysteine proteases (31, 32).

ABPP Probes That Target the Serine Hydrolase Superfamily—Serine hydrolases are one of the largest and most diverse

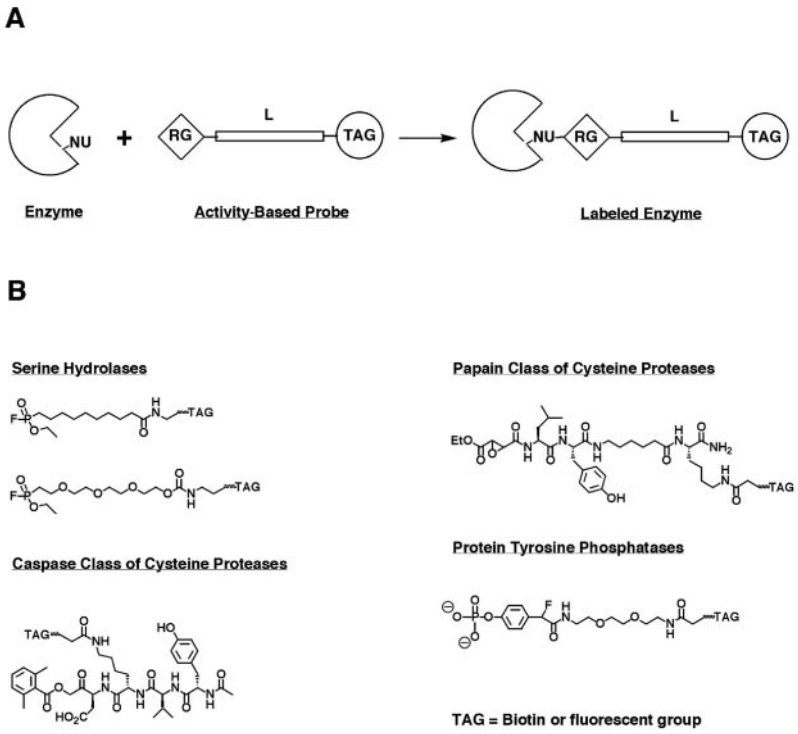

FIG. 3. **ABPP with chemical proteomics probes.** *A*, general mechanism by which activity-based probes label the active sites of targeted enzymes where *NU* = nucleophilic amino acid residue, *RG* = reactive group, *L* = linker, and *TAG* = detection/affinity tag. *B*, representative ABPP probes directed toward specific classes of enzymes. See text for details.

classes of enzymes in higher eukaryotes, representing ~1% of the predicted protein products encoded by the human genome (33, 34). Representative serine hydrolases include proteases like thrombin (35), trypsin (36), and urokinase plasminogen activator (37); lipid hydrolases like phospholipase A_2 (38); esterases like acetylcholinesterase (39); and amidases like fatty acid amide hydrolase (40). Due to shared features of their catalytic mechanism, nearly all members of the serine hydrolase superfamily are irreversibly inactivated by fluorophosphonate reagents (41). Accordingly, for the design of ABPP probes that target serine hydrolases, Liu and colleagues (29, 30) synthesized compounds consisting of: 1) a fluorophosphonate (FP) reactive group, 2) an alkyl or polyethylene glycol chain linker, and 3) a biotin tag (Fig. 3*B*). These FP probes were found to label numerous members of the serine hydrolase superfamily directly in complex proteomes. Additionally, FP probes were shown to read out the functional state of serine hydrolases, labeling, for example, active proteases but not their inactive zymogen and/or inhibitor-bound forms. Kidd and colleagues (30) exploited the activity-based nature of FP-proteome reactions to detect multiple brain serine hydrolases sensitive to trifluoromethyl ketone inhibitors of fatty acid amide hydrolase, demonstrating that ABPP can serve as a screen to evaluate the potency and selectivity of enzyme inhibitors. Notably, these inhibitor selectivity screens were conducted directly in complex proteomes, thus alleviating the need to recombinantly express or purify the enzymes under investigation. Finally, by establishing a covalent link between the labeled enzymes and a biotin tag, FP probes provided a straightforward route for the affinity purification and molecular identification of targeted proteins by avidin chromatography and mass spectrometry procedures, respectively (30).

Although valuable for the affinity purification of probe-reactive proteins, biotin-conjugated ABPP probes displayed several shortcomings for the systematic detection of enzyme activities in complex proteomes. In particular, biotin labeling events must be visualized indirectly, typically with avidin-horseradish peroxidase complexes and chemiluminescent substrates. These assays are limited in sensitivity, throughput, and dynamic range, thus hindering efforts to rapidly and quantitatively compare large numbers of proteomic samples. To address these limitations, the FP reactive group has been conjugated to fluorescent tags (either rhodamine (42) or fluorescein (29)), permitting the use of direct in-gel fluorescence scanning as a rapid, sensitive, and quantitative screen for activity-based protein labeling events (42). Notably, Patricelli and colleagues (42) have estimated that in-gel fluorescence scanning can detect on the order of 100 amol of FP-rhodamine-labeled enzyme, a detection limit nearly 2 orders of magnitude more sensitive than that of biotin-conjugated probes (29). Thus, a two-tiered strategy for ABPP has since been adopted in which, first, fluorescent probes are used for rapid and quantitative comparative proteome analysis, and second, biotin probes are applied to affinity enrich and identify differentially expressed enzyme activities.

Capitalizing on the technical advantages afforded by both fluorescence and biotin-avidin ABPP methods, Jessani and colleagues (43) set out to comparatively profile a panel of human cancer cell lines to determine whether a global anal-

ysis of serine hydrolase activities would yield proteomic information of sufficient quantity and quality to depict higher order cellular properties. In this study, cancer cell proteomes were split into three fractions prior to characterization: the secreted, membrane, and cytosolic fractions. Profiling of these proteomic fractions from 11 breast and melanoma cancer lines identified a cluster of serine hydrolase activities that distinguished these lines based on tissue of origin. Interestingly, however, nearly all of these enzymes were down-regulated in the most invasive cancer lines examined, which instead up-regulated a distinct set of serine hydrolase activities that included the protease urokinase and a novel membrane-associated enzyme, KIAA1363. A more detailed analysis revealed that most of the serine hydrolase activities responsible for classifying cancer cells into subtypes based on tissue of origin and/or state of invasiveness resided in the secreted and membrane proteome, suggesting that these proteomic fractions were particularly enriched in enzyme markers of cell behavior. Collectively these studies demonstrate that ABPP can generate molecular profiles that accurately depict higher order cellular properties and, in the process, identify uncharacterized enzyme activities, like KIAA1363, that may represent new biomarkers and/or therapeutic targets for the diagnosis and treatment of human disease.

ABPP Probes That Target Cysteine Proteases—For the design of activity-based chemical probes that label cysteine proteases, researchers have also exploited well characterized active site-directed covalent inhibitors. For example, Thornberry and colleagues (44) appended (acyloxy)methyl ketone inhibitors of caspases with biotin (Fig. 3*B*), creating first generation ABPP probes for this subclass of cysteine proteases (~15 predicted caspases are encoded by the human genome). Biotinylated variants of caspase inhibitors have since been applied to several model systems to identify members of this enzyme family associated with cellular events like apoptosis (31, 45). Recently Winssinger and colleagues (46) coupled caspase-directed reactive groups to peptide nucleic acids, permitting the detection of activated caspase-3 on a glass slide microarray bearing complementary oligonucleotide sequences. Although this approach may offer a more high throughput and miniaturized assay platform for the detection of probe-labeled proteins, it remains unclear whether such a strategy is applicable to the majority of ABPP probes, which each target multiple enzymes (*e.g.* FP probes where the fluorescent signal on a given microarray spot would likely represent a complicated sum of the levels of several enzyme activities).

To generate ABPP probes for the papain class of cysteine proteases, Greenbaum and colleagues (32) have synthesized biotinylated variants of the peptide epoxide natural product E-64, a covalent inhibitor of several papain family members (Fig. 3*B*). The papain class of proteases includes cathepsins, a family of lysosomal proteases (~15 predicted members encoded by the human genome), and calpains, a group of calcium-dependent cytosolic proteases (~10 predicted members encoded by the human genome). E-64-based ABPP probes have been used to identify cathepsin activities that correlate with skin cancer progression (32) and calpain activities implicated in cataract formation (47). Recently Greenbaum and colleagues (48) created variants of E-64-based ABPP probes coupled to BODIPY dyes and showed that these fluorescent reagents could be used to visualize cathepsin activities in living cells by fluorescence microscopy. Additionally, these probes were applied in combination with libraries of peptide epoxides to identify selective irreversible inhibitors of cathepsin B (48).

ABPP Probes That Target Tyrosine Phosphatases—Lo and colleagues (49) have reported the synthesis and application of first generation activity-based probes to profile members of the protein tyrosine phosphatase (PTP) family. These probes were comprised of a mechanism-based reactive group (a 4-fluoromethyl-1-phosphophenyl substituent (50)), a diethylene glycol linker, and a biotin or dansyl tag (Fig. 3*B*). The authors hypothesized that PTP-catalyzed hydrolysis of the phosphate group would promote a 1,6-elimination of the fluorine atom to form a highly reactive quinone methide that might label phosphatase active sites. Consistent with this notion, the tyrosine phosphatase PTP-1B, but not other proteins like phosphorylase *b* and albumin, was covalently modified by these mechanism-based probes. Still, high probe concentrations (1 mM) were required to label PTP-1B, and further studies will be needed to determine whether such conditions are compatible with profiling members of the PTP family in complex proteomes.

Non-directed ABPP: the Design and Application of Libraries of Activity-based Chemical Probes That Target Multiple Classes of Enzymes

As described above, the creation of activity-based probes for some enzyme classes, like serine and cysteine hydrolases, has been conceptually straightforward. Because active site-directed affinity labels were already known for these enzymes, researchers in chemical proteomics could couple these "reactive group" elements to an appropriate linker and detection/isolation tag to generate probes for ABPP. For many enzyme classes, however, cognate affinity labels do not yet exist, thus limiting the scope of such directed ABPP efforts. To expand the number of enzyme classes addressable with ABPP methods, Adam and colleagues (51, 52) have introduced a non-directed or combinatorial strategy in which libraries of candidate probes are synthesized and screened against complex proteomes for activity-dependent protein reactivity.

To demonstrate the feasibility of non-directed approaches for ABPP, a relatively small library of candidate probes was synthesized that incorporated the following elements: 1) a variable alkyl/aryl binding group, 2) a sulfonate ester reactive group, 3) an aliphatic linker, and 4) a rhodamine or biotin tag (for the detection and affinity isolation of protein targets, re-

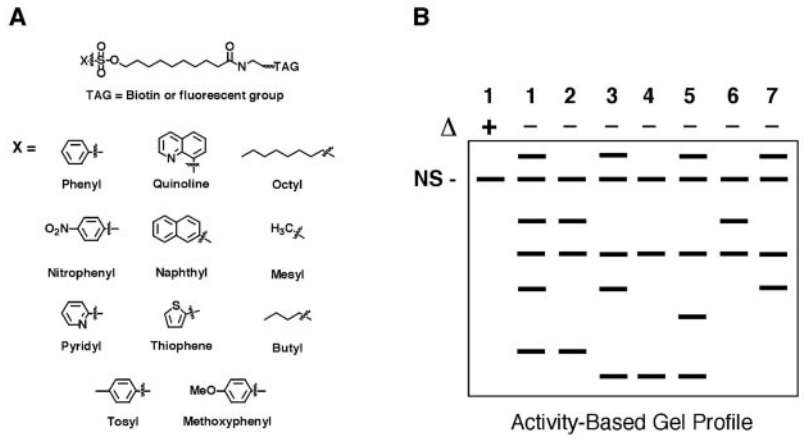

FIG. 4. *A*, representative library of chemical probes for non-directed ABPP where a variable alkyl or aryl binding group is attached to a sulfonate ester reactive group (see Refs. 51 and 52). *B*, schematic depiction of proteomic data generated with non-directed ABPP where specific and nonspecific (*NS*) targets are defined as those that show heat-sensitive and heat-insensitive probe labeling, respectively.

spectively) (Fig. 4*A*). By selecting a carbon electrophile (sulfonate ester) as the reactive group element of the probe library, it was hoped that the probes would label the active sites of enzymes from several different mechanistic classes. In support of this hypothesis, natural products bearing carbon electrophiles have been identified that covalently modify the active sites of a diverse number of enzymes including wortmannin, which targets kinases (53), microcystin, which targets phosphatases (54), and fumagillin, which targets metalloproteases (55). Rhodamine-tagged members of the sulfonate probe library were applied to tissue and cell line proteomes in a screen for specific protein reactivities, which were defined as those that occurred in native but not heat-denatured proteomes (conversely, proteins showing heat-insensitive reactivity were considered nonspecific targets) (Fig. 4*B*). The authors hypothesized that proteins reacting with sulfonate probes in a heat-sensitive manner would possess structured sites for small molecule interactions and that these sites would often determine the biological activity of the protein (*e.g.* the active site of enzyme). In these initial studies, several heat-sensitive sulfonate targets were detected in both soluble and membrane proteomes (51, 52). Interestingly, most proteins showed preferential labeling with specific members of the sulfonate library, indicating that the varied binding group element was at least in part specifying the proteome reactivity of the probes. Biotin-conjugated sulfonate probes were used to affinity isolate several protein targets, and these proteins were identified by mass spectrometry methods as enzymes from nine mechanistically distinct classes (Table II). Each enzyme was recombinantly expressed in COS-7 cells to confirm its sensitivity to sulfonate labeling. For several enzymes, additional evidence was obtained that sulfonate probes were modifying the active site. For example, the addition of cofactors or substrates was found to reduce the labeling of some enzymes (52, 56), while the sulfonate reactivity of other proteins was either enhanced or inhibited by known allosteric regulators of catalytic activity (56). For one enzyme, aldehyde dehydrogenase, the sulfonate probes were also shown to act as active site-directed irreversible inhibitors (51). Finally, it is notable that several of the sulfonate targets, including glutathione *S*-transferase GSTO1-1 (51), tissue transglutaminase (56), and platelet-type phosphofructokinase (56), were found to be up-regulated in invasive breast cancer cells, indicating that non-directed methods for ABPP can identify novel protein markers of discrete pathological states.

In summary, through the development and application of non-directed methods for ABPP, Adam and colleagues have shown that activity-based chemical probes compatible with whole-proteome analysis can be generated for numerous enzyme classes (51, 52, 56). Strikingly, none of the enzymes labeled by the sulfonate library represented targets of previously described proteomics probes. This finding suggests that researchers in proteomics are still far away from saturating the amount of "active site space" addressable with chemical probes. Nonetheless, successful attempts to further expand the scope of ABPP will likely require probe libraries of considerable chemical and structural diversity as well as efficient strategies to screen proteomes for targets of these profiling tools. The future design of ABPP probes would also benefit from a deeper understanding of the parameters that drive the probe-enzyme reactions observed to date with the sulfonate library. The absence of a shared catalytic mechanism among the sulfonate targets argues that other features are dictating probe labeling. Efforts to identify the specific sites of sulfonate modification on the targeted enzymes may help to define the molecular properties that support active site-directed labeling events. Finally, it is worth considering how often active site-directed labeling actually equates with an "activity-based" event. For certain enzymes like tissue transglutaminase, sulfonate labeling appears to provide an exquisite readout of catalytic activity as both properties show a strict requirement for calcium and inhibition by the allosteric regulator GTP (56). On the other hand, for some enzymes, active site modification may occur on non-catalytic residues, akin to the manner in which microcystin labels a non-catalytic cysteine residue in the active sites of serine/threonine phosphatases (54). Can such labeling events be considered activity-based? From a pure mechanistic standpoint, the answer

TABLE II
Representative enzyme activities identified from mouse and human proteomes by non-directed ABPP methods using a sulfonate ester probe library

Enzyme	Enzyme class	Proteome source
Acetyl-CoA acetyltransferase[a]	Thiolase	Mouse heart
Aldehyde dehydrogenase 1[b]	Aldehyde dehydrogenase	Mouse heart, rat liver
Aldehyde dehydrogenase 7[a]	Aldehyde dehydrogenase	Mouse heart
Dihydrodiol dehydrogenase[a]	NAD/NADP-dependent oxidoreductase	Mouse heart
Enoyl-CoA hydratase, peroxisomal[a]	Enoyl-CoA hydratase	Mouse heart Human breast cancer line
Epoxide hydrolase, cytoplasmic[a]	Epoxide hydrolase	Mouse heart
GSTO1-1[a]	Glutathione *S*-transferase	Human breast cancer line
3β-Hydroxysteroid dehydrogenase Δ5-isomerase-1[c]	3β-Hydroxysteroid dehydrogenase	Mouse testis
Platelet phosphofructokinase[c]	Phosphofructokinase	Human breast cancer line
Type II tissue transglutaminase[c]	Transglutaminase	Human breast cancer line

[a] See Ref. 52.
[b] See Ref. 51.
[c] See Ref. 56.

would be no; however, from a more biological perspective, if, as is often the case, enzyme activity is regulated *in vivo* by autoinhibitory domains, protein partners, and/or small molecules that sterically obstruct the active site (7), then any probe that is sensitive to these molecular interactions would provide an effective readout of the functional state of the enzyme in the context of the cell biology of the proteome.

CONCLUSIONS AND FUTURE DIRECTIONS

In this review, we have highlighted a promising new class of proteomic methods that has united the fields of synthetic chemistry, protein biochemistry and cell biology to create powerful tools and assays for the global analysis of protein expression and function. Chemical approaches like the ICAT method offer researchers in proteomics the opportunity to compare the expression level of low abundance proteins in samples of high complexity (16, 18). The extension of ICAT methods to chemical probes specific for phosphorylated peptides has engendered assays to monitor changes in the post-translational modification state of proteins in cell and tissue proteomes (26–28). Finally, both directed and non-directed strategies for ABPP have produced a menu of chemical probes that can be used either separately or in combination to discover enzyme activities associated with discrete physiological and/or pathological states (29, 32, 43, 52, 56). The value of ABPP as a method for functional proteome analysis has been further highlighted by its application as a screen to evaluate the potency and selectivity of enzyme inhibitors (30, 32). Nonetheless, despite the considerable advances made to date, chemical approaches for proteome analysis still face significant technical challenges. Perhaps most notably, an unsatisfying trade-off seems to exist between the need for high sample throughput and the desire for in-depth analysis of individual proteomes. For example, with a one-dimensional gel format, hundreds of proteomic samples treated with ABPP probes can be readily analyzed in a single day by a given academic laboratory (43). However, the modest resolution afforded by one-dimensional gels likely will result in some low abundance and/or co-migrating protein targets eluding detection. In contrast, proteomic investigations that utilize LC as a separation method can achieve exceptional resolution of chemically tagged peptides but with a much lower sample throughput. In the end, the optimal platform with which to analyze probe-labeled proteomes will likely depend on the biological question being addressed. Indeed, if extensive detail is sought on a select number of samples, then one may wish to apply all of the proteomic methods described above, thereby approaching a complete picture of dynamics in protein abundance, modification state, and activity.

Acknowledgments—We thank the Cravatt and Sorensen laboratories for helpful discussions.

* This work was supported by NCI, National Institutes of Health Grant CA87660, the California Breast Cancer Research Program, Activx Biosciences, and the Skaggs Institute for Chemical Biology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Dept. of Cell Biology, The Scripps Research Inst., 10550 N. Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-784-8633; Fax: 858-784-2798; E-mail: cravatt@scripps.edu.

REFERENCES

1. Anderson, N. L., and Anderson, N. G. (1998) Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19,** 1853–1861
2. Pandey, A., and Mann, M (2000) Proteomics to study genes and genomes. *Nature* **405,** 837–846
3. Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J. C. (2000) The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **21,** 1104–1115
4. Nelson, P. S., Han, D., Rochon, Y., Corthals, G. L., Lin, B., Monson, A., Nguyen, V., Franza, B. R., Plymate, S. R., Aebersold, R., and Hood, L. (2000) Comprehensive analysis of prostate gene expression: convergence of expressed sequence tag databases, transcript profiling and proteomics. *Electrophoresis* **21,** 1823–1831

5. Santoni, V., Molloy, M., and Rabilloud, T. (2000) Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **21,** 1054–1070
6. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9390–9395
7. Kobe, B., and Kemp, B. E. (1999) Active site-directed protein regulation. *Nature* **402,** 373–376
8. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae. Nature* **403,** 623–627
9. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 4569–4574
10. MacBeath, G., and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289,** 1760–1763
11. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105
12. Cravatt, B. F., and Sorensen, E. J. (2000) Chemical strategies for the global analysis of protein function. *Curr. Opin. Chem. Biol.* **4,** 663–668
13. Griffin, T. J., and Aebersold, R. (2001) Advances in proteome analysis by mass spectrometry. *J. Biol. Chem.* **276,** 45497–45500
14. Patton, W. F. (2000) A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* **21,** 1123–1144
15. Lopez, M. F., Berggren, K., Chernokalskaya, E., Lazarev, A., Robinson, M., and Patton, W. F. (2000) A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis* **21,** 3673–3683
16. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999
17. Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19,** 946–951
18. Zhou, H., Ranish, J. A., Watts, J. D., and Aebersold, R. (2002) Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nat. Biotechnol.* **19,** 512–515
19. Smolka, M., Zhou, H., and Aebersold, R. (2002) Quantitative protein profiling using two-dimensional gel electrophoresis, isotope-coded affinity tag labeling, and mass spectrometry. *Mol. Cell. Proteomics* **1,** 19–29
20. Ahn, N. G., and Resing, K. A. (2001) Toward the phosphoproteome. *Nat. Biotechnol.* **19,** 317–318
21. Kang, S., Liao, P., Gage, D. A., and Esselman, W. J. (1997) Identification of *in vivo* phosphorylation sites of CD45 protein-tyrosine phosphatase in 70Z/3.12 cells. *J. Biol. Chem.* **272,** 11588–11596
22. Immler, D., Gremm, D., Kirsch, D., Spengler, B., Presek, P., and Meyer, H. E. (1998) Identification of phosphorylated proteins from thrombin-activated human platelets isolated by two-dimensional gel electrophoresis by electrospray ionization-tandem mass spectrometry (ESI-MS/MS) and liquid chromatography-electrospray ionization mass-spectrometry (LC-ESI-MS). *Electrophoresis* **19,** 1015–1023
23. Andersson, L., and Porath, J. (1986) Isolation of phosphoproteins by immobilized metal (Fe^{3+}) affinity chromatography. *Anal. Biochem.* **154,** 250–254
24. Blaydes, J. P., Vojtesek, B., Bloomberg, G. B., and Hupp, T. R. (2000) The development and use of phospho-specific antibodies to study protein phosphorylation. *Methods Mol. Biol.* **99,** 177–189
25. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae. Nat. Biotechnol.* **20,** 301–305
26. Oda, Y., Nagasu, T., and Chait, B. T. (2001) Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **19,** 379–382
27. Goshe, M. B., Conrads, T. P., Panisko, E. A., Angell, N. H., Veenstra, T. D., and Smith, R. D. (2001) Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.* **73,** 2578–2586
28. Zhou, H., Watts, J. D., and Aebersold, R. (2001) A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19,** 375–378
29. Liu, Y., Patricelli, M. P., and Cravatt, B. F. (1999) Activity-based protein profiling: the serine hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 14694–14699
30. Kidd, D., Liu, Y., and Cravatt, B. F. (2001) Profiling serine hydrolase activities in complex proteomes. *Biochemistry* **40,** 4005–4015
31. Faleiro, L., Kobayashi, R., Fearnhead, H., and Lazebnik, Y. A. (1997) Multiple species of CPP32 and Mch2 are the major active caspases present in apoptotic cells. *EMBO J.* **16,** 2271–2281
32. Greenbaum, D., Medzihradszky, K. F., Burlingame, A., and Bogyo, M. (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* **7,** 569–581
33. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921
34. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., *et al.* (2001) The sequence of the human genome. *Science* **291,** 1304–1351
35. Goldsack, N. R., Chambers, R. C., Dabbagh, K., and Laurent, G. J. (1998) Thrombin. *Int. J. Biochem. Cell Biol.* **30,** 641–646
36. Kraut, J. (1977) Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46,** 331–358
37. Andreasen, P. A., Kjoller, L., Christensen, L., and Duffy, M. J. (1997) The urokinase-type plasminogen activator system in cancer metastasis: a review. *Int. J. Cancer* **72,** 1–22
38. Dessen, A. (2000) Structure and mechanism of human cytosolic phospholipase A(2). *Biochem. Biophys. Res. Commun.* **1488,** 40–47
39. Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S. S., Friedmann, T., and Taylor, P. (1986) Primary structure of acetylcholinesterase: implications for regulation and function. *Fed. Proc.* **45,** 2976–2981
40. Cravatt, B. F., Giang, D. K., Mayfield, S. P., Boger, D. L., Lerner, R. A., and Gilula, N. B. (1996) Molecular characterization of an enzyme that degrades neuromodulatory fatty acid amides. *Nature* **384,** 83–87
41. Walsh, C. T. (1979) *Enzymatic Reaction Mechanisms*, pp. 53–107, W. H. Freeman and Company, New York
42. Patricelli, M. P., Giang, D. K., Stamp, L. M., and Burbaum, J. J. (2001) Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1,** 1067–1071
43. Jessani, N., Liu, Y., Humphrey, M., and Cravatt, B. (2002) Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 10335–10340
44. Thornberry, N. A., Peterson, E. P., Zhao, J. J., Howard, A. D., Griffin, P. R., and Chapman, K. T. (1994) Inactivation of interleukin-1β converting enzyme by peptide (acyloxy)methyl ketones. *Biochemistry* **33,** 3934–3940
45. Martins, L. M., Kottke, T., Mesner, P. W., Basi, G. S., Sinha, S., Frigon, N., Jr., Tatar, E., Tung, J. S., Bryant, K., Takahashi, A., Svingen, P. A., Madden, B. J., McCormick, D. J., Earnshaw, W. C., and Kaufmann, S. H. (1997) Activation of multiple interleukin-1β converting enzyme homologues in cytosol and nuclei of HL-60 cells during etoposide-induced apoptosis. *J. Biol. Chem.* **272,** 7421–7430
46. Winssinger, N., Ficarro, S., Schultz, P. G., and Harris, J. L. (2002) Profiling protein function with small molecule microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 11139–11144
47. Baruch, A., Greenbaum, D., Levy, E. T., Nielsen, P. A., Gilula, N. B., Kumar, N. M., and Bogyo, M. (2001) Defining a link between gap junction communication, proteolysis, and cataract formation. *J. Biol. Chem.* **276,** 28999–29006
48. Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K. F., and Bogyo, M. (2002) Chemical approaches for functionally probing the proteome. *Mol. Cell. Proteomics* **1,** 60–68
49. Lo, L.-C., Pang, T.-L., Kuo, C.-H., Chiang, Y.-L., Wang, H.-Y., and Lin, J.-J. (2002) Design and synthesis of class-selective activity probes for protein tyrosine phosphatases. *J. Proteome Res.* **1,** 35–40

50. Myers, J. K., and Widlanski, T. S. (1993) Mechanism-based inactivation of prostatic acid phosphatase. *Science* **262,** 1451–1453
51. Adam, G. C., Cravatt, B. F., and Sorensen, E. J. (2001) Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem. Biol.* **8,** 81–95
52. Adam, G. C., Sorensen, E. J., and Cravatt, B. F. (2002) Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat. Biotechnol.* **20,** 805–809
53. Wymann, M. P., Bulgarelli-Leva, G., Zvelebil, M. J., Pirola, L., Vanhaesebroeck, B., Waterfield, M. D., and Panayotou, G. (1996) Wortmannin inactivates phosphoinositide 3-kinase by covalent modification of Lys-802, a residue involved in the phosphate transfer reaction. *Mol. Cell. Biol.* **16,** 1722–1733
54. Runnegar, M., Berndt, N., Kong, S.-M., Lee, E. Y. C., and Zhang, L. (1995) *In vivo* and *in vitro* binding of microcystin to protein phosphatases 1 and 2A. *Biochem. Biophys. Res. Commun.* **216,** 162–169
55. Liu, S., Widom, J., Kemp, C. W., Crews, C. M., and Clardy, J. (1998) Structure of human methionine aminopeptidase-2 complexed with fumagillin. *Science* **292,** 1324–1327
56. Adam, G. C., Sorensen, E. J., and Cravatt, B. F. (2002) Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes. *Mol. Cell. Proteomics* **1,** 828–835