

AU1

## Inhibitor Focusing: Direct Selection of Drug Targets from Proteomes Using Activity-Based Probes

Tyzoon K. Nomanbhoy, Jonathan Rosenblum, Arwin Aban, and Jonathan J. Burbaum

**Abstract:** In the latter stages of drug discovery and development, assays that establish drug selectivity and toxicity are important when side effects, which are often due to lack of specificity, determine drug candidate viability. There has been no comprehensive or systematic methodology to measure these factors outside of whole-animal assays, and such phenomenological assays generally fail to establish the additional targets of a given small molecule, or the molecular origin of toxicity. Consequently, small-molecule development programs destined for failure often reach advanced stages of testing, and the money and time invested in such programs could be saved if information on selectivity were available early in the process. Here, we present a methodology that utilizes chemical ABPs in combination with small-molecule inhibitors to selectively label small-molecule binding sites in whole proteomic samples. In principle, the ABP and small molecule will compete for similar binding sites, such that the small molecule will protect against modification by the ABP. Thus, after removal of the small molecule, the binding site for the ABP will be revealed, and a second probe can then be used to label the small-molecule binding sites selectively. To demonstrate this experimentally, we mapped the binding sites of the DPP4 inhibitor, IT, in a number of different tissue types.

### Introduction

NOW THAT THE SEQUENCE OF THE HUMAN GENOME has been substantially finished, the leading edge of pan-biological analysis has turned to proteomics.<sup>1-3</sup> Proteomics is defined as “the analysis of complete complements of proteins . . . {including} not only the identification and quantification of proteins, but also the determination of their localization, modifications, interactions, activities, and, ultimately, their function.”<sup>4</sup> However, the methodologies of proteomics, compared with those of genomics, are relatively insensitive and cumbersome. For instance, in proteomics, there are no equivalent methods to the PCR-based amplification and high-throughput sequencing methods that have proven to be essential for genomics. Further, although one organism contains but a single genome, the same organism contains a very large number of proteomes that generally

change with tissue type, developmental stage, experimental perturbation, and disease condition. This enormous complexity, combined with the relatively cumbersome analytical methods of protein science (when compared with those used in nucleic acid analysis), presents both a hurdle to the widespread application of proteomics to drug discovery and an opportunity for technological advancement.

Here, we present a general experimental approach toward the detection and quantification of the proteomic fraction having the highest level of interest in the drug discovery community—the fraction that contains active proteins that bind small-molecule drug candidates. This approach uses ABPs, which have been developed both at ActivX Biosciences and elsewhere<sup>5-8</sup> (Fig. 1). Such probes react class-specifically, but nonselectively, with broad classes of proteins by covalently modifying particular conserved residues in the active site of the protein

F1

ActivX Biosciences, Inc., La Jolla, California

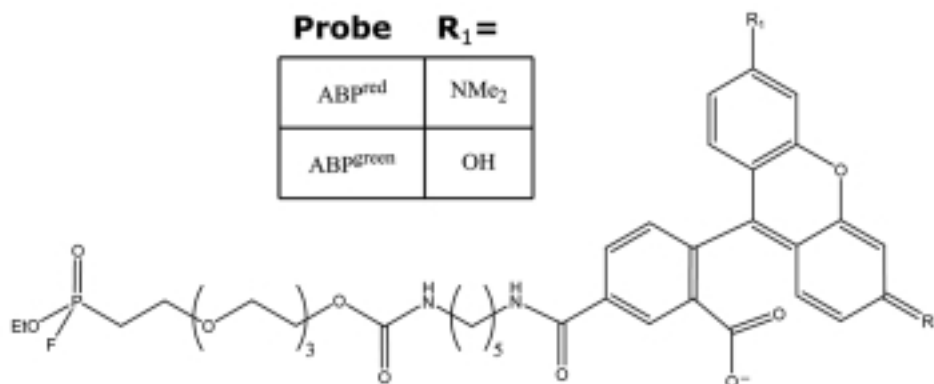
\*The use of two-dimensional gel electrophoresis, a staple of proteomics analysis, also results in multiple proteins per band.<sup>16</sup>

**ABBREVIATIONS:** ABP, activity-based probe; 1D, one-dimensional; DPP4, DPP8, and DPP9, dipeptidyl peptidase 4, 8, and 9, respectively; FP, fluorophosphonate; IT, isoleucyl thiazolidine; LC, liquid chromatography; MS, mass spectrometry; PEG, polyethylene glycol; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TAMRA, tetramethylrhodamine.

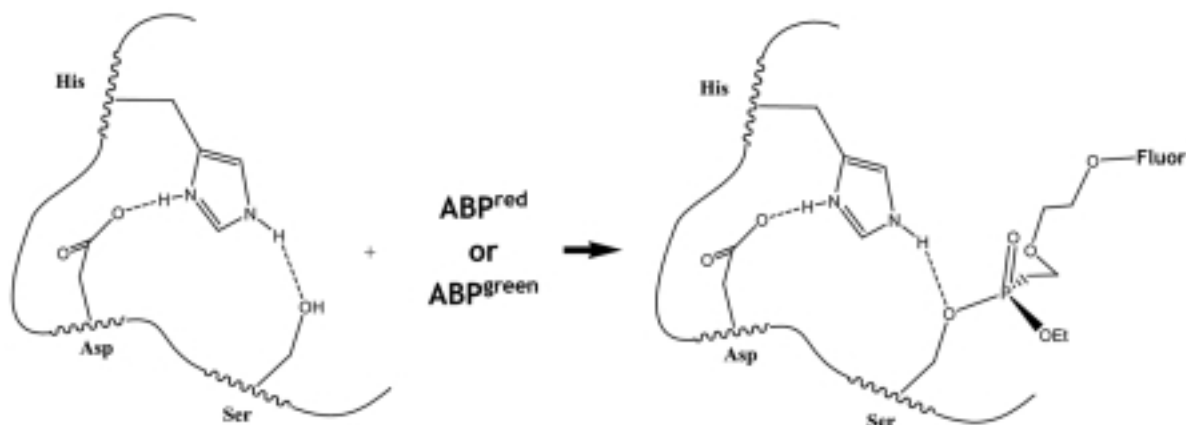
AU2

AU3

A



B



**FIG. 1.** Interaction of ABP with serine hydrolase family members. **(A)** The structure of the two ABPs used in this study, ABP<sup>red</sup> and ABP<sup>green</sup>, are shown. ABPs in general consist of a class-selective reactive group (ethoxyfluorophosphate, in this case), a linker, and a detection group (TAMRA for ABP<sup>red</sup> and fluorescein for ABP<sup>green</sup>). **(B)** The active site of the serine hydrolases is shown in schematic form. The classic “catalytic triad” of the serine hydrolases consists of an aspartic acid, a histidine, and a serine, which serves as a nucleophilic catalyst for hydrolysis of carboxamides, carboxylate esters, and the like. The reactive group of the ABP recognizes the active site, and reacts specifically with the serine nucleophile to provide a covalent adduct (with the release of fluoride as a leaving group).

class. Because small-molecule drug candidates also interact with these active sites, the presence of a bound drug candidate will generally block the interaction of ABPs with their targets. Upon removal of the drug, the reactive residues will once again become available for reaction with the probe. This unmasking of reactivity provides the basis for a chemical footprinting approach that provides not only the identification of all sites of interaction within a proteomic sample, but also a methodology for the specific enrichment of proteins that possess a particular small-molecule binding site.

Other methods for protein footprinting have been described. Recently, Richards *et al.* reported a method using the very reactive molecule methylene, coupled with electrospray MS, to define solvent-accessible, nonpolar surfaces in proteins.<sup>9</sup> The use of limited proteolysis (akin to nuclease protection assays in the nucleotide world) to footprint protein interactions was quite common before

the advent of high-resolution structural methods, and continues to provide some useful information.<sup>10</sup> Such methods are valuable for interactions that occur over large surface areas, but would be expected to have little value in small-molecule binding. Methods that seek to label metal-ion binding sites inferred from chemical protection data have also been reported.<sup>11</sup> Generally, however, these methods have been used in highly purified systems to provide insight into specific issues of protein structure. We describe here a method that can be used in crude proteomic mixtures to discover proteins actively engaged in binding small molecules.

The significance of this methodology lies in the unmet need for mechanistic insight into clinical drug candidates. Generally, clinical candidates are optimized against one “primary” target, with selectivity and/or toxicity information considered as additional “secondary” criteria for optimization. Targets that are used to analyze a drug can-

didate's selectivity are generally chosen in an *ad hoc* manner, often through sequence similarity or availability of a suitable assay or panel of assays. Such targets rarely receive the in-depth biological validation of the primary targets. Toxicity-related protein targets, which represent a special case of selectivity where the biological effect is harmful, are generally completely unanticipated. The importance of understanding such "off-target" activities is, however, vital. Species-specific efficacy and/or toxicity are often the death knell of a drug development program. Millions of dollars could be saved through early termination of such drug development programs if the interactions of the drug candidate were better understood up front, at a proteomic level, especially if the method were amenable to small-scale assays with little or no optimization necessary.

F2

Here, we present a method that uses two ABPs that target the same proteins (Fig. 2). In this method, a proteomic sample containing active proteins of interest is first saturated with the small-molecule candidate, and then labeled with one ABP. The small molecule is subsequently removed (*e.g.*, via dialysis or gel filtration) and the sample is re-labeled with a second ABP. The first probe need not contain a detection group, although in practice, this allows for data normalization and quality control. The second probe identifies the sites for small-molecule interaction. Because the reaction of a probe with a target protein provides both the foundation for a biochemical assay for the target regardless of its identity and a handle for purification and identification of the targeted proteins, this simple analysis provides for a more comprehensive and unbiased analysis of off-target activity and its consequent toxicity than existing methods.

## Materials and Methods

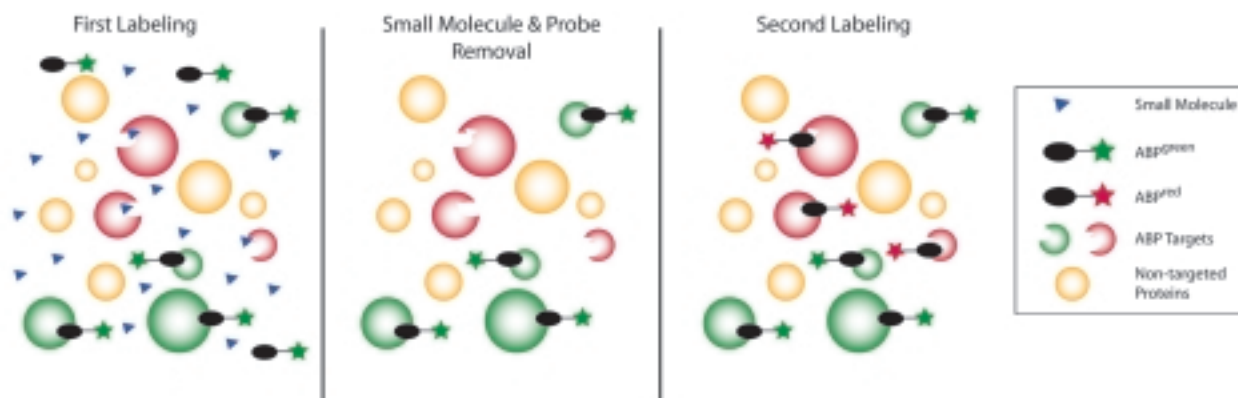
### Reagents

The synthesis of the fluorescent FP derivatives has been previously described.<sup>8</sup> ABP<sup>green</sup> is an FP-based probe that targets serine hydrolases. In this probe, the reactive group is tethered with a PEG spacer to a fluorescein reporter group, and was previously described as FP-PEG-fluorescein. In contrast, ABP<sup>red</sup> is identical except for the fluorescent group, which is TAMRA, and was previously described as FP-PEG-TAMRA. IT was synthesized according to the published procedures.<sup>12</sup> Other reagents used were obtained from Sigma (St. Louis, MO, U.S.A.).

### Preparation of proteome samples

Frozen mouse kidneys were purchased from Pel-Freeze Biologicals. The fractionation of the tissue into soluble and membrane components has been previously described.<sup>8</sup> The A431 human epidermoid cell line was maintained in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. Upon reaching confluence, the cells were washed once with Ca<sup>2+</sup>- and Mg<sup>2+</sup>-free Hanks' medium, and then incubated in the same buffer for 30 min at 37°C, resulting in the detachment of the cells from the tissue culture plates. The cells were then collected by centrifugation at 500 g for 10 min. The cells were resuspended in HEPES buffer (20 mM, pH 7.5) containing 100 mM NaCl, and lysed by Dounce homogenization. The soluble and membrane components were then separated by centrifugation at 15,000 g for 15 min at 4°C. After preparation, protein concentrations were de-

4C



**FIG. 2.** Schematic of Inhibitor Focusing approach. In the first labeling step, the small molecule (represented by blue triangles) is incubated at a concentration exceeding its dissociation constant from the proteins of interest, and excess ABP is added (ABP<sup>green</sup>). The ABP then recognizes the active-site configuration and reacts with proteins that are not targeted by the small molecule (the green spheres). In the second step, the (reversible) small molecule, as well as the excess ABP<sup>green</sup>, is removed via dialysis or gel filtration. Then the proteins that were targets of the small molecule (the red proteins) become available for reaction with ABP in a second labeling step. The labeling in this instance uses another color ABP (ABP<sup>red</sup>) to select these proteins.

terminated using the Bio-Rad Protein Assay kit, following the manufacturer's directions. Proteome samples were either used immediately or stored in aliquots at  $-80^{\circ}\text{C}$ .

#### *ABP labeling of proteomes*

Mouse kidney membranes were diluted to 3 mg/ml with 20 mM HEPES buffer (20 mM, pH 7.5) containing 100 mM NaCl and 0.1% CHAPS. The membranes were then centrifuged at 15,000  $g$  for 15 min to remove detergent-insoluble material. The supernatant, containing solubilized membrane proteins, was labeled with ABP (6  $\mu\text{M}$ ) for 60 min, in the absence or presence of IT (100  $\mu\text{M}$ ). The concentration of the soluble A431 lysate was adjusted to 2 mg/ml with 20 mM HEPES (pH 7.5), 100 mM NaCl. The sample was labeled with ABP (6  $\mu\text{M}$ ) for 60 min, in the absence or presence of IT (250  $\mu\text{M}$ ).

#### *Direct selection of drug targets*

Either solubilized kidney membranes (400  $\mu\text{l}$  at 3 mg/ml) or soluble A431 lysate (400  $\mu\text{l}$  at 2 mg/ml) was reacted with ABP<sup>green</sup> (6  $\mu\text{M}$ ) in the presence of IT (100  $\mu\text{M}$  for solubilized kidney membranes, 250  $\mu\text{M}$  for the A431 lysate) for 1 h. The free ABP<sup>green</sup> and the free IT were removed by running the sample through an Amersham HiTrap Desalting Column (Amersham Biosciences, Piscataway, NJ, U.S.A.). The resulting samples were then reacted with ABP<sup>red</sup> (6  $\mu\text{M}$ ) for 1 h.

#### *Image analysis and data reduction*

Proteomes labeled with ABP were separated on 8 cm  $\times$  10 cm Laemmli minigels (12.5% acrylamide/bisacrylamide, 37.5:1)<sup>13</sup> for 45 min at 250 V. The gels were then scanned on a Hitachi FMBio II flatbed fluorescence scanner, with excitation provided by the 532-nm line of a 50-mW neodymium-doped yttrium-aluminum-garnet (Nd:YAG) laser. A 605-nm bandpass filter was used to detect ABP<sup>red</sup> fluorescence, whereas a 505-nm bandpass was used to detect ABP<sup>green</sup>. The scanned images were subsequently analyzed with Image Analysis 3.0 software (MiraiBio, Alameda, CA, U.S.A.), which provided the side traces shown in the comparisons.

#### *MS for protein identification*

ABP<sup>red</sup>-labeled proteins were isolated by an immunoaffinity approach (Patricelli *et al.*, manuscript in preparation). The isolated proteins were separated on Laemmli minigels as described above. Gel bands of interest were excised using a spot picker (Amersham Pharmacia Biosciences, Piscataway, NJ, U.S.A.) into polypropylene 96-well microplates following the manufacturer's instructions. After excision, gel pieces were washed three times sequentially with water (50  $\mu\text{l}$ ) followed by ammonium bicarbonate (100 mM)/acetonitrile

(70:30; 50  $\mu\text{l}$ ). The gel pieces were dried using a commercial evaporator (TurboVap<sup>®</sup>-96, Zymark, Hopkinton, MA, U.S.A.). For tryptic digest, gel pieces were incubated at  $37^{\circ}\text{C}$  overnight in porcine trypsin (20 ng per gel piece: Promega, Madison, WI, U.S.A.) in ammonium bicarbonate (20 mM). Samples were then extracted with 40:60:0.1 water/acetonitrile/trifluoroacetic acid, then analyzed by nanospray LC/MS. Peptides were eluted using a gradient from 10% solvent A (water + 0.1% formic acid) to 95% solvent B (acetonitrile + 0.08% formic acid) over 40 min. Detection was achieved with an LC/Q Deca XP mass spectrometer (Thermo Finnigan, San Jose, CA, U.S.A.), with data analysis performed using Mascot (Matrix Science Ltd., London, U.K.). Spectra were collected in data-dependent mode (one full MS scan followed by two MS/MS scans on the first and second most intense ions).

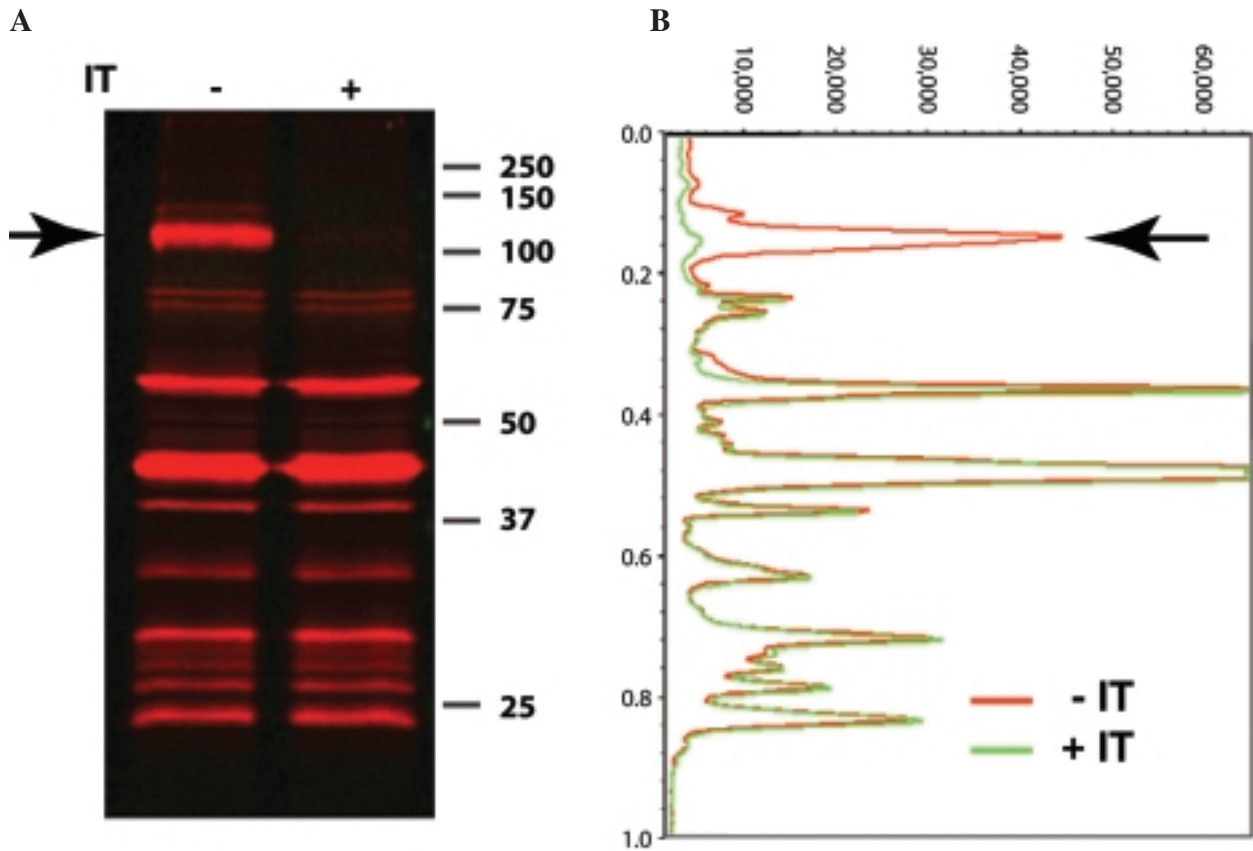
## **Results**

Chemical reagents known as ABPs have been developed to target several families of proteins. For the serine hydrolase superfamily of proteins, these probes consist a detection group (in the first generations,<sup>14,15</sup> biotin) tethered to an FP electrophile that reacts specifically with the active-site serine. Labeling of a proteome with this probe, followed by separation of the components using 1D SDS-PAGE, allows for the detection of up to 40 distinct proteins.<sup>8</sup> The number of bands that can be distinguished depends on the proteome used. Although this approach significantly reduces the complexity of the proteome (from some  $10^5$  proteins down to  $\sim 10^1$ – $10^2$ ), many bands still consist of more than a single protein.\* More significantly, the dynamic range in the abundance (estimated to be  $\sim 10^7$  for all proteins<sup>17</sup>) exceeds the dynamic range of fluorescence detection ( $\sim 10^3$ ) or identification ( $\sim 10^2$ ), resulting in the loss of potentially valuable information. Hence, it is likely that important information from the less abundant proteins may be overlooked. We have therefore considered alternative methods to reduce the complexity of a proteome labeled with a probe still further.

ABPs can be used to identify reversible inhibitors for particular enzymes in a proteome.<sup>14</sup> In this process, a proteome is labeled with an ABP in the presence or absence of an inhibitor. The presence of the inhibitor blocks the ABP from the active sites of its target. Upon SDS PAGE, this can be detected by the disappearance of one or more bands when compared with the proteome labeled with the ABP in the absence of the inhibitor. As an example,

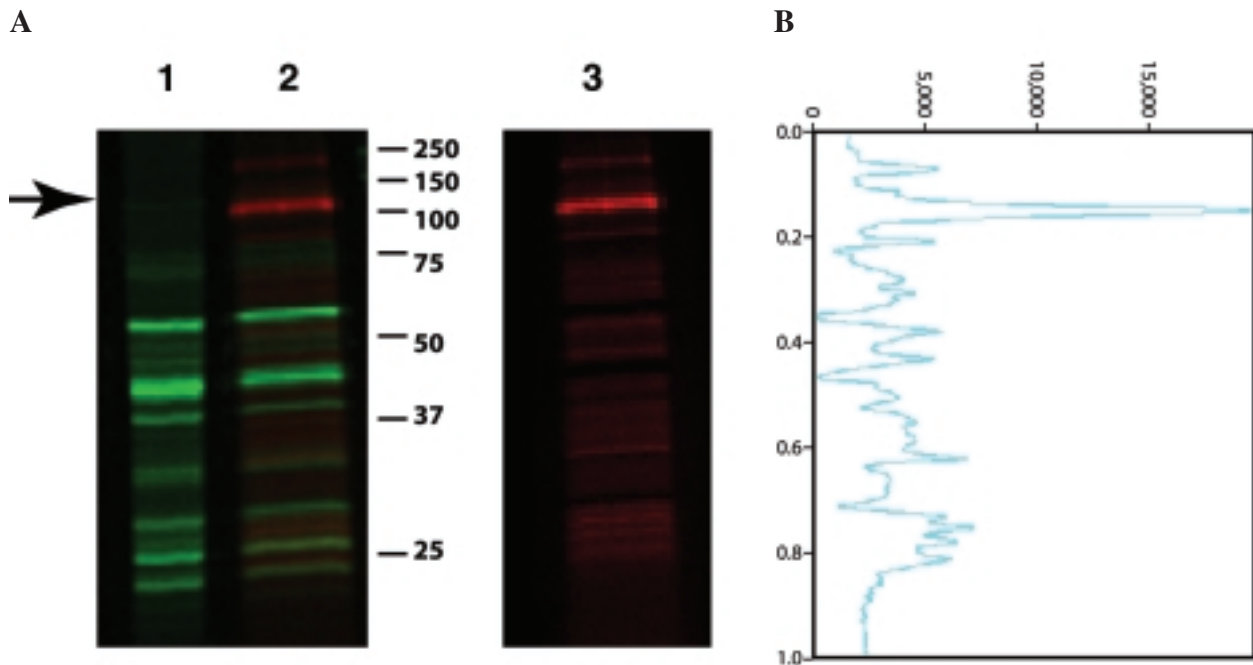
The use of two-dimensional gel electrophoresis, a staple of proteomics analysis, also results in multiple proteins per band.<sup>16</sup>

QU3



4C

**FIG. 3.** Inhibition of proteins in mouse kidney membranes by IT as observed by 1D SDS-PAGE. Proteome labeling was carried out as described, in the absence (- lane) and presence (+ lane) of IT. **A:** shows the raw fluorescence scan of the gels. **B:** shows the side traces of the gel data, and the arrow denotes the position of the major inhibited band (DPP4).



4C

**FIG. 4.** Inhibition of proteins in mouse kidney membranes by IT as observed by Inhibitor Focusing. In lane 1, proteome labeling was carried out as in Fig. 2, except ABP<sup>green</sup> was used. Lane 2 represents the result of the Inhibitor Focusing procedure used on this sample, using ABP<sup>red</sup> as the second probe. In this case, the green channel detects both ABP<sup>green</sup>-labeled proteins, whereas the red channel detects ABP<sup>red</sup>. Lane 3 is identical to lane 2, with only the red channel displayed. The side trace depicts the ABP<sup>red</sup> signal from lane 3.

F3

mouse kidney membranes labeled with ABP<sup>red</sup> leads to the observation of ~15 bands (Fig. 3). A prominent band at ~100 kDa is observed (Fig. 3, lane 1, arrow). In the presence of the DPP4 inhibitor, IT, the 100-kDa band disappears (Fig. 3, lane 2). Mouse DPP4 is a glycosylated protein whose polypeptide portion is 87,436 Da,<sup>18</sup> suggesting that the disappearing band is, in fact, DPP4. To demonstrate this conclusively, an immunoaffinity approach (Patricelli *et al.*, manuscript in preparation) was used to isolate the modified proteins, followed by gel separation and MS-based identity determination. The band at 100 kDa, when analyzed, was found to contain a number of peptides derived from mouse DPP4 (with a total coverage of >30%). We conclude that this band is, in fact, the mouse DPP4 protein, in concordance with the simple explanation. However, this overall approach points out an important shortcoming of such a simple technique, as well as a key feature of our Inhibitor Focusing approach: Because ABPs label active proteins, and small-molecule inhibitors block labeling, the simple approach (*e.g.*, Fig. 3) requires the inference of identity through the loss of a signal. Furthermore, when a small molecule causes only the partial disappearance of a band (due to the band consisting of more than one protein, for example), the resulting loss of signal may be missed altogether. On the other hand, by inverting the labeling approach, the mode of assay measurement is inverted and identity is determined directly (and more robustly) through the gain of a signal.

To reduce further the complexity of ABP targets in a proteome, we exploited the ability of reversible inhibitors to compete with the labeling of specific enzymes by an ABP (see Fig. 2). The labeling of a proteome with an ABP, ABP<sup>green</sup>, in the presence of an inhibitor, will lead to the covalent attachment of ABP<sup>green</sup> to the active sites of all its targets, except for the targets whose active sites were blocked by the inhibitor. The free inhibitor (and free ABP<sup>green</sup>) can be easily and gently removed through dialysis or gel filtration. The subsequent addition of a second ABP (ABP<sup>red</sup>), having the same specificity as ABP<sup>green</sup> but with a distinct fluorophore, will result in the fluorescent signal from ABP<sup>red</sup> associated primarily with targets initially blocked by the inhibitor, thus reducing the complexity of the ABP targets.

F4

Solubilized kidney membranes were labeled with ABP<sup>green</sup> in the presence of IT. Due to inhibition by IT, DPP4 was not labeled in this step (Fig. 4A, lane 1, arrow). The inhibitor and unreacted probe were then removed by gel filtration, followed by labeling of the protein fraction with ABP<sup>red</sup>. At this stage, the label from ABP<sup>red</sup> was primarily found on a band at 100 kDa, which corresponds to DPP4 (Fig. 4A, lane 2 and 3). When kidney membranes were labeled with ABP<sup>red</sup> in the absence of IT (Fig. 3), DPP4 was only the third most intense band seen. In contrast, using Inhibitor Focusing of the pro-

teomic fraction that binds IT, DPP4 was the most significant band (Fig. 4B); no other distinct peaks were observed.

We examined whether Inhibitor Focusing could identify off targets for IT in other tissues by looking at the human epidermoid carcinoma cell line A431, which was not previously known to contain the primary target of IT, DPP4. Further, the soluble fraction derived from these cells was evaluated, whereas DPP4 is strongly associated with the membrane fraction. Upon treatment of A431 cell lysates with ABP<sup>red</sup> in the absence or presence of IT in the manner described above, we observed the inhibition and selective labeling of a single band at ~95 kDa (Fig. 5, arrow). In contrast to DPP4 in kidney membranes, this band is significantly less abundant than the other FP-labeled proteins. The Inhibitor Focusing procedure was performed on the A431 cell lysate as described previously, and the 95-kDa band was selectively labeled in this experiment, and is in fact the major band labeled with ABP<sup>red</sup> (Fig. 6, arrow). The only other significant band labeled with ABP<sup>red</sup> consists of a much larger (~250 kDa) protein. This protein however is the most abundant target for the ABP used in these experiments (Fig. 5), and as it does not appear to be inhibited by IT, its labeling by ABP<sup>red</sup> during the Inhibitor Focusing procedure may reflect the possibility that the labeling by ABP<sup>green</sup> had not gone to completion. However, further exploration of the reaction profile (following the kinetics of labeling with and without inhibitor, for example) would be necessary to fully exclude this protein as a low-affinity target. Nevertheless, the Inhibitor Focusing approach has successfully enriched the signal from a particular ABP (ABP<sup>red</sup>) on the protein targets initially inhibited by IT. Although Inhibitor Focusing can suggest targets and help to enrich proteomic samples for subsequent identification, to demonstrate conclusively that the band is a relevant target requires additional experimentation. To this end, the 95 kDa band indicated by the blue arrow in Fig. 5, a suggested target for the DPP4 inhibitor, IT, was identified by MS. The results from MS analysis are summarized in Table 1. Two proteins were identified in the excised band. The two highest scoring proteins consist of recently identified homologues of DPP4, DPP9<sup>18</sup> and DPP8.<sup>19</sup>

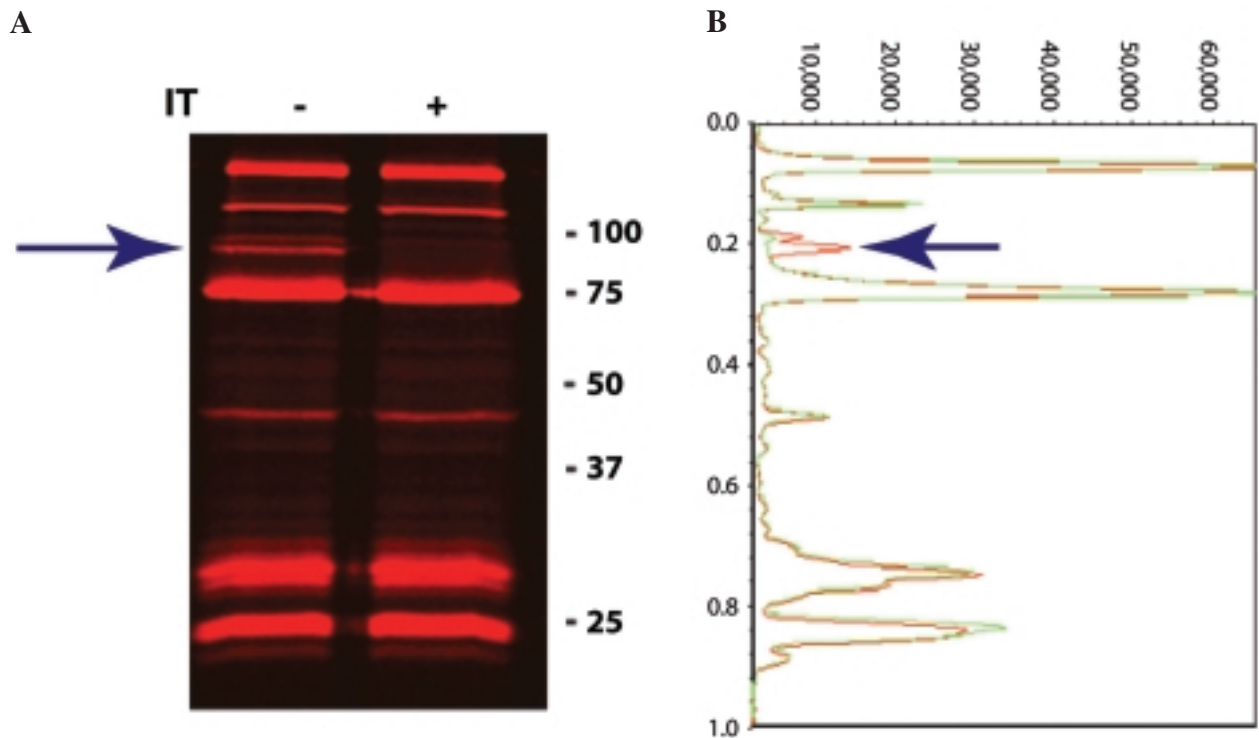
F5

F6

T1

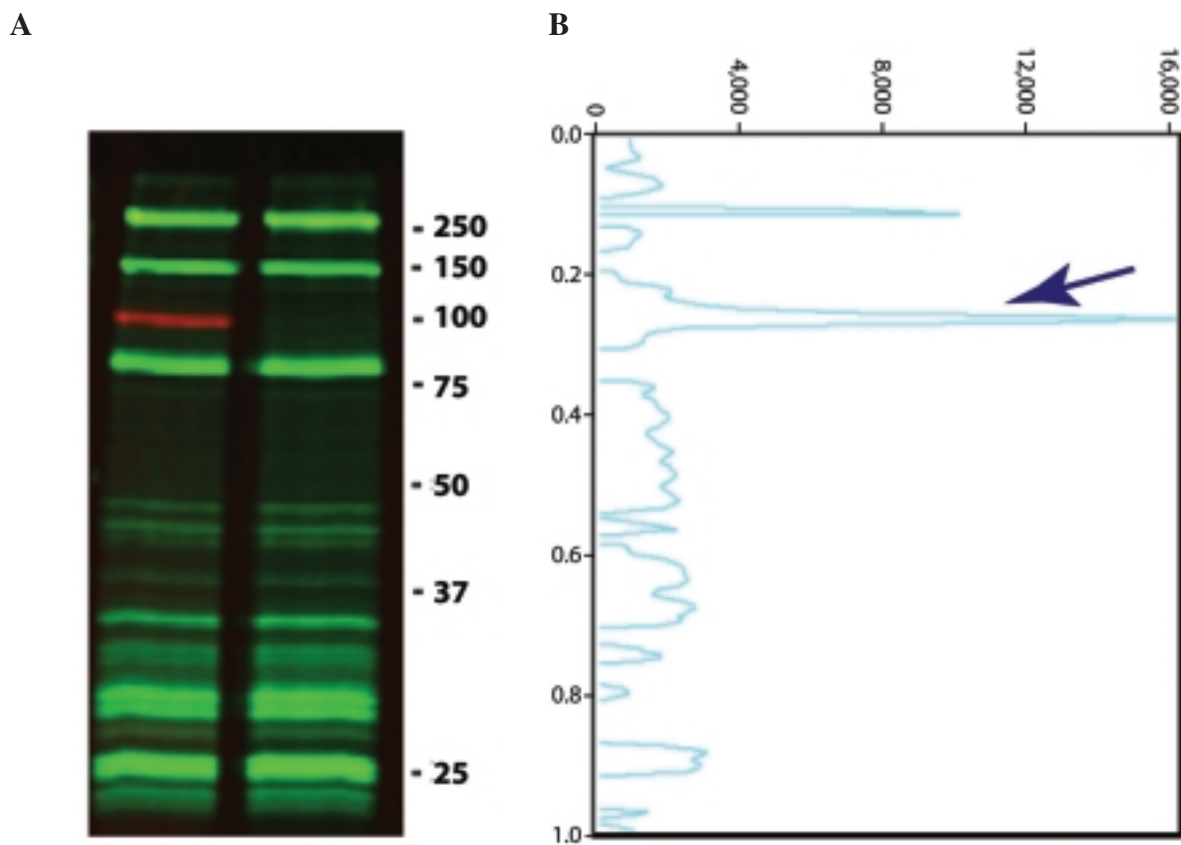
## Discussion

The interaction of IT with DPP4 was chosen as a model system for small molecule–drug target interactions. This choice was made for several reasons. First, DPP4 is a member of the serine hydrolase superfamily and is targeted by FP-based probes. Second, IT is being evaluated for the treatment of type II diabetes, based on the rationale that DPP4 is the principal degradative enzyme for



4C

**FIG. 5.** Inhibition of proteins in A431 soluble fraction by IT as observed by 1D SDS-PAGE. Proteome labeling was carried out as described, in the absence (– lane) and presence (+ lane) of IT. (A) The raw fluorescence scan of the gels, denoting labeling by ABP<sup>red</sup>. (B) The side traces of the gel data. The arrow denotes a significant band, well separated from other signals, that is a target for IT inhibition, identified in the text as DPP8 and DPP9.



4C

**FIG. 6.** Inhibition of proteins in A431 soluble fraction by IT as observed by Inhibitor Focusing. (A) The raw image data. Lane 2 is the result of the first labeling (with ABP<sup>green</sup> in the presence of IT), whereas lane 1 is the result of the overall Inhibitor Focusing procedure (with both ABP<sup>green</sup> and ABP<sup>red</sup>). (B) The ABP<sup>red</sup> signal from lane 1. The arrow denotes the major band apparent in Fig. 4.

TABLE 1. IDENTIFICATION OF NOVEL IT TARGETS

<i>Protein identification</i>	<i>Mascot score</i>	<i>Number of unique peptides observed</i>	<i>Sequence coverage</i>
DPP9	1,020	29	31%
DPP8	279	8	18%

The band indicated by the arrow in Fig. 5 was analyzed by LC/MS/MS. The two proteins tabulated were found to be present in the sample excised from the gel. The Mascot score is related to the quality of the MS.<sup>22</sup> More significantly, the number of unique peptides (which were identified not only by the overall molecular weight of the peptide, but also by MS/MS sequencing of the peptide) indicates that the protein is present in its intact form, as judged from the distribution of the peptides throughout the protein sequence. The sequence coverage represents the fraction of residues that were actually observed, lending further credence to the identification.

glucagon-like peptide 1, a peptide hormone that induces, among other biological effects, the stimulation of glucose-dependent insulin secretion and insulin biosynthesis.<sup>20</sup> Finally, because the drug appears to require relatively high dosages for efficacy (20 mg/kg twice daily in a rat study), the chance for side effects is presumptively high, rendering the clinical development pathway somewhat narrow.

The methods were developed to maximize the signal from DPP4 in kidney extracts. Some of the other targets were seen to react only partially in the first step (labeling with ABP<sup>green</sup>; Fig. 4, lane 1), such that the residual activity resulted in background in the red channel upon completion of the second step. This was observed in the major labeled band in liver (45 kDa; Fig. 3, -IT), and also in the direct-selection enhanced profile (Fig. 4, lane 3). However, the sensitivity we observed for inhibitor targets is significantly better than simple subtraction of side traces from analyzed lanes (*e.g.*, Fig. 3). Simple subtraction provides information only about the inhibition of major species; it is impractical to differentiate minor bands that lie too close to major bands because of the uncertainties in determining small differences between very large signal intensities. Further, under such conditions, it would be an experimental challenge to identify conclusively the inhibited band due to interference by the major species. Through Inhibitor Focusing (*e.g.*, Fig. 6), it is apparent that considerable enhancement of minor species can be effected. Thus, we believe that this approach will allow for the selective enrichment of low-abundance target proteins, and may allow for their unequivocal identification. However, at this stage of development, the techniques described need to be validated further. In the fully developed system, where the kinetics of labeling (both before and after removal of the inhibitor) are more fully understood, the complete com-

putational and experimental removal of background signals can be anticipated, leading to a stand-alone method.

In the general case, the appearance of a band in the directly selected profile (*e.g.*, Fig. 4) may have several possible interpretations. Depending on the dissociation kinetics of the complex, gel filtration to remove the free inhibitor may be insufficient to completely reactivate the target proteins. In the case of DPP4, the association of IT is reversible. Consequently, IT dissociates upon gel filtration, freeing the DPP4 to react with ABP<sup>red</sup>. However, this may not be the case with other targets or inhibitors, and conclusions about specific proteins will require additional experiments to determine the precise nature of the interaction. These types of experiments are also enabled using ABPs, because ABP labeling is quantitatively linked to enzyme activity. Thus, standard enzymological measurements to determine inhibitor reversibility<sup>21</sup> can be performed without further purification of the sample. Another potential pitfall is that improper or idiosyncratic sample handling can decrease or abolish the activity of selected targets, resulting in false negative results. In this particular case, taking into account the twofold dilution of the protein sample as a result of the gel filtration step, the DPP4 peak from the proteome selected for inhibitor binding is ~80% as intense as the DPP4 labeled in the absence of selection. Hence, the initially inhibited DPP4 can be subsequently labeled close to completion. We conclude that the complexity of ABP<sup>red</sup>-labeled targets in this particular proteome has been reduced significantly.

The 95-kDa band that was a target for IT in A431 cells was analyzed by MS and found to be a mixture of primarily two proteins, DPP8 and DPP9. Both these proteins were identified fairly recently and have not been particularly well characterized to date. DPP8 was identified as a novel homologue of DPP4 in 2000, as a protein of 51% similarity to both DPP4 and its close relative, fibroblast activation protein. Preliminary characterization of this protein<sup>19</sup> indicates that it migrates at the size expected from the primary sequence (~100 kDa) and shows a similar inhibitory profile and substrate specificity to DPP4. The other element of these inhibited bands, DPP9, was identified in 2002 and has 37% identity to DPP4. This protein was shown to migrate at a molecular mass of 98 kDa.<sup>16</sup> These literature results are consistent with those reported here. Careful examination of the lane trace in Fig. 5 indicates that the band at 95 kDa, in fact, consists of a doublet (Fig. 5, blue arrow). Based on polypeptide chain length and the lack of reported posttranslational modifications for DPP8 and DPP9, the upper and less abundant of the two proteins probably consists of DPP8, whereas the lower and more abundant of the two proteins probably consists of DPP9. This assignment is also consistent with DPP9

having a higher Mascot score<sup>22</sup> than DPP8 (1,020 and 279, respectively) from MS analysis. The single band seen after the Inhibitor Focusing procedure appears to correspond to DPP9, based on gel mobility. The absence of DPP8 may reflect the possibility that IT is a less potent inhibitor of DPP8 than DPP9. Hence, IT would have a higher dissociation rate from DPP8 than from DPP9, and this would result in the DPP8 active site being more accessible to ABP<sup>green</sup> when IT is present. More extensive biochemical evaluation of the proteins involved will be needed to confirm this possibility.

We have demonstrated here that, in the tissues and cell lines evaluated, IT is reasonably selective toward its principal target, DPP4. We have also demonstrated that profiling using ABPs can detect off-target activity, both in major bands from tissues that lack DPP4 (like A431 cells) and in minor bands in mouse kidney that are highlighted by the process. In the context of a drug discovery program, additional congeners of the drug, or candidates from other lead series, would be available for testing, and the correlation between these off targets and the structure of the inhibitor would naturally lead to physiologically relevant selectivity assays.

## Conclusions

The results presented here illustrate the further application of activity-based proteomic profiling to identify drug targets through the direct detection of binding sites for small-molecule drug candidates. The methodology is conceptually straightforward and is applicable to the full breadth of active site-directed chemical probes, regardless of mechanism. Because methods based on chemical reactivity are not restricted to previously established targets, are independent of species of origin, and are useful on a small scale, the more widespread adoption of such chemical probe-based methods during the drug discovery process will certainly reap significant rewards. Using this method, off targets and plausible effectors of toxicity will be identified more objectively or even discovered in relevant samples. Significantly, the use of ABP-based methods requires neither pure enzyme nor knowledge of a unique, assayable substrate. In addition, through the utilization of ABPs that have broader chemical reactivity than that exhibited by the FPs, it is possible to assess simultaneously off-target activity from a small molecule across several enzyme classes. Finally, potential therapeutics can be evaluated biochemically in the crucial tissues of interest, either using human-derived samples or in the most relevant animal model and tissue. In this way, ABP-based selection of small-molecule targets promises to provide an important addition to the drug discovery process.

## Acknowledgments

We would like to acknowledge Gabriela Tobal for editorial assistance with the manuscript, John Kozarich and Michael Dunn for critical reading, Jennie Lill, Jane Wu, and Jennifer Hanson for MS analysis, and the NIH (SBIR no. 97462-01) for financial assistance.

## References

1. Blackstock WP, Weir MP: Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* 1999;17:121–127.
2. Dove A: Proteomics: translating genomics into products? *Nat Biotechnol* 1999;17:233–236.
3. Martin DB, Nelson PS: From genomics to proteomics: techniques and applications in cancer research. *Trends Cell Biol* 2001;11:S60–S65.
4. Fields S: Proteomics in Genomeland. *Science* 2001;291:1221–1224.
5. Adam GC, Cravatt BF, Sorensen EJ: Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem Biol* 2001;8:81–95.
6. Cravatt BF, Sorensen EJ: Chemical strategies for the global analysis of protein function. *Curr Opin Chem Biol* 2000;4:663–668.
7. Bogyo M, Verhelst S, Bellingard-Dubouchaud V, Toba S, Greenbaum D: Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. *Chem Biol* 2000;7:27–38.
8. Patricelli MP, Giang DK, Stamp LM, Burbaum JJ: Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* 2001;1:1067–1071.
9. Richards F, Lamed R, Wynn R, Patel D, Olack G: Methylene as a possible universal footprinting reagent that will include hydrophobic surface areas: Overview and feasibility: properties of diazirine as a precursor. *Protein Sci* 2000;9:2506–2517.
10. Shea M, Sorensen B, Pedigo S, Verhoeven A: Proteolytic footprinting titrations for estimating ligand-binding constants and detecting pathways of conformational switching of calmodulin. *Methods Enzymol* 2000;323:254–301.
11. Qin K, Yang Y, Mastrangelo P, Westaway D: Mapping Cu(II) binding sites in prion proteins by diethyl pyrocarbonate modification and matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometric footprinting. *J Biol Chem* 2002;277:1981–1990.
12. Hildebrandt M, Arck P, Kruber S, Demuth HU, Reutter W, Klapp B: Inhibition of dipeptidyl peptidase IV (DP IV, CD26) activity abrogates stress-induced, cytokine-mediated murine abortions. *Scand J Immunol* 2001;53:449–454.
13. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227:680–685.
14. Kidd D, Liu Y, Cravatt BF: Profiling serine hydrolase activities in complex proteomes. *Biochemistry* 2001;40:4005–4015.
15. Liu Y, Patricelli MP, Cravatt BF: Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* 1999;96:14694–14699.

16. Olsen C, Wagtmann N: Identification and characterization of human DPP9, a novel homologue of dipeptidyl peptidase IV. *Gene* 2002;299:185–193.
17. Corthals G, Wasinger V, Hochstrasser D, Sanchez J: The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 2000;21:1104–1125.
18. Bernard A-M, Mattei M-G, Pierres M, Marguet D: Structure of the mouse dipeptidyl peptidase IV (CD26) gene. *Biochemistry* 1994;33:15204–15214.
19. Abbott C, Yu D, Woolatt E, Sutherland G, McCaughan G, Gorrell M: Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8. *Eur J Biochem* 2000;267:6140–6150.
20. Pauly RP, Demuth H-U, Rosche F, Schmidt J, White HA, Lynn F, McIntosh CH, Pederson RA: Improved glucose tolerance in rats treated with the dipeptidyl peptidase IV (CD26) inhibitor Ile-thiazolidide. *Metabolism* 1999;48:385–389.
21. Fersht A: *Enzyme Structure and Mechanism*. W.H. Freeman & Co., New York, 1984.
22. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;20:3551–3567.

Address reprint requests to:

*Tyzoon K. Nomanbhoy*

*ActivX Biosciences, Inc.*

*11025 N. Torrey Pines Road, #120*

*La Jolla, CA 92037*

*E-mail: tyzoonn@activx.com*

## **AU1**

**Is this a trademark?**

## **AU2**

**Is FP as meant?**

## **AU3**

**Where does corresponding asterisk belong in text?**