

High-resolution functional proteomics by active-site peptide profiling

Eric S. Okerberg, Jiangyue Wu, Baohong Zhang, Babak Samii, Kelly Blackford, David T. Winn, Kevin R. Shreder, Jonathan J. Burbaum, and Matthew P. Patricelli*

ActivX Biosciences, 11025 North Torrey Pines Road, Suite 120, La Jolla, CA 92037

Communicated by JoAnne Stubbe, Massachusetts Institute of Technology, Cambridge, MA, February 15, 2005 (received for review August 17, 2004)

Characterization and functional annotation of the large number of proteins predicted from genome sequencing projects poses a major scientific challenge. Whereas several proteomics techniques have been developed to quantify the abundance of proteins, these methods provide little information regarding protein function. Here, we present a gel-free platform that permits ultrasensitive, quantitative, and high-resolution analyses of protein activities in proteomes, including highly problematic samples such as undiluted plasma. We demonstrate the value of this platform for the discovery of both disease-related enzyme activities and specific inhibitors that target these proteins.

capillary electrophoresis | fluorescence | MS | protease

The completion of several major genome sequencing projects has both accelerated the pace and broadened the scale of modern biology. For example, it is estimated that a complete molecular understanding of the human organism will require the characterization of at least 100,000 proteins (1). Several innovations in the field of proteomics have advanced to meet the demands of postgenome biology. Particularly, improvements in the resolving capacity and reproducibility of 2D gel electrophoresis (2) and the increased sensitivity and accuracy of modern MS methods (3) have enabled the global characterization of protein expression levels in complex biological systems. Nonetheless, the large dynamic range of protein expression in biological tissues and fluids renders these methods ineffective at detecting and/or quantitating low-abundance proteins in the absence of labor-intensive prefractionation or enrichment procedures (2, 4). Moreover, abundance-based proteomic strategies fail to provide direct information regarding protein function or activity.

To address these limitations, a chemical proteomic technology called activity-based protein profiling (ABPP) (5) has emerged that utilizes active site-directed probes to enable the parallel measurement of the activity of many enzymes in complex proteomic mixtures (5–8). Activity-based probes (ABPs) enable the detection of changes in enzyme activity independent of alterations in protein abundance, and provide a specific enrichment strategy for low-copy-number enzymes without interference from more abundant proteins. To date, ABPP has relied on gel-based methods for proteome analysis that are difficult to automate and often fail to resolve highly related protein species (5, 6, 9, 10). Here, we describe a gel-free platform for ABPP, termed Xsite, that permits the rapid and systematic quantification and identification of enzyme activities in complex protein mixtures.

Materials and Methods

General Materials and Methods. Fluorophosphonate probes were synthesized as described (5). Synthesis of the cathepsin probe, AX6429, is detailed in supplemental materials (Fig. 3, which is published as supporting information on the PNAS web site). Rat FAAH was expressed in *Escherichia coli* and purified as described (11). Porcine trypsin and equine butyrylcholinesterase were purchased from Sigma. Urokinase plasminogen activator,

tissue plasminogen activator, prostate-specific antigen, and bovine trypsin were purchased from Calbiochem.

Preparation and Digestion of Proteomes. Generally, proteomes were labeled with ABPs at concentrations of 1–10 μ M for 5–60 min. After labeling, proteomes were denatured by adding 1 volume of 12 M urea (prepared at 50°C to allow for complete dissolution of the urea) and 1/50th volume of DTT from a fresh 1 M stock. The mixture was then heated to 65°C for 15 min. Iodoacetamide was added to 40 mM from a fresh 1 M solution, and the solution was incubated at 37°C for 30 min. The buffer was exchanged (10 mM ammonium bicarbonate/2 M urea) by gravity flow gel filtration to remove excess reagents and unreacted probe, and reduce the urea concentration to a suitable concentration for trypsin activity. Gel filtration of small volume samples, for CE-only analysis, was performed by using Sephadex G-25 superfine resin (Amersham Pharmacia Biosciences) in custom 96-well filter plates from Innovative Microplate (Chicopee, MA). Larger volume samples to be analyzed by MS were gel-filtered using Econo-Pac 10-DG columns (Bio-Rad). After gel filtration, the samples were treated with 150 ng of trypsin (sequencing grade modified trypsin, Promega) per 10 μ l of sample for 1 h at 37°C. Digests were stopped and prepared for CE by adding 200 mM citrate and 10% Triton X-100 at 1/20th of the sample volume.

Preparation of Internal Standard Proteins. Bovine trypsin, porcine trypsin, and porcine elastase were diluted to 0.25 mg/ml in reaction buffer and labeled with 20 μ M fluorophosphonate polyethylene glycol (FP-PEG)-4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacine-3-propionic acid (BODIPY FL) for 30 min. Excess probe was removed by gel filtration on Econo-Pac 10-DG columns (Bio-Rad). The proteins were mixed together before use. The internal standard mix was added at 1/30th of the volume of the sample immediately after the quenching of the reactions with urea. Importantly, the standard proteins require proper reduction, alkylation, and tryptic digestion to give the proper migration pattern.

Capillary Electrophoresis (CE). CE was performed by using a Beckman PACE MDQ CE system. Excitation was provided by a 100 mW, frequency-doubled (532 nm) neodymium:yttrium-aluminum-garnet laser (Coherent). Fluorescence was collected into two channels, one for the carboxytetramethylrhodamine (TAMRA) signal and one for the BODIPY FL signal. The TAMRA channel contained a 532-nm notch filter, followed by a 580-nm bandpass filter (Omega Optical). The BODIPY FL

Freely available online through the PNAS open access option.

Abbreviations: ABP, activity-based probe; CE, capillary electrophoresis; LIF, laser-induced fluorescence detection; LC, liquid chromatography; PF-PEG, fluorophosphonate polyethylene glycol; BODIPY FL, 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacine-3-propionic acid; TAMRA, carboxytetramethylrhodamine.

*To whom correspondence should be addressed. E-mail: mattp@activx.com.

© 2005 by The National Academy of Sciences of the USA

channel contained a 520-nm short-pass filter, followed by a 510-nm bandpass filter (Omega Optical).

Samples were separated in 100 μM \times 60 cm eCAP DNA capillaries (Beckman Coulter). Coated capillaries demonstrated the best resolution, reproducibility, and lifetime characteristics, but small amounts of detergent were required to prevent adsorption of labeled peptides to the capillary walls. The running buffer contained 50 mM aspartic acid, 10 mM Hepes, and 0.2% Triton X-100 (pH 3.1 without adjustment). Acidic buffer was chosen to ensure that all tryptic peptides would have a net positive charge and migrate toward the cathode. Before each run, capillaries were rinsed at 20 psi for 1.5 min. Samples were injected for 6 s at 0.1 psi (20-nl estimated injection volume) then separated by applying 25 kV (inlet = anode) for 45 min. Data from TAMRA and BODIPY FL were collected in separate fluorescence channels during CE analysis, and the BODIPY FL standard traces were used to normalize both relative migration times and peak intensities in the sample (TAMRA) channel. Typical separation efficiencies ranged from 200,000 to 500,000 theoretical plates.

Affinity Capture of ABP-Labeled Peptides. After digestion, peptides were incubated with agarose beads containing a monoclonal anti-TAMRA antibody (5 mg of antibody per milliliter of beads) for 1 h. The beads were washed three times with buffer, and three times with deionized water before eluting with two volumes of 50% acetonitrile/0.1% trifluoroacetic acid (TFA). Samples were then dried under vacuum and resuspended in 10% acetonitrile/0.1% TFA for liquid chromatography (LC)-MS/MS analysis. The efficiency and selectivity of this method were determined (see *Supporting Text*, which is published as supporting information on the PNAS web site).

LC-MS/MS Analysis. Peptides were separated by using 300 μM ID C18 reversed-phase silica columns (Agilent Technologies, Palo Alto, CA) interfaced with a Finnigan LCQ Deca LC-MS/MS system with a nanospray source. Data were collected in data-dependent mode with dynamic exclusion.

To correlate identified peptides to CE peaks, a flow splitter was incorporated just after the LC column, which sent $\approx 50\%$ of the eluent through a line that passed through a fluorescence detector (Picometrics, Ramonville, France) and terminated in a 96-well fraction collector (Dionex, Sunnyvale, CA). The time delay between the fraction collector and the MS data were determined by overlaying the fluorescence and ion intensity profiles from the run. The LC migration times of peptides of interest was then determined, and the appropriate fraction was analyzed by CE to determine its migration position. The internal standard peptides described above were included in all fractions (See *Supporting Text* for additional details).

Identification of Probe-Labeled Peptides. LC-MS/MS data were searched against protein databases by using the SEQUEST algorithm. Several modifications were introduced to facilitate the interpretation of the SEQUEST results (see *Supporting Text*). Before searching, known probe fragment ions were removed from the MS/MS spectra. This procedure was found to significantly improve the search results. The results yielded by the SEQUEST program were then filtered to remove any results that corresponded to unmodified peptides, or peptides with more than one probe modification. These adjustments resulted in an increased number of relatively high-scoring peptides corresponding to serine hydrolase active-site-labeled fragments. However, many other peptides, which were most likely false-positive results, were still found with comparable scores.

Supporting Information. Detailed method information is contained in *Supporting Text*.

Results

CE Screening Method. We envisioned that the analysis of ABP-labeled peptides from digested proteomic samples, rather than intact proteins, would allow for a high-resolution, protein size-independent separation, overcoming a major limitation of SDS/PAGE-based ABP analysis (Fig. 1A). The strength of this approach derives from the fact that ABPs, in principle, react selectively with a single active-site amino acid on each target enzyme. Thus, the number of ABP-labeled peptide species in an enzymatic digest of an ABP-labeled proteome should equal the number of labeled proteins present before digestion. Furthermore, because ABPs react covalently, the mass spectral characterization of probe-labeled peptides from complex mixtures should not only reveal the identity of the labeled protein but also the precise site of labeling. This information is invaluable for assessing the selectivity and mechanism of probe reactivity.

The success of the Xsite platform requires highly reproducible enzymatic digests and a robust, sensitive, and high-resolution separation strategy. To allow profiling of enzyme activities across a broad range of proteomes, both the enzymatic digest protocol and separation strategy must produce quantitative, reproducible data, regardless of the nature of the background proteome. To achieve these goals, a sample preparation method was developed that includes disulfide reduction, thiol alkylation, and gel filtration steps before enzymatic digestion to prevent oxidation of free thiol groups and to remove unreacted probe (and other potential interfering small molecules). CE with laser-induced fluorescence detection (CE-LIF) was chosen for the analysis of probe-labeled peptides because of its high-sensitivity, high-resolution, and potential for high-throughput separations. A CE-LIF method operating at low-run buffer pH in coated capillaries afforded high-resolution, reproducible separations by minimizing electroosmotic flow and adsorption of peptides to the capillary surface. A set of probe-labeled internal standard proteins was generated, by using an alternate fluorophore, to enable normalization of both the sample preparation and the separation processes. Importantly, all steps in this procedure, including the gravity-flow gel filtration steps, are performed in 96-well plates from an initial sample volume of 10 μl with standard liquid handling robotics.

To validate the general utility of this screening method for proteomic analysis, eight purified serine hydrolases were pre-labeled with a serine hydrolase directed ABP (FP-PEG TAMRA; ref 10), added to eight different proteomes, processed according to our standardized protocol, and analyzed by CE-LIF. The eight proteomes included two soluble proteomes, two membrane proteomes, two secreted proteomes, whole yeast extract, and undiluted human plasma. The protein concentration of these proteomes ranged from ≈ 0.5 mg/ml for the secreted proteomes to 50 mg/ml for the undiluted plasma. Migration times and peak heights for all eight proteins were remarkably uniform among all proteomes tested (Fig. 1B and Tables 3 and 4, which are published as supporting information on the PNAS web site). Peak migration time CV's ranged from 0.02% to 0.1%, whereas peak height and area reproducibility were between 15% and 30%. The extremely low variability in migration time and the high separation efficiencies are among the best reported for peptide analysis by CE (12), and are particularly noteworthy, considering the presence of problematic background proteomes such as undiluted plasma.

To assess the sensitivity of this analysis method, known amounts of a purified serine hydrolase (butyrylcholinesterase) were added to mouse heart proteome before labeling with 10 μM FP-PEG-TAMRA (Fig. 1C). At 190 pM, in 5 mg/ml background proteome, the active-site peptide from butyrylcholinesterase exhibited a signal-to-noise ratio of 40, corresponding to an absolute detection limit of 230 molecules of active protein per

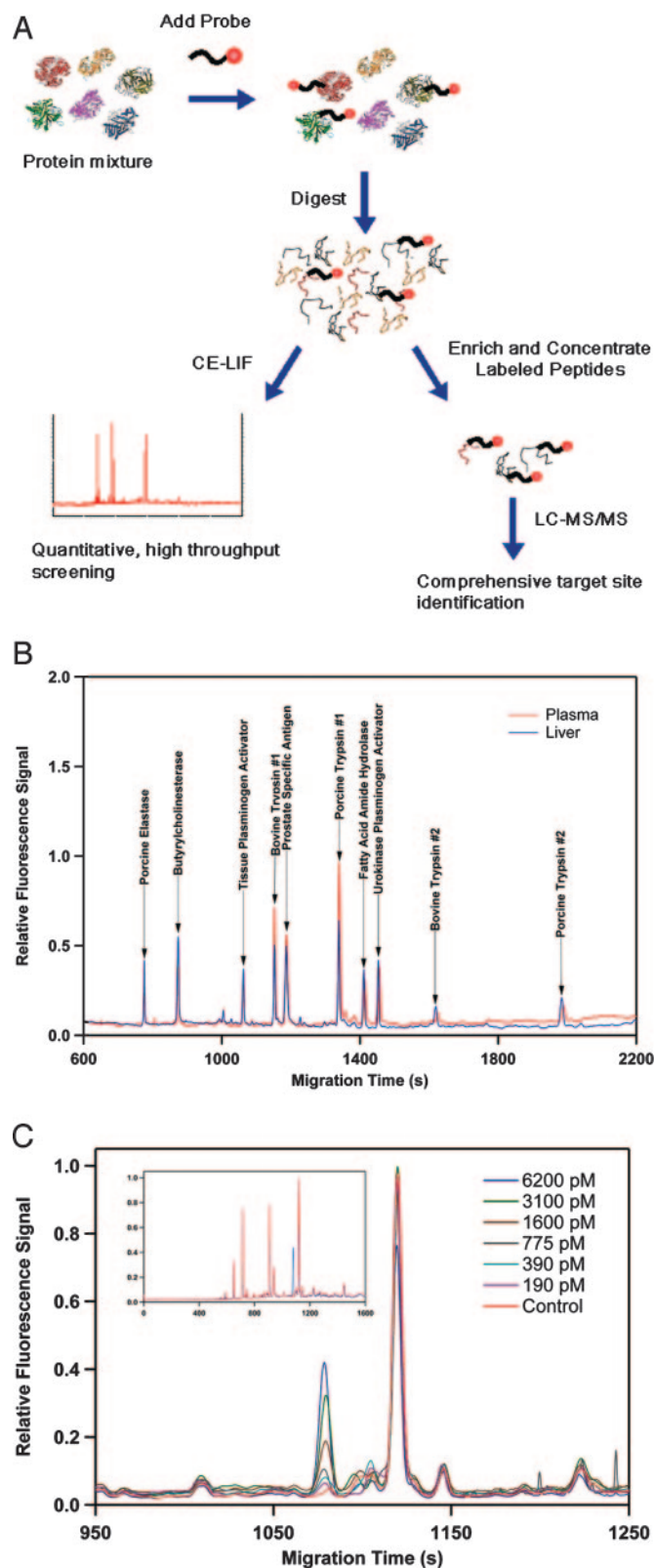


Fig. 1. Validation of the Xsite method. (A) A diagrammatic view of the Xsite platform. Proteomic samples are first reacted with an ABP, followed by an enzymatic digestion step to yield a mixture of peptides, some of which are ABP-modified. The ABP-modified peptides can then either be rapidly and quantitatively profiled without further sample handling using CE-LIF, or purified by affinity chromatography and analyzed by LC-MS/MS to determine the identities of the labeled proteins. (B) Eight serine hydrolases pre-labeled with the FP-PEG TAMRA probe were added to either mouse liver proteome

cell (with a signal to noise ratio of 4) by using only 5×10^5 cells. It is important to consider that this analysis consumed ≈ 20 nl of sample, yielding an absolute detection limit of ≈ 100 zmol (1×10^{-19} mol) for the CE-LIF instrument.

Active-Site Peptide Identification. The CE-based screening method detailed above provides rapid, extremely sensitive, and quantitative analyses of ABP-labeled samples. However, the ability to assign the molecular identities for each peak observed during a CE run is critical to the success of this platform. To design a broadly applicable method for molecular annotation of CE traces, several factors were considered. First, the method for preparing samples for MS should not significantly alter the relative levels of ABP-labeled species. Second, because ABPs generally react with relatively scarce proteins, an enrichment method must be implemented to concentrate and purify the ABP-labeled peptides from unlabeled material. Finally, to obtain confident protein identifications, a single ABP-labeled peptide must be sufficient to determine the parent protein's identity.

To identify the labeled species in a given CE run, samples were processed and digested according to the protocol described for the CE-screening method, except that larger sample volumes (0.2–1.0 ml) were typically used. After digestion, ABP-labeled peptides were enriched and concentrated by using a monoclonal anti-TAMRA antibody immobilized on agarose beads. The enriched, labeled peptides were eluted from the affinity capture beads and analyzed by LC-MS/MS. The overall yield of this peptide enrichment protocol was determined by CE-LIF to be near 100%, and no significant unlabeled material was detected by LC-MS/MS analysis, even when a single labeled species was present in a proteomic mixture at only 0.0002% (Fig. 4, which is published as supporting information on the PNAS web site).

Three FP-PEG-TAMRA-labeled mouse proteomes (liver, submaxillary, and pancreas) were enriched by this “peptide-capture” method and analyzed by nano-electrospray LC-MS/MS. The expected increase in mass due to probe modification was used as a variable modification on the expected target residue for searching the tandem mass spectra against a non-redundant mouse database by using the SEQUEST algorithm (13). Raw data files were prefiltered to remove known probe fragments and to predict peptide charge states, and SEQUEST results were interpreted by using optimized peptide scoring methods (*Supporting Text*). The scoring procedures were designed to both increase the number of high-confidence identifications and to statistically define the level of confidence in all peptides identified (Figs. 5 and 6, which are published as supporting information on the PNAS web site). Analysis of these three proteomes resulted in the identification of 49 serine hydrolases with the probe modification site at the known or predicted active-site serine nucleophile, with $>90\%$ confidence (Table 1 and *Supporting Text*). In total, the high-confidence serine hydrolase identifications listed in the table represent ≈ 760 MS/MS spectra. In this same data set, only 30 MS/MS spectra with search results $>90\%$ confidence did not correspond to an active-site-labeled serine hydrolase (3.6%), and only 30 MS/MS spectra with search results lower than 90% confidence corresponded to active-site-labeled serine hydrolases (3.6%). This method, therefore, achieves remarkably comprehensive coverage of probe-

(1 mg/ml) or undiluted human plasma (≈ 50 mg/ml). The samples were processed and digested with trypsin under identical conditions and then analyzed by CE-LIF. All eight peptides exhibited excellent migration and peak height/area reproducibility between these very different proteomes (see also Table 3). (C) Butrylcholinesterase was added to mouse lung proteome (5 mg/ml) at the indicated concentrations to determine the sensitivity of the Xsite CE method.

Table 1. Serine hydrolases identified in mouse tissues

Protein	Reference	Tissue distribution	Molecular mass, kDa
Candidate tumor suppressor OVCA2	Q9D7E3	●	24.2
Lysophospholipase 2	Q9WTL7	●	24.8
Monoglyceride lipase	O35678	●	33.4
Kynurenine formamidase	Q8K4H1	●	34.2
Carboxylesterase ML1	Q924V7	●	41.5
Similar to platelet activating factor acetylhydrolase	Q8VDG7	●	43.5
Arylacetamide deacetylase	Q8VCF2	●	45.3
Putative carboxylesterase	Q91XD5	●	58.2
Liver carboxylesterase	ESTM_MOUSE	●	61.5
Similar to carboxylesterase 2	Q8QZR3	●	61.9
Putative carboxylesterase	Q8BK48	●	62.3
Carboxylesterase 1	Q8VCC2	●	62.7
Cholinesterase	CHLE_MOUSE	●	68.5
Prolyl-oligopeptidase (putative)	Q8BK56	●	75.9
Prolyl endopeptidase	PPCE_MOUSE	●	80.8
Dipeptidyl peptidase IV	DPP4_MOUSE	●	87.4
Tripeptidyl-peptidase II	TPP2_MOUSE	●	140.0
Lysophospholipase	Q8BWM6	●, ■	22.8
Putative lipase	Q9DB29	●, ■	28.0
Esterase 10	Q9R0P3	●, ■	31.3
Williams–Beuren syndrome critical region 21	Q8K4F5	●, ■	33.6
Platelet-activating factor acetylhydrolase	PAFA_MOUSE	●, ■	49.4
Liver carboxylesterase 22	ES22_MOUSE	●, ■	61.6
Carboxylesterase 3	Q8VCT4	●, ■	61.8
Putative carboxylesterase	Q8R097	●, ■	62.7
Bile-salt-activated lipase	BAL_MOUSE	●, ■	65.8
Similar to <i>N</i> -acylaminoacyl-peptide hydrolase	Q8R146	●, ■	81.4
Dipeptidyl peptidase VIII	Q80YA7	●, ■	10.2
Trypsin II, anionic (Pretrypsinogen II)	TRY2_MOUSE	■	26.2
Pancreatic trypsin	Q9R0T7	■	26.3
Chymotrypsin B	Q9CR35	■	27.8
Chymopasin	Q9ER05	■	28.1
Putative elastase 1	Q91X79	■	28.9
Putative elastase 3B	Q9CQ52	■	28.9
Elastase 2	EL2_MOUSE	■	28.9
Palmitoyl-protein thioesterase 2	PPT2_MOUSE	●, ▲	34.4
Liver carboxylesterase	ESTN_MOUSE	●, ■, ▲	61.1
Kallikrein K22	KLKM_MOUSE	▲	28.4
Kallikrein K26	KLKZ_MOUSE	▲	28.5
Kallikrein K8	KLK8_MOUSE	▲	28.5
Kallikrein K13	KLKD_MOUSE	▲	28.7
Gamma-renin	KLKG_MOUSE	▲	28.7
Kallikrein K11	KLKB_MOUSE	▲	28.7
Kallikrein 27	Q9JM71	▲	28.7
Kallikrein K6	KLK6_MOUSE	▲	28.8
Kallikrein 21	Q9JM70	▲	28.7
Kallikrein K3	KLK3_MOUSE	▲	29.0
Kallikrein K9	KLK9_MOUSE	▲	28.9
Kallikrein K1	KLK1_MOUSE	▲	29.0

●, mouse liver; ■, mouse pancreas; ▲, mouse submaxillary.

labeled proteins with a high degree of accuracy in both the identity of the labeled protein and the site of probe labeling.

Proteomic Profiling. The ABP-labeled liver, pancreas, and submaxillary samples contained remarkably distinct serine hydrolase profiles based on the above LC-MS/MS analysis (Table 1). In fact, only one serine hydrolase, liver carboxylesterase, was confidently identified in all three tissues. However, by using fluorescence-based SDS/PAGE analysis, the same three proteomic samples appeared substantially similar, with a dominant probe-reactive protein band at ≈ 25 kDa (Fig. 2A). Considering

the molecular masses of the proteins identified in these tissues, it is readily apparent that the major 25-kDa bands on the gel correspond to a variety of serine proteases. The submaxillary sample contained high levels of 11 different members of the kallikrein family, whereas the pancreas contained high levels of trypsin and several other trypsin-related proteases, none of which are members of the kallikrein family. By using the CE-LIF-based screening platform, these three proteomes appeared very distinct, which is consistent with the diversity of proteins identified by the LC-MS/MS analysis (Fig. 2B).

A similar improvement in sensitivity and resolution was

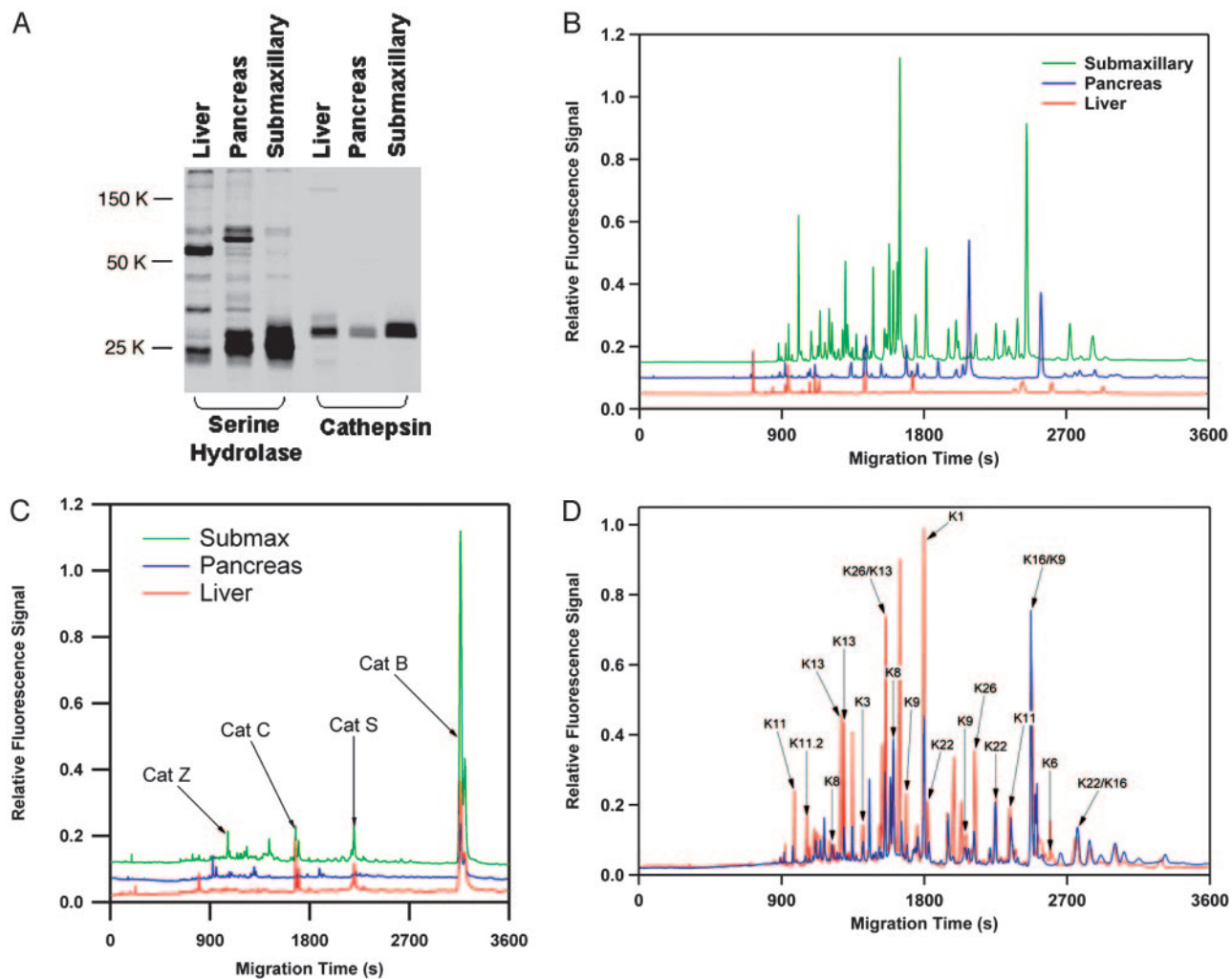


Fig. 2. Proteomic profiling of serine and cysteine proteases with the Xsite method. (A) SDS/PAGE analysis of mouse liver, pancreas, and submaxillary proteomes after labeling with the serine hydrolase probe FP-PEG-TAMRA and the cathepsin probe AX6429. Analysis of the same three proteomes by using the CE-LIF method after digestion of the proteomes with trypsin shows greatly increased resolution for both serine hydrolases (B), and cathepsins (C). Despite the similarity between serine hydrolase profile of the pancreas and submaxillary by using SDS/PAGE, the CE-LIF method suggests these proteomes are highly distinct. Additionally, the CE-LIF separation suggests the presence of many more proteins than the number of bands on the SDS/PAGE gels. This result is because of the presence of multiple labeled species in these proteomes with common molecular weights. (D) The submaxillary proteome was analyzed by LC-MS/MS, and the identities of the CE peaks were determined (red trace, annotations). Several members of the kallikrein group of serine proteases were found. Addition of the protease inhibitor nafamostat (blue trace) indicated that this compound selectively inhibited a subset of these enzymes.

achieved when these proteomes were analyzed with a probe directed against the cathepsin family of cysteine proteases (see Fig. 3 for probe structure). SDS/PAGE analysis indicated a crowded region of labeling containing several poorly resolved bands near 30 kDa. MS analysis of the submaxillary proteome revealed the presence of four cathepsins (B, C, S, and Z). By using the CE-LIF method, the probe-labeled active-site peptides from these cathepsins were nearly equally spaced over a separation range of 45 min. Significant differences in relative expression levels were observed between the three tissues, with only the submaxillary sample containing significant levels of active cathepsin Z. Previous work with cathepsin ABPs (14) has required the use of 2D electrophoresis to achieve complete resolution of these highly homologous enzymes.

Drug Selectivity Analysis. The ability of the CE-LIF method to distinguish the various kallikrein enzymes present in the submaxillary sample was confirmed by determining the molecular identity of the observed CE peaks. To correlate the identified

proteins with the peaks observed in the CE trace, fractions were collected during a LC-MS/MS run from a split flow line fitted with a fluorescence detector and fraction collector. The collected fractions were then analyzed by CE to determine the migration positions of the components present (Figs. 2D and 7, which is published as supporting information on the PNAS web site). This analysis confirmed that the CE-LIF separation had successfully separated the majority of the kallikrein enzymes present in this sample, despite primary sequence identity between these enzymes of up to 96%.

To demonstrate the power of the Xsite platform for rapid and quantitative profiling of highly related enzyme families, the selectivity of a clinically relevant serine protease inhibitor, nafamostat (Fig. 8, which is published as supporting information on the PNAS web site), was assessed directly in the mouse submaxillary proteome. nafamostat is approved for clinical use in Japan for diseases where serine hydrolases are activated, including disseminated intravascular coagulation and acute pancreatitis. Its ability to inhibit kallikreins has been suggested to be

Table 2. IC₅₀ Values for nafamostat inhibition of mouse kallikreins

Kallikrein	IC ₅₀ , μM
K13	5
K22	9
K11	11
K9	27
K26	31
K1	34
K3	51
K16	280
K8	>500

the source for nafamostat's prevention of injection pain (15). Several targets of nafamostat were apparent in soluble protein fractions from the submaxillary gland by using the CE-LIF screening method (Fig. 2D, compare red and blue traces). IC₅₀ values were determined for the various kallikreins by quantitating the peak heights corresponding to each enzyme within a range of nafamostat concentrations (Table 2). Importantly, these data were generated very rapidly by using the CE screening platform, because MS-based analysis was no longer necessary. Samples were prepared for CE in 96-well plates by using standard robotics in ≈4 h.

The data reveal that nafamostat exhibited unpredicted selectivity within the set of kallikreins examined (Table 2). Interestingly, most of the kallikreins examined contain an aspartic acid at the P1 site, priming them for binding substrates with a positively charged S1 moiety. nafamostat contains two symmetrically arranged positively charged guanidine groups that would be expected to interact with the P1-binding pocket of the enzyme targets. Two kallikreins in the set examined (K8 and K16) do not contain an aspartic acid at the P1 site. Very weak inhibition of K16 was observed (IC₅₀ = 280 μM), whereas K8 was not significantly inhibited at concentrations as high as 500 μM. The

potencies of inhibition for kallikreins with aspartic acid at P1 ranged by approximately one order of magnitude, indicating additional factors contribute to the potency of the interaction between nafamostat and the kallikreins examined since. Overall, these data suggest the possibility that selective kallikrein inhibitors could be developed by using this screening method.

The tissue kallikrein gene family is a largely uncharacterized group of serine proteases whose members include validated tumor markers such as PSA (16) and several recently discovered members with largely unknown functions (17). Tissue kallikreins have been implicated in many disease related biological processes, including tumorigenesis (18, 19), angiogenesis (19), and inflammation (20). As a result, the diagnostic and therapeutic potential of kallikreins is currently the subject of much investigation. These efforts have been complicated by the close sequence homology among members, overlapping tissue distributions, and overlapping substrate selectivities. Further, the correlation between kallikrein expression and activity is complicated by the presence of at least one endogenous inhibitor (19, 21, 22). The Xsite method overcomes several obstacles in the characterization of both the disease-related expression patterns of active kallikreins and the development of selective kallikrein inhibitors.

Discussion

The Xsite platform enables rapid functional analyses of enzyme activities in crude protein mixtures and proteomes. The sensitivity and resolution of the analysis methods greatly exceed those attainable by other approaches and enable the facile analysis of samples that were previously intractable through other known methods. The platform is extremely well suited for the analysis of inhibitor selectivity and potency and provides the ability to profile clinical samples for the discovery of disease-related enzyme activities.

We thank John Kozarich, Ben Cravatt, Jonathan Rosenblum, and Katrin Szardenings for helpful suggestions with the manuscript and for abundant technical advice.

- Harrison, P. M., Kumar, A., Lang, N., Snyder, M. & Gerstein, M. (2002) *Nucleic Acids Res.* **30**, 1083–1090.
- Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F. & Sanchez, J. C. (2000) *Electrophoresis* **21**, 1104–1115.
- Aebersold, R. & Mann, M. (2003) *Nature* **422**, 198–207.
- Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y. & Aebersold, R. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 9390–9395.
- Liu, Y., Patricelli, M. P. & Cravatt, B. F. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 14694–14699.
- Bogyo, M., Verhelst, S., Bellingard-Dubouchaud, V., Toba, S. & Greenbaum, D. (2000) *Chem. Biol.* **7**, 27–38.
- Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K. F. & Bogyo, M. (2002) *Mol. Cell. Proteomics* **1**, 60–68.
- Adam, G. C., Sorensen, E. J. & Cravatt, B. F. (2002) *Mol. Cell. Proteomics* **1**, 781–790.
- Greenbaum, D. C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K. F., Engel, J., DeRisi, J., Holder, A. A. & Bogyo, M. (2002) *Science* **298**, 2002–2006.
- Patricelli, M. P., Giang, D. K., Stamp, L. M. & Burbaum, J. J. (2001) *Proteomics* **1**, 1067–1071.
- Patricelli, M. P., Lashuel, H. A., Giang, D. K., Kelly, J. W. & Cravatt, B. F. (1998) *Biochemistry* **37**, 15177–15187.
- Hutterer, K. M. & Jorgenson, J. W. (1999) *Anal. Chem.* **71**, 1293–1297.
- Yates, J. R., III, Eng, J. K., McCormack, A. L. & Schieltz, D. (1995) *Anal. Chem.* **67**, 1426–1436.
- Greenbaum, D., Medzihradszky, K. F., Burlingame, A. & Bogyo, M. (2000) *Chem. Biol.* **7**, 569–581.
- Iwama, H., Nakane, M., Ohmori, S., Kaneko, T., Kato, M., Watanabe, K. & Okuaki, A. (1998) *Br. J. Anaesth.* **81**, 963–964.
- Balk, S. P., Ko, Y. J. & Bublely, G. J. (2003) *J. Clin. Oncol.* **21**, 383–391.
- Yousef, G. M. & Diamandis, E. P. (2001) *Endocr. Rev.* **22**, 184–204.
- Diamandis, E. P. & Yousef, G. M. (2002) *Clin. Chem.* **48**, 1198–1205.
- Miao, R. Q., Agata, J., Chao, L. & Chao, J. (2002) *Blood* **100**, 3245–3252.
- Colman, R. W. (1999) *Immunopharmacology* **43**, 103–108.
- Chao, J., Miao, R. Q., Chen, V., Chen, L. M. & Chao, L. (2001) *Biol. Chem.* **382**, 15–21.
- Chai, K. X., Chen, L. M., Chao, J. & Chao, L. (1993) *J. Biol. Chem.* **268**, 24498–24505.