Direct Imaging and Identification of Proteoforms up to 70 kDa from Human Tissue

Authors

Pei Su,[1] John P. McGee,[1] Kenneth R. Durbin,[1] Michael A. R. Hollas,[1] Manxi Yang,[2] Elizabeth K. Neumann,[3] Jamie L. Allen,[3] Bryon S. Drown,[1] Fatma Ayaloglu Butun,[4] Joseph B. Greer,[1] Bryan P. Early,[1] Ryan T. Fellers,[1] Jeffrey M. Spraggins,[3,5] Julia Laskin,[2] Jeannie M. Camarillo,[1,4] Jared O. Kafader,[1,4] and Neil L. Kelleher[1,4,6]*

Affiliations

[1]Departments of Molecular Biosciences, Chemistry, Chemical and Biological Engineering, and the Feinberg School of Medicine, Northwestern University, Evanston, IL, USA

[2]Department of Chemistry, Purdue University, West Lafayette, IN, USA

[3]Department of Biochemistry and Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA

[4]Proteomics Center of Excellence, Evanston, IL, USA

[5]Department of Chemistry, Cell and Development Biology, Vanderbilt University, Nashville, TN, USA

[6]Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

*Correspondence to: Neil L Kelleher, n-kelleher@northwestern.edu 2145 Sheridan Rd, Evanston, IL 60208

Abstract

Imaging of proteoforms in human tissues is hindered by low molecular specificity and limited proteome coverage. Here, we introduce proteoform imaging mass spectrometry (PiMS), which increases the size limit for proteoform detection and identification by 4-fold compared to reported methods, and reveals tissue localization of proteoforms at <80 µm spatial resolution. PiMS advances proteoform imaging by combining liquid sampling (nanospray desorption electrospray ionization, nano-DESI) with ion detection using individual ion mass spectrometry (I^2MS). We demonstrate the first proteoform imaging of human kidney, identifying 169 of 400 proteoforms <70 kDa using top-down mass spectrometry and database lookup from the human proteoform atlas, including dozens of key enzymes in primary metabolism. Moreover, PiMS images visualize kidney anatomical structures and cellular neighborhoods in the vasculature versus the medulla or the cortex. The benefits of PiMS are poised to increase proteome coverage for label-free protein imaging of intact tissues.

Teaser

Nano-DESI combined with individual ion mass spectrometry generates images of proteoforms up to 70 kDa.

Introduction

Proteoforms are the protein-level products of gene expression and post-translational modifications functioning as key effectors in human health and disease (*1, 2*). In addition to understanding of their molecular compositions, interactions, and biological function, comprehensive characterization of the human proteoform landscape also requires mapping of their spatial distributions in human tissues and organs (*3*). Protein-level imaging using antibody-based optical microscopy has revealed distinct cell types, functional tissue units, and subcellular structures (*4*). These techniques employ enzymes and fluorophores as reporters to obtain high-resolution maps of protein targets in tissues (*5*). In recent years, highly multiplexed antibody-based imaging assays such as CODEX (*6, 7*), IBEX (*8*), and Cell-DIVE (*9*) has drastically increased the number of protein targets that can be probed in a single experiment. Alternatively, mass spectrometry (MS)-based imaging assays, including imaging mass cytometry (IMC) and multiplexed ion beam imaging (MIBI), utilize antibodies labeled with rare earth metals to detect protein localization for up to ~60 protein targets at once (*10, 11*). Despite the significant advances in spatial resolution and sensitivity, antibody-based approaches require prior knowledge of the protein targets and lack full molecular specificity like that provided by proteoform-level information (*12, 13*).

MS-based top-down proteomics (*14, 15*) has been widely used for proteoform characterizations (*16*). Modern MS instrumentation has reached the sensitivity for spatially-resolved top-down proteomics suitable for imaging experiments (*17, 18*). Matrix-assisted laser desorption/ionization (MALDI) is widely used for protein imaging (*19, 20*) due to the broad mass range to the sampling of the proteome (*21-23*). However, MALDI predominantly generates singly-charged ions, which gives limited fragment information for direct top-down identification of intact proteins (*24*). This challenge may be addressed using matrix-assisted laser desorption electrospray ionization (MALDESI), which combines MALDI with extractive ESI to generate

multiply-charged ions of peptides and proteins extracted from tissues (*25*). Alternatively, multiply-charged protein ions may be generated using liquid extraction-based ambient ionization methods (*26*) including desorption electrospray ionization (DESI) (*27*), liquid extraction surface analysis (LESA) (*28*), and nanospray desorption electrospray ionization (nano-DESI) (*29*). These techniques are particularly advantageous in top-down analysis of intact proteoforms in the imaging mode. Among these techniques, nano-DESI that utilizes a sub-nanoliter dynamic liquid bridge as a sampling probe enables imaging of biomolecules in tissues with a spatial resolution down to 10 µm (*30, 31*).

One major challenge in proteoform imaging using liquid extraction-based techniques is the detection of low-abundance, high mass proteoforms in the congested MS spectra produced by ionizing complex mixtures of biomolecules extracted from the sample. Until now, imaging and identification of intact proteoforms directly from tissue has been limited to <20 kDa species (*27, 28, 32*), with one report leading to the identification of subunits from a 43 kDa trimeric protein complex (*33*). Here, we address this challenge using individual ion mass spectrometry (I^2MS). I^2MS is a new, orbitrap-based charge detection technique (*34-36*) for single ion detection and results in a 500-fold improved sensitivity with a 10-fold higher resolving power with a high dynamic range (*37*). In particular, we combine nano-DESI ionization with I^2MS (*37*) to create proteoform imaging mass spectrometry (PiMS) for tissue imaging and direct identification of proteoforms up to ~70 kDa. We show isotopically-resolved proteoform assignment from human kidney and confidently identify proteoforms up to 53 kDa, illuminating differences in kidney architecture from the medulla, cortex, and vasculature. Incorporating I^2MS, PiMS yielded 169 proteoform assignments/identifications at 80 µm spatial resolution and demonstrates the potential to illuminate the proteinaceous structures comprising our tissues.

Results

Overview of PiMS workflow

Proteoform imaging mass spectrometry (PiMS) illustrated in Fig. 1 combines nanospray desorption electrospray ionization (nano-DESI) imaging with data acquisition and processing for individual ions. Specifically, we perform nano-DESI line scans on tissue, during which proteoforms were sampled as multiply-charged ions distributed across multiple charge states (Fig. 1, top left). Instead of unresolved protein signals typically observed for ensembles of ions, we obtain charge-assigned individual protein ions from each pixel on the tissue section to allow for detection of proteoforms with resolution of their ^{13}C isotopic peaks (Fig. S1). This allows for confident assignment of proteoform mass to <2 parts-per-million at one sigma in each pixel of the imaging data (Fig. 1, middle left). Beyond proteoform-specific images (Fig. 1, bottom left), molecular identification is achieved using either direct top-down MS off the tissue (Fig. 1, top right) or database searching from known proteoforms at the intact masses (Fig. 1, bottom right).

Human Kidney Proteoforms Detected by PiMS

We used PiMS to examine the proteoforms and their localizations in a 10 µm-thick human kidney tissue section. Encouragingly, we immediately expanded the mass detection range for proteoform imaging to >70 kDa. Fig. 2A shows the full PiMS spectrum from 5-72 kDa from a sum of 16500 MS scans (~8 million single ions). This spectrum contains ~400 proteoform masses above 0.1% relative abundance that are isotopically resolved. A complex group of proteoforms in the 68-80 kDa range were observed and not individually resolved (fig. S2). Spectral attributes include a dynamic range of ~200 (using S/N 3 as the limit of detection) and a mass resolution ($m/\Delta m$) of ~100,000. PiMS images can be constructed for any of the 242 proteoform masses with relative abundance above 1% in the full PiMS spectrum of Fig. 2A.

To identify these proteoforms, we manually annotated the full PiMS spectrum shown in Fig. 2A based off of the results from an intact mass tag (IMT) search against a custom database. In this approach, we compared the shape and mass accuracy of the isotopic distributions of the proteoforms in the PiMS spectrum with the theoretical proteoforms in the database. The database was constructed from the top 500 most abundant proteins identified in a bottom-up proteomics study of human kidney tissues (*38*). To maximize the number of proteoform identification, we combined additional matches from top-down identification (details discussed in the next section) and manually inspected post-translational modifications (PTMs) recorded in the Swiss-Prot database (*38*). As a result, we manually annotated 169 proteoforms in the entire mass range using a ±5 ppm mass tolerance (list of proteoforms shown in table S1). Fig. 2D and 2E show two zoomed regions of the full PiMS spectrum with theoretical proteoform matches highlighted in color. Fig. 2D shows various unmodified proteoforms in the mass range of 22.2-22.7 kDa captured by the custom database, demonstrating that the search included a variety of proteins in the kidney proteome. The mass range of 41.4-42.1 kDa shown in Fig. 2E contains proteoforms of actin, cytoplasmic 2 (highlighted in pink) and 3-ketoacyl-CoA thiolase (highlighted in blue) with a few types of PTMs and their combinations. Aside from the monoacetylated + dimethylated proteoform of actin, cytoplasmic 2 identified by top-down MS, other modified proteoforms in the displayed mass range were manually annotated. Clearly, further investigation is needed to estimate the false discovery rate for automated identification in PiMS, including the use of Bayesian priors.

We annotated the most abundant proteins in Fig. 2A and Fig. 2B to demonstrate the portion of the human kidney proteome captured by PiMS. Not surprisingly, blood proteins (hemoglobin subunits, apolipoprotein A-1, and albumin, Fig. 2A) were found at highest abundance due to the highly vascularized nature of the kidney. Meanwhile, we captured many proteins prevalent in cellular pathways that are naturally abundant in human cells. In Fig. 2B, we

labeled the most abundant proteins to give a brief overview of the molecular functions and biological pathways observed in PiMS. From low to high mass range shown in Fig. 2B, we found molecular chaperone (alpha-crystallin B chain, CRYAB), signaling modulator (phosphatidylethanolamine-binding protein 1, PEBP1), proteins for cellular detoxification (superoxide dismutase [Mn], mitochondrial, SOD2, glutathione S-transferase A1&A2, GSTA1&GSTA2) and homeostasis (carbonic anhydrase 1&2), and structural proteins (actins, vimentin). Moreover, proteins participating in central metabolic pathways are dominant. In particular, we found key enzymes in Krebs cycle (malate dehydrogenase, MDH) and gluconeogenesis (fructose-1,6-bisphosphatase 1, FBP1), and 27 subunits of protein complexes in the electron transport chain of oxidative phosphorylation (blue asterisks for 6 subunits in 18-56 kDa mass range, others recorded in table S1). More intriguingly, four key enzymes in glycolysis (triosephosphate isomerase, TPI, glyceraldehyde-3-phosphate dehydrogenase, GAPDH, phosphoglycerate kinase, PGK, and alpha-enolase, ENO) were captured (Fig. 2C, red asterisks in Fig. 2A and 2B). PiMS images of the four detected glycolytic enzymes were shown in Fig. 2C with largely even distributions across the entire tissue section spanning from the kidney medulla to the cortex, confirming the presence of glycolysis in nearly all kidney cell types. A more complete investigation of the biological pathways found in the 169 identified proteoforms were demonstrated by a Gene Ontology (GO) analysis shown in Fig. 2F, showing cellular metabolic processes as the predominant biological pathway observed in the PiMS experiment.

Top-Down Characterizations of Kidney Proteoforms in PiMS

Human kidney proteoforms were further identified and characterized using tandem MS (MS/MS). Direct fragmentation of protein ions >20 kDa is challenging due to the low abundance of their fragment ions, particularly larger ones (>15 kDa) that result from cleavage in the middle of the protein sequence (*39*). To overcome this limitation, we employed I^2MS for the readout of

top-down fragmentation spectra to capture those large fragment ions typically buried under the noise level in ensemble MS/MS experiments (*39*). In particular, we selected 20-50 kDa proteoforms observed at >4% relative abundance for on-tissue MS/MS aiming for where they are in highest abundance on the tissue section. For each target proteoform, we selected a <0.8 *m/z* wide isolation window corresponding to the most abundant charge state of the proteoform obtained from PiMS data at MS^1-level (Fig. 1, right and Materials and Methods section). By matching the originally observed intact mass with the subsequent fragmentation data, we confidently identified 21 proteoforms >20 kDa with E-values ranging from 10^{-12} to 10^{-160} (Table 1 and fig. S7a-u).

In Fig. 3, we highlight two representative >20 kDa proteoforms confidently identified by MS/MS () The 11-14.5 kDa region of the fragmentation spectrum of monoacetylated glutathione S-transferase A1 (GSTA1, 25,542 Da, Fig. 3A) contains abundant complementary fragments of a 14-amino-acid-long sequence tag (B95-B108, Y113-Y126), contributing to the confident identification of this proteoform (Graphical Fragment Map, GFM, shown in Fig. 3A). PiMS image of GSTA1 proteoform in Fig. 3A show localization of this proteoform to the kidney cortex region. On the higher mass end, vimentin (53,530 Da) was also identified (fragment spectrum and GFM shown in Fig. 3B), with 63 isotopically-resolved >15 kDa fragment ions above 1% relative abundance matching sequence fragments of vimentin (a few representative ones shown in Fig. 3B). PiMS image of vimentin (Fig. 3B) with localizations to the vasculature also confirms the identity of the proteoform.

At larger masses, identification of the proteoform by MS/MS is increasingly challenging. We obtained 3% sequence coverage for a proteoform of ~66.4 kDa, putatively identified as albumin (fig. S7t). Poor sequence coverage is likely due to presence of 17 disulfide linkages known to occur in albumin, thereby making PTM localization challenging. In one attempt to identify a proteoform centered at 70,900 Da, we found that the precursor proteoform with the best

database retrieval score was mesothelin isoform 2 (fig. S7u). The deviation in precursor mass (67,938 Da compared to 70,900 Da) may be attributed to modifications and/or isoform expression. Nevertheless, using top-down MS in PiMS, we were able to readily identify 21 human proteoforms ranging from 20 to 70 kDa in molecular mass.

Creation of a Kidney Proteoform Map

PiMS images allow for direct visualization of sub-mm anatomical structures and functional tissue units of human kidney sections with proteoform-level precision. PiMS of kidney tissue containing cortex, medulla, and vasculature regions shows distinct differences in the distribution of proteoforms across these vastly different anatomical regions (Fig. 4, optical image shown in Fig. 4A). The identification of kidney internal structures was supported by autofluorescence microscopy (Fig. 4B) (*40*) and periodic acid-Schiff (PAS) staining histology (Fig. 4C).

PiMS images of proteoforms in the kidney show distinct localizations (Fig. 4D-J). apolipoprotein A-1 (Fig. 4D) and GSTA2 (Fig. 4E) were found to be medulla-enriched and cortex-enriched proteoforms, respectively. Blood-abundant albumin (Fig. 4F) and alpha-crystallin B chain (Fig. 4G), a small heat shock protein, were ubiquitously expressed in most of the regions of the kidney, albeit at different relative abundances. Additionally, transgelin-2 (Fig. 4H) and vimentin (Fig. 4J) were found only abundant in highly-focused regions near the artery.

Assessing Proteoform Biology: Differences in Space and Molecular Composition

Next, we created composite PiMS proteoform images to enable more efficient readout of anatomical regions and functional tissue units in the kidney. Using the abovementioned single PiMS images, we created a series of tricolor composite PiMS images (Fig. 4K-M). By combining images of apolipoprotein A-1 (Fig. 4D, blue), GSTA2 (Fig. 4E, red) and albumin (Fig. 4F, green),

we highlighted vascular regions within the kidney (Fig. 4K, labeled as bulk vasculature). This confirms the highly vascularized inner medulla, along with distinct points of vascularization in the cortex aligning with glomeruli. Fig. 4L is a composite PiMS image of transgelin-2 (Fig. 4H), GSTA2 (Fig. 4E, blue) and alpha-crystallin B chain (Fig. 4G, red). Using this image, we highlight the large artery via the specific localization of transgelin-2 (Fig. 4L, green). Previous studies have shown the specific localization of transgelin-2 to the smooth muscle cells (SMCs). This observation is consistent with the abundance of SMC in arteries whereas veins are mainly comprised of stromal cells (*41*). Finally, the localization of vimentin (Fig. 4J, green) shows a more refined vasculature (Fig. 4M) when combined with GSTA2 (Fig. 4E, blue) and alpha-crystallin B chain (Fig. 4G, red). Vimentin is a filament protein found in most of the blood vessels and connective tissues. The localization of vimentin to the bulk blood vessel regions by PiMS confirms this. A majority of the scattered vimentin spots in the cortex fit well into the dark spots corresponding to glomeruli. Moreover, vimentin is found in many scattered locations in the inner medulla region. These illuminated spots correspond to the spatially-dispersed peritubular capillaries, which form a complex three-dimensional network in the medulla and become dispersed when the tissue is cross-sectioned.

Another major advantage of PiMS over antibody-based imaging approaches lies in the ability to determine the molecular composition of proteoforms in an untargeted fashion. Antibody-based imaging approaches do not distinguish different proteoforms of a single protein, whereas PiMS can capture sequence differences and modifications. Subtle differences of protein sequences and modifications become especially challenging to detect for high-mass proteins. Direct top-down identification of >20 kDa proteoforms from tissue enabled by PiMS allows for the characterization of highly similar kidney protein isoforms that originate from allelic coding single nucleotide polymorphisms (cSNPs).

Three major proteoforms of N-terminal acetylated GST subunits, GSTA1 and GSTA2, were observed in kidney tissue with localizations to the cortex region (Fig. 5B, 5D, and 5E). Fig. 5C shows the mass domain spectrum of GSTA1 and GSTA2 proteoforms. A single proteoform was detected from GSTA1 (25,542 Da), showing two alleles with the same sequence. In contrast, we identified two GSTA2 proteoforms, the canonical form at 25,573 Da, and another form at 25,587 Da representing a 14 Da mass shift. Both proteoforms were observed at similar abundances with highly similar tissue localization, which is characteristic of non-specific biallelic tissue expression resulting from a cSNP (Fig. 5D and 5E). Fragmentation data was able to localize this mass shift to the region between 110 and 113 from the N-terminus (regions highlighted in red, Fig. 5D and 5E and fig. S7d). A UniProt search shows a Ser111→Thr natural variant of GSTA2 (highlighted in blue), confirming that the 14 Da mass shift corresponds to a proteoform resulting from a common cSNP (allele frequency >40% according to dbSNP entry No. rs2180314). This exemplifies the power of PiMS to the probing of gene expression in tissues directly at the proteoform level, which is in necessity and complementary to genomic and transcriptomic predictions.

Discussion

Imaging methods have boomed in recent years. Highly multiplexed affinity reagent-based methods allow for the detection of more than 50 protein targets in a single assay (*6*). The increased throughput of these approaches has provided an opportunity to develop comprehensive maps of human tissues at a rapid pace. However, antibody-based imaging approaches do not distinguish different proteoforms of a single protein, whereas PiMS can capture kidney protein isoforms from different gene family members or allelic cSNPs. With the concept that proteoforms define cell types better than proteins (*42*), driving the spatial resolution of PiMS down to <10 μM is a major goal to.

Moreover, PiMS presents low-bias to the protein masses compared to multiplex imaging assays, with the ability of capturing a majority of the abundant cytosolic proteins in cells. In fact, while performing IMT search of PiMS data against a database containing 100 most abundant proteins from bottom-up proteomics of human kidney tissue, 70 protein candidates were detected by PiMS. This provides an exceptional opportunity of using PiMS to delve into the abundant biological pathways in tissue in a spatially-resolved manner. In the Gene Ontology (GO) analysis result shown in Fig. 2F, a variety of central cellular metabolic pathways were found enriched in PiMS. Notably, we were able to detect and image four key enzymes in glycolysis and 27 subunits of the complexes in the respiratory electron transport chain, which are critical component of ATP metabolic process, the highest enriched pathway.

Another advantage of whole proteoform measurement is the robust detection of diverse types of post-translational modifications with low bias and knowledge of their stoichiometry. For example, two proteoforms of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were identified in the kidney dataset, both missing the start methionine and one containing a dimethylation site at Lys65. While both of these post-translational modifications are known (*43*, *44*), the dimethylation of Lys65 was shown to be very low in abundance and inconsistently observed (*44*). Our data indicates that the dimethylated proteoform is about ~5% of total GAPDH in human kidney. The function of the dimethylation is unknown, but its colocalization with the unmethylated proteoform suggests a possible role for the modification in the GAPDH dimerization or catalytic rate (*45*).

In conclusion, we present the PiMS approach that combines nano-DESI with I^2MS technology, which enables the direct imaging and molecular identification of human kidney proteoforms up to ~70 kDa. This approach increases observable masses by nearly 4-fold and resolving power by 10-fold compared to prior work (*37*). PiMS as a disseminatable imaging approach opens up exciting opportunities to infuse proteoform knowledge into the multi-omic

approaches being evaluated for inclusion into the Human Reference Atlas (*46*). By providing spatial localization of proteoforms to anatomical regions, cell types, and functional tissue units, PiMS promises applications molecular tissue mapping, biomarker discovery, and disease diagnostics.

Materials and Methods

Tissue Preparation

Human kidney tissue sections were prepared according to published protocols (*47*). Mouse brain tissue sections were sectioned at −21°C to a 12 µm thickness using a CM1850 Cryostat (Leica Microsystems, Wetzlar, Germany). Tissue sections were thaw mounted onto glass microscope slides (IMEB, Inc Tek-Select Gold Series Microscope Slides, Clear Glass, Positive Charged) and stored at −80 °C before mass spectrometry imaging analysis.

Human kidney tissue sections were thawed under slight vacuum at room temperature, fixed and desalted via successive immersion in 70%/30%, 90%/10%, and 100%/0% ethanol/H_2O solutions for 20 s each, delipidated by 99.8% chloroform for 60s, and dried under slight vacuum right before nano-DESI imaging experiments. These sample preparation steps allow for the *in situ* precipitation of proteins and removal of lipids avoiding suppression of protein signals upon ionization into the mass spectrometer (*19, 48*).

Nano-DESI Ion Source

A custom-designed nano-DESI source was used for all data acquisition. The experimental details of nano-DESI MSI have been described elsewhere (*30, 31*). Briefly, the nano-DESI probe is comprised of a primary (OD 150 µm, ID 20 µm) and a nanospray capillary (OD 150 µm, ID 40 µm) with the spray side of the nanospray capillary positioned close to the MS inlet. The probe was fabricated using fused silica capillary tubings (Molex, Thief River Falls, MN). A liquid bridge formed at the location where the two capillaries meet is brought into contact with the tissue

section for analyte extraction. The liquid bridge is dynamically maintained by solvent propulsion from the primary capillary and instantaneous vacuum aspiration through the nanospray capillary. The extracted analytes are continuously transferred to a mass spectrometer inlet and ionized by ESI. Imaging experiments are performed by moving the sample under the nano-DESI probe in lines. The optimal scan rate is discussed in the next section. The strip step between the line scans was set to 150 µm to avoid overlap between the adjacent line scans. To ensure the stability of the nano-DESI probe during the imaging experiment, we applied a surface tilt angle correction to the tissue sample by defining a three-point plane prior to the imaging experiment (*49*). All samples were electrosprayed under denaturing conditions in a 60%/39.4% acetonitrile/water and 0.6% acetic acid solution compatible with both protein extraction and ionization. All the experiments were performed in positive ionization mode. The source conditions on the mass spectrometer were set as follows: ESI voltage: 3 kV; in-source CID: 15 eV; S-lens RF level: 70%; capillary temperature: 325 ºC.

PiMS Conditions and Data Acquisition

PiMS data acquisition was performed in the individual ion mass spectrometry (I^2MS) mode which has been described previously on a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany, fig. S1) (*37*). In particular, rather than collecting typical ensemble ion MS spectra, ion signals were attenuated down to the individual ion regime by limiting the ion collection time in the C-trap before injection into the Orbitrap analyzer. The MS acquisition rate was set at 1 scan every 2 s. During PiMS data acquisition, proteoforms were sampled by a nano-DESI probe producing multiply-charged ions distributed across multiple charge states. To enable downstream I^2MS analysis, the majority of ions in one detection period was collected in the individual ion regime, corresponding to a singular ion signal at a defined *m/z* (or frequency) value. In particular, rastering line scans on tissue were performed at a reduced rate of 2.5-4 µm/s, which corresponds to a 5-8 µm sampling distance between adjacent pixels. The

nano-DESI solvent flow rate was kept at 600 nL/s to allow for efficient extraction and dilution of the proteins. The injection time for a specific set of tissue sections is typically optimized by acquiring line scans on an adjacent section prior to an imaging experiment, and it may vary from 100 to 500 ms. For 10 µm thin sections of human kidney tissue presented in this study, a 300 ms injection time was employed for all the sections from the same subject. We note that for extremely dominant proteoforms (e.g., hemoglobin subunits in this study), "multiple ion events" were commonly observed. The strip step between the line scans was set to 150 µm to avoid overlap between the adjacent line scans.

Additional MS instrument conditions in the PiMS experiment are mentioned here: the Orbitrap central electrode voltage was adjusted to -1 kV to improve the ion survival rate under denatured conditions. HCD pressure level was kept at 0.2 (UHV pressure $< 2 \times 10^{-11}$ Torr) to reduce collision-induced ion decay within the Orbitrap analyzer without completely losing the ion signal. Additional relevant data acquisition parameters were adjusted as follows: mass range: 400-2500 *m/z*; AGC mode was disabled and the maximum injection time was held constant at 300 ms; enhanced Fourier transform: off; averaging: 0; microscans: 1. Time-domain data files were acquired at detected ion frequencies and recorded as Selective Temporal Overview of Resonant Ions (STORI) files (*50*). STORI enabled: Enabled.

PiMS Data Analysis & Image Generation

Ion images were generated using a MATLAB script developed in-house. Mass-domain spectra were constructed by co-adding all individual ions obtained from the entire tissue section. In specific cases where computation power was limited, charge assignment and image construction were performed in sections with an upper limit of 50 million ions per portion. In the first step, all ion signals were subjected to STORI analysis to filter out decayed and multiple ion events. The neutral masses of the protein ions were calculated by:

$$Mass = \left(\frac{m}{z} \times z\right) - (z \times M_{proton})$$

Charge state (*z*) is obtained from the slope of induced image current determined by the STORI analysis (*50*). Accurate charge assignment of each ion was statistically evaluated by comparing the slopes of its isotopologues across different charge states from the entire tissue section. In particular, an iterative voting methodology was employed for filtering out ions with a lower probability score in the process, which allows for the construction of mass-domain isotopic distribution of a proteoform with statistical confidence. In this step, we utilized a Kernel Density Estimation (KDE) approach to convert centroid masses of individual ions to uniform distributions. Accurate masses of the isotopes were obtained from the center of the summed individual ion profiles.

For image generation, ions composing the mass-domain isotopic envelope of a protein were registered back to their spatial origins on the tissue section for PiMS image generation. A ±10 ppm isotopic mass tolerance was used to select individual ions for image generation. A raw image was first generated using absolute ion counts at different x-y locations. In the kidney PiMS images presented in this study, each pixel was constructed from three adjacent MS scans corresponding to ~10 µm × 150 µm area. The raw image was normalized using a total ion count matrix, which accounts for the fluctuation of sampling conditions at different locations.

Intact Mass Tag (IMT) Search & Gene Ontology (GO) analysis

The summed mass-domain PiMS spectrum was converted to *.mzML* format and processed using a custom version of TD Validator (Proteinaceous, Evanston, IL) implemented with an MS^1 IMT search function. The PiMS spectrum was shifted by +4 ppm according to the accurate masses of six MS/MS identified proteoforms in the 20-50 kDa mass range. A human protein database constructed from top 500 most abundant proteins in a bottom-up proteomics study of human kidney tissues was used for the search (Table S2). Methionine on/off and monoacetylation

were considered as possible proteoform modifications in the database. IMT search was performed with a ±5 ppm mass tolerance. Additional proteoform matches were curated by spectrum inspection and manual annotation of putative modifications recorded in Swiss-Prot human proteome database. The final 169 proteoform matches include MS/MS-identified proteoforms and all IMT-identified proteoforms discussed above.

Gene Ontology (GO) analysis was performed using Metascape (https://metascape.org/) (*51*). Specifically, a list of Entrez Gene ID was retrieved for the 169 identified proteoforms on Uniprot and submitted to Metascape for GO analysis. The result contains the top-level GO biological processes.

On-tissue top-down proteomics data analysis

Targeted MS/MS experiments were performed on a tissue section adjacent to the imaged section using higher-energy collisional dissociation (HCD). In the first step, a target proteoform was selected in the mass-domain spectrum. We utilized the PiMS image of the target proteoform to select a target area on the section where the proteoform abundance is enhanced. For the selected area, the mass-domain PiMS spectrum was convoluted back to *m/z* domain, from which a proper isolation window that contains predominantly the target proteoform was selected. A 0.8 *m/z* isolation window was typically employed for most of the targets; in special cases, 0.5-0.6 *m/z* window was used to avoid overlapping signal. MS/MS experiments were performed by scanning the nano-DESI probe over the selected region with the selected isolation window at 2.5-4 µm/s scan rate. MS/MS data acquisition was conducted in the I^2MS mode with an Orbitrap detection period of 2 s (HCD pressure setting = 0.5) (*39*). HCD collision energy and injection time was optimized to maximize the population of individual ion fragments. Typical ranges of collision energy and injection time used in this study were 7-14 eV and 200-1500 ms, respectively. Total data acquisition time for each target varied from 1-5 hours.

MS/MS data was first subjected to I^2MS processing for fragment ion charge assignment and mass-domain spectrum construction following the same procedure as described above. Mass-domain spectrum was converted to *.mzML* format subjected to MS^2 search function implemented in ProSight Native (Proteinaceous, Evanston, IL) to look for possible candidates from the entire human protein database. For each search, the top 1-5 candidates were manually validated using a custom version of TDValidator (Proteinaceous, Evanston, IL) to identify the best matching proteoform. Proteoform E-values were obtained from ProSight Native and TD Validator reports.

Acknowledgments

Funding:

National Institutes of Health UH3 CA246635 (NLK)

National Institutes of Health P41 GM108569 (NLK)

National Institutes of Health P30 DA018310 (NLK)

National Institutes of Health P30 CA060553 (awarded to the Robert H. Lurie Comprehensive Cancer Center)

National Institutes of Health UH3CA255132 (JL)

National Institutes of Health U54DK120058 (JMS)

National Institute of Environmental Health Sciences T32ES007028 (EKN).

NIH National Cancer Institute 5 UM1 CA183727-08 (Cooperative Human Tissue Network at Vanderbilt University Medical Center).

Author contributions:

Conceptualization: PS, JMC, JOK, NLK

Methodology: PS, MY, FAB, JOK, JL, NLK

Resources: MY, EKN, JLA

Software: KRD, MARH, JBG, BPE, RTF

Investigation: PS, EKN, FAB

Visualization: PS, JPM, KRD, MARH, JBG, BPE, JOK

Supervision: JMC, JOK, NLK

Writing—original draft: PS, JMC, NLK

Writing—review & editing: PS, JPM, KRB, MARH, MY, EKN, JLA, BSD, FAB, JBG, BPE, RTF, JMC, JOK, JMS, JL, NLK

Competing interests:

N.L.K., K.R.D, and J.O.K. report a conflict of interest with I^2MS technology, currently being commercialized by Thermo Fisher Scientific.

Data and materials availability:

Custom compiled code used to process and create I^2MS files is already available (*52*). Additional software and data that support the findings of this study are available from the corresponding authors upon request.

Figures and Tables

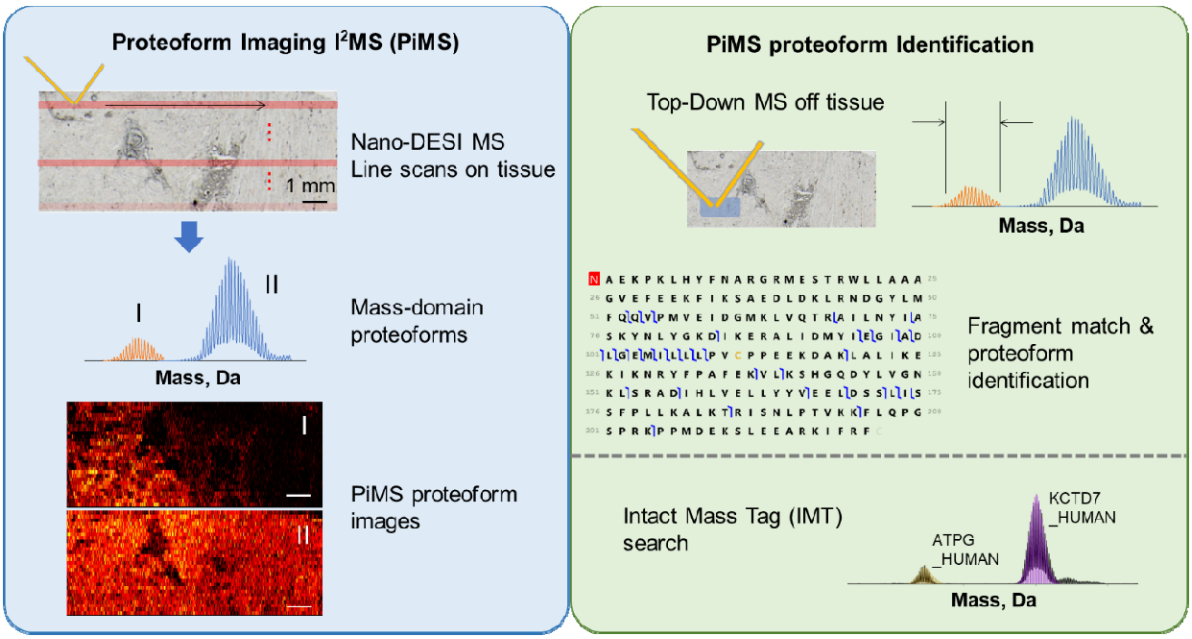

Fig. 1. **Illustration of the PiMS workflow for proteoform imaging and identification.** The left panel shows the scanning approach (top), detection of proteoforms in the mass domain (middle) and image reconstruction (bottom). The right panel depicts the two approaches to identify proteoforms using either direct fragmentation of proteoform ions and spectral readout by individual ion MS/MS (top) or database lookup of accurate mass values (IMT, bottom).

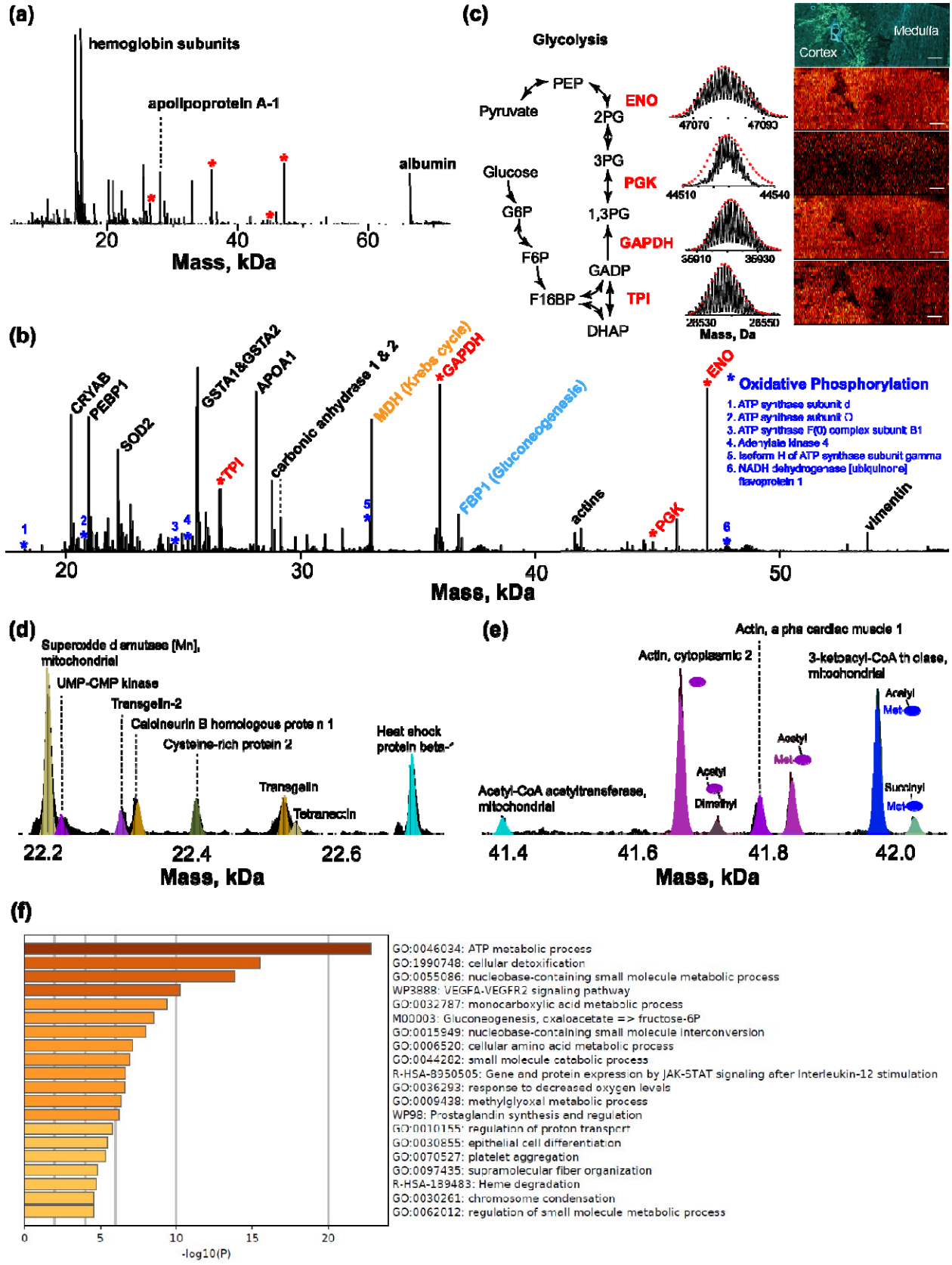

Fig. 2. Sum of mass domain spectrum obtained from PiMS of human kidney. (a) Full scale PiMS spectrum in 5-72 kDa summed from 16700 MS scans; regions occupied by abundant blood

proteins (hemoglobin subunits, albumin) are labeled in the spectrum; red asterisks denote the key glycolytic enzymes found in the spectrum. (b) PiMS full spectrum in the 18-56 kDa range zoomed in from (a); major proteins identified using a variety of approaches are labeled in the spectrum; aside from glycolytic enzymes (red asterisks), enzymes involved in a few other major metabolic pathways (Krebs cycle, gluconeogenesis, oxidative phosphorylation) found in PiMS are labeled in the spectrum. (c) Theoretical (red triangles) and experimentally-observed (black trace) isotopic distributions of the four glycolytic enzymes (labeled with red asterisks in (a) and (b)) together with their PiMS images depicted in a schematic diagram of the glycolysis metabolic pathway. The cortex and the medulla of the kidney section imaged are labeled in the autofluorescence image at top. Scale bar: 1 mm. Selected mass range of 22.2-22.7 kDa (d) and 41.4-42.1 kDa (e) showing identified proteoforms (spectrum in black, theoretical isotopic distributions in color). (f) GO analysis of biological pathways found enriched in the 169 identified proteoforms from the PiMS experiment shown in the -log(10)P scale.

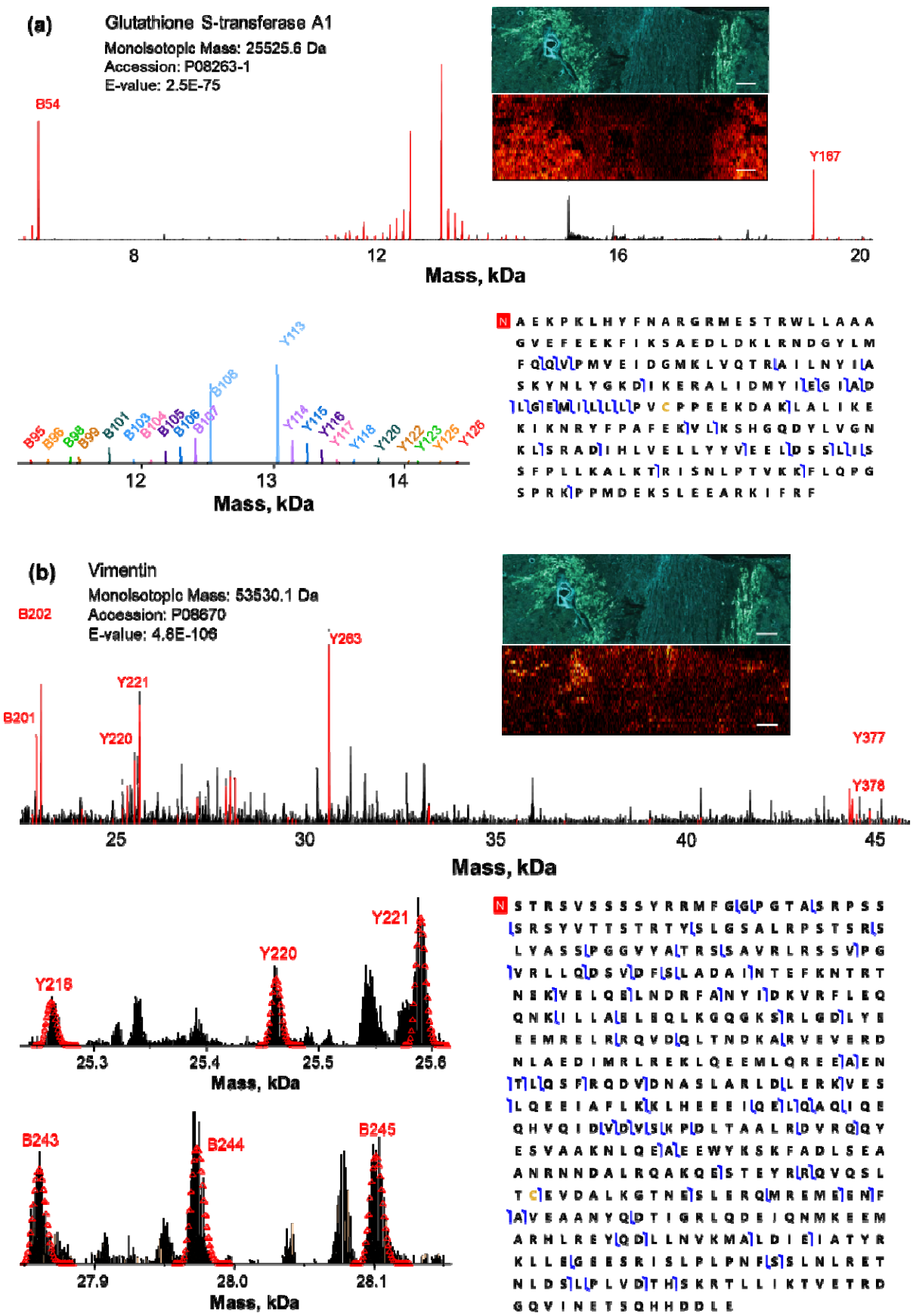
(a)
Glutathione S-transferase A1
Monoisotopic Mass: 25525.6 Da
Accession: P08263-1
E-value: 2.5E-75
B54
Y167
8
12
16
20
Mass, kDa
Y113
B108
B95
B96
B98
B99
B101
B103
B104
B105
B106
B107
Y114
Y115
Y116
Y117
Y118
Y120
Y122
Y123
Y125
Y126
12
13
14
Mass, kDa
(b)
Vimentin
Monoisotopic Mass: 53530.1 Da
Accession: P08670
E-value: 4.8E-106
B202
B201
Y221
Y220
Y263
Y377
Y378
25
30
35
40
45
Mass, kDa
Y218
Y220
Y221
25.3
25.4
25.5
25.6
Mass, kDa
B243
B244
B245
27.9
28.0
28.1
Mass, kDa

Fig. 3. On-tissue identification of proteoforms by MS/MS. Two representative human kidney proteoforms, glutathione S-transferase A1 (a) and vimentin (b). The PiMS images of the two proteoforms are shown on the top right of each panel along with an autofluorescence image of an adjacent section as a reference. On the bottom left of each panel, expanded regions of the fragment spectrum are displayed with the major matching fragment ions annotated. GFMs are shown on the bottom right of each panel. Scale bar: 1 mm.

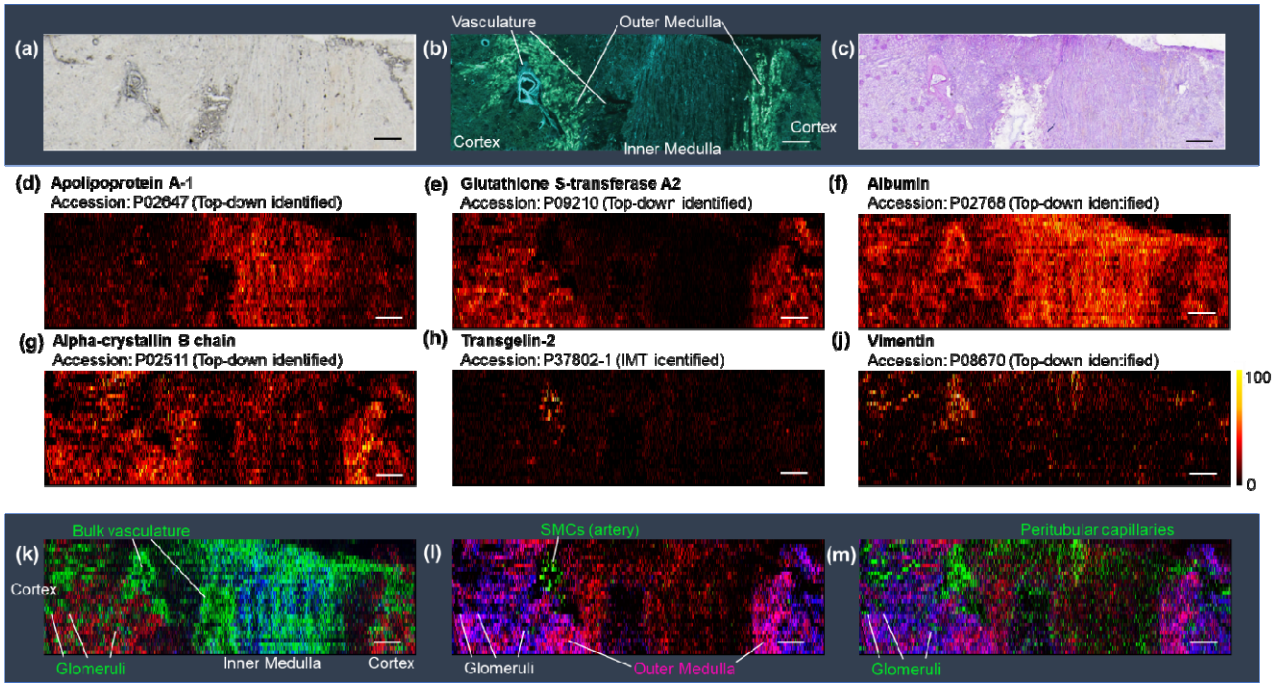

Fig. 4. Kidney Proteoform Maps. Optical (a) and autofluorescence (b) images of adjacent human kidney sections containing the cortex, medulla, and vasculature regions; (c) PAS staining of an adjacent section from the same kidney; (d)-(j) PiMS images of individual proteoforms that selectively illuminate different anatomical regions and cellular neighborhoods, the name and UniProt accession of the proteoforms are depicted next to the images with their color scale. Composite image of (k) apolipoprotein A-1 (blue), glutathione S-transferase A2 (red), and albumin (green); (l) glutathione S-transferase A2 (blue), alpha-crystallin B chain (red), and Transgelin-2 (green); and (m) glutathione S-transferase A2 (blue), alpha-crystallin B chain (red), and vimentin (green). Scale bar: 1 mm.

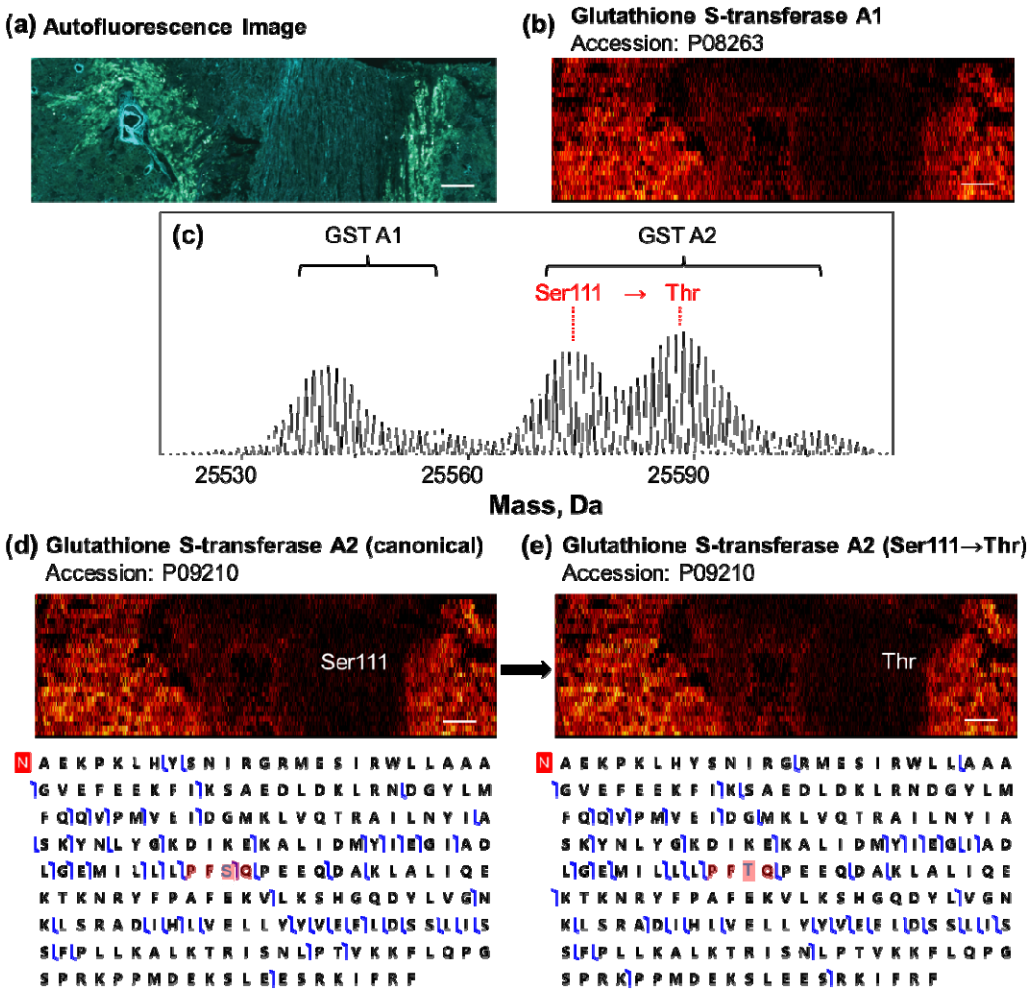

Fig. 5. Mass spectrometric detection of gene differences. (a) Autofluorescence and PiMS images of alpha GST enzymes of the human kidney section (b, d, and e). Mass domain PiMS spectra of ~25.6 kDa range (c) shows GSTA1 & GSTA2 enzyme proteoforms. The GFMs of the two proteoforms of GSTA2 from a known biallelic cSNP are shown in (d) and (e) below the PiMS images. The sequence variation of the two cSNP GSTA2 proteoforms are highlighted in the GFMs in blue.

509 **Table 1. Proteoforms identified by on-tissue MS/MS.**

Protein description	Gene	Accession	PFR	Observed mass
Alpha-crystallin B chain	CRYAB	P02511	PFR00000325920	20200.47
Superoxide dismutase [Mn], mitochondrial	SOD2	P04179-1	PFR00000001332	22204.25
Membrane-spanning 4-domains subfamily A member 4A	M4A4A	Q96JQ5	N/A	25431.57
Glutathione S-transferase A1	GSTA1	P08263-1	N/A	25540.66
Glutathione S-transferase A2	GSTA2	P09210	PFR00000432952	25573.626
Glutathione S-transferase A2	GSTA2	P09210	N/A	25587.64
V-type proton ATPase subunit E 1	VATE1	P36543-1	PFR00000053332	26055.743
Apolipoprotein A-1	APOA1	P02647	N/A	28078.5
Adaptin ear-binding coat-associated protein 1	NECAP1	Q8NC96-1	N/A	29764.03
N(G),N(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1	O94760	PFR00000058601	31031.95
Malate dehydrogenase, mitochondrial	MDHM	P40926-1	PFR00000002776	32999.46
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001134	35921.4
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001124	35950.39
Fructose-1,6-bisphosphatase 1	FBP1	P09467	PFR00006225388	36724.76
Actin, cytoplasmic 2	ACTG1	P63261	PFR00000001511	41660.74
Actin, cytoplasmic 1	ACTB	P60709	N/A	41660.74
Actin, cytoplasmic 2	ACTG1	P63261	N/A	41716.67
Alpha-enolase	ENOA	P06733	PFR00000001065	47078.34
Vimentin	VIME	P08670	PFR00000001800	53561.13
Albumin	ALBU	P02768	N/A	66438.68
Mesothelin	MSLN	Q13421-3	N/A	70900.52

Theoretical mass	Mass accuracy, ppm	Sequence coverage	E-value	Modifications
20200.46	0.50	35%	1.06E-131	Acetyl@N
22204.27	0.90	16%	2.08E-58	truncation 1-24
25431.5	2.75	4%	2.66E-12	Met-off, acetyl@N, phosphoryl@S
25540.67	0.39	15%	2.50E-75	Met-off, acetyl@N
25573.63	0.16	17%	1.38E-57	Met-off, acetyl@N
25587.64	0.00	20%	8.64E-87	Met-off, acetyl@N, 114 S to T
26055.77	1.04	14%	4.16E-54	2-225, alpha-amino acetylated residue@N
28078.53	1.07	17%	1.70E-65	Truncation 24-266
29764.21	6.05	18%	3.52E-75	Omega-N-methyl@R180, methyl@N
31031.95	0.00	16%	1.28E-74	Met-off, acetyl@N
32999.49	0.91	18%	6.08E-94	Truncation 1-24
35921.41	0.28	14%	4.48E-76	Met-off
35950.07	8.90	14%	8.64E-63	Met-off, dimethyl@K65
36724.82	1.63	22%	1.12E-121	Met-off, acetyl@N, 217 R to K
41660.82	1.92	24%	1.12E-154	Met-off
41660.78	0.96	24%	3.84E-160	Met-off, acetyl@N, methyl@H72
41716.84	4.08	20%	1.73E-115	Met-off, acetyl@N, methyl@H72/K83
47078.36	0.42	12%	1.60E-74	Met-off, acetyl@N
53561.13	0.00	16%	4.80E-106	Met-off, acetyl@N
66438.78	1.51	3%	8.32E-22	Truncation 1-24, disulfide bonds all Cys,
67980.42	N/A	7%	1.60E-45	Met-off, acetyl@N

References

1. M.-S. Kim, S. M. Pinto, D. Getnet, R. S. Nirujogi, S. S. Manda, R. Chaerkady, A. K. Madugundu, D. S. Kelkar, R. Isserlin, S. Jain, J. K. Thomas, B. Muthusamy, P. Leal-Rojas, P. Kumar, N. A. Sahasrabuddhe, L. Balakrishnan, J. Advani, B. George, S. Renuse, L. D. N. Selvan, A. H. Patil, V. Nanjappa, A. Radhakrishnan, S. Prasad, T. Subbannayya, R. Raju, M. Kumar, S. K. Sreenivasamurthy, A. Marimuthu, G. J. Sathe, S. Chavan, K. K. Datta, Y. Subbannayya, A. Sahu, S. D. Yelamanchi, S. Jayaram, P. Rajagopalan, J. Sharma, K. R. Murthy, N. Syed, R. Goel, A. A. Khan, S. Ahmad, G. Dey, K. Mudgal, A. Chatterjee, T.-C. Huang, J. Zhong, X. Wu, P. G. Shaw, D. Freed, M. S. Zahari, K. K. Mukherjee, S. Shankar, A. Mahadevan, H. Lam, C. J. Mitchell, S. K. Shankar, P. Satishchandra, J. T. Schroeder, R. Sirdeshmukh, A. Maitra, S. D. Leach, C. G. Drake, M. K. Halushka, T. S. K. Prasad, R. H. Hruban, C. L. Kerr, G. D. Bader, C. A. Iacobuzio-Donahue, H. Gowda, A. Pandey, A draft map of the human proteome. *Nature* **509**, 575-581 (2014).
2. N. L. Anderson, A. D. Matheson, S. Steiner, Proteomics: applications in basic and applied biology. *Current Opinion in Biotechnology* **11**, 408-412 (2000).
3. M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, Tissue-based map of the human proteome. *Science* **347**, (2015).
4. S. K. Saka, Y. Wang, J. Y. Kishi, A. Zhu, Y. Zeng, W. Xie, K. Kirli, C. Yapp, M. Cicconet, B. J. Beliveau, S. W. Lapan, S. Yin, M. Lin, E. S. Boyden, P. S. Kaeser, G. Pihan, G. M. Church, P. Yin, Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* **37**, 1080-1090 (2019).
5. E. Lundberg, G. H. H. Borner, Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 285-302 (2019).
6. Y. Goltsev, N. Samusik, J. Kennedy-Darling, S. Bhate, M. Hale, G. Vazquez, S. Black, G. P. Nolan, Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e915 (2018).
7. E. K. Neumann, N. H. Patterson, E. S. Rivera, J. L. Allen, M. Brewer, M. P. deCaestecker, R. M. Caprioli, A. B. Fogo, J. M. Spraggins, Highly multiplexed immunofluorescence of the human kidney using co-detection by indexing. *Kidney International*, (2021).
8. A. J. Radtke, E. Kandov, B. Lowekamp, E. Speranza, C. J. Chu, A. Gola, N. Thakur, R. Shih, L. Yao, Z. R. Yaniv, R. T. Beuschel, J. Kabat, J. Croteau, J. Davis, J. M. Hernandez, R. N. Germain, IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proceedings of the National Academy of Sciences* **117**, 33455 (2020).
9. M. J. Gerdes, C. J. Sevinsky, A. Sood, S. Adak, M. O. Bello, A. Bordwell, A. Can, A. Corwin, S. Dinn, R. J. Filkins, D. Hollman, V. Kamath, S. Kaanumalle, K. Kenny, M. Larsen, M. Lazare, Q. Li, C. Lowes, C. C. McCulloch, E. McDonough, M. C. Montalto, Z. Pang, J. Rittscher, A. Santamaria-Pang, B. D. Sarachan, M. L. Seel, A. Seppo, K. Shaikh, Y. Sui, J. Zhang, F. Ginty, Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of Sciences* **110**, 11982 (2013).
10. C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature methods* **11**, 417-422 (2014).

11. L. Keren, M. Bosse, S. Thompson, T. Risom, K. Vijayaragavan, E. McCaffrey, D. Marquez, R. Angoshtari, N. F. Greenwald, H. Fienberg, MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Science advances* **5**, eaax5851 (2019).
12. L. M. Smith, N. L. Kelleher, M. Linial, D. Goodlett, P. Langridge-Smith, Y. Ah Goo, G. Safford, L. Bonilla*, G. Kruppa, R. Zubarev, J. Rontree, J. Chamot-Rooke, J. Garavelli, A. Heck, J. Loo, D. Penque, M. Hornshaw, C. Hendrickson, L. Pasa-Tolic, C. Borchers, D. Chan, N. Young*, J. Agar, C. Masselon, M. Gross*, F. McLafferty, Y. Tsybin, Y. Ge, I. Sanders*, J. Langridge, J. Whitelegge*, A. Marshall, P. The Consortium for Top Down, Proteoform: a single term describing protein complexity. *Nature Methods* **10**, 186-187 (2013).
13. R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P. J. Hergenrother, C. G. Huber, A. R. Ivanov, O. N. Jensen, M. C. Jewett, N. L. Kelleher, L. L. Kiessling, N. J. Krogan, M. R. Larsen, J. A. Loo, R. R. Ogorzalek Loo, E. Lundberg, M. J. MacCoss, P. Mallick, V. K. Mootha, M. Mrksich, T. W. Muir, S. M. Patrie, J. J. Pesavento, S. J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schlüter, S. Sechi, S. A. Slavoff, L. M. Smith, M. P. Snyder, P. M. Thomas, M. Uhlén, J. E. Van Eyk, M. Vidal, D. R. Walt, F. M. White, E. R. Williams, T. Wohlschlager, V. H. Wysocki, N. A. Yates, N. L. Young, B. Zhang, How many human proteoforms are there? *Nature Chemical Biology* **14**, 206-214 (2018).
14. N. Siuti, N. L. Kelleher, Decoding protein modifications using top-down mass spectrometry. *Nature methods* **4**, 817-821 (2007).
15. D. P. Donnelly, C. M. Rawlins, C. J. DeHart, L. Fornelli, L. F. Schachner, Z. Lin, J. L. Lippens, K. C. Aluri, R. Sarin, B. Chen, Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature methods* **16**, 587-594 (2019).
16. M. Wilhelm, J. Schlegl, H. Hahne, A. M. Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, B. Kuster, Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582-587 (2014).
17. D. Unsihuay, D. M. Sanchez, J. Laskin, Quantitative Mass Spectrometry Imaging of Biological Systems. *Annual Review of Physical Chemistry* **72**, 307-329 (2021).
18. P. D. Piehowski, Y. Zhu, L. M. Bramer, K. G. Stratton, R. Zhao, D. J. Orton, R. J. Moore, J. Yuan, H. D. Mitchell, Y. Gao, Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100-μm spatial resolution. *Nature communications* **11**, 1-12 (2020).
19. E. H. Seeley, R. M. Caprioli, Molecular imaging of proteins in tissues by mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18126-18131 (2008).
20. P. Chaurand, S. A. Schwartz, R. M. Caprioli, Profiling and Imaging Proteins in Tissue Sections by MS. *Anal. Chem.* **76**, 86 A-93 A (2004).
21. A. Van Remoortere, R. J. Van Zeijl, N. Van den Oever, J. Franck, R. Longuespée, M. Wisztorski, M. Salzet, A. M. Deelder, I. Fournier, L. A. McDonnell, MALDI imaging and profiling MS of higher mass proteins from tissue. *Journal of the American Society for Mass Spectrometry* **21**, 1922-1929 (2010).
22. H. Liu, M. Han, J. Li, L. Qin, L. Chen, Q. Hao, D. Jiang, D. Chen, Y. Ji, H. Han, A Caffeic Acid Matrix Improves In Situ Detection and Imaging of Proteins with High Molecular Weight Close to 200,000 Da in Tissues by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging. *Analytical Chemistry*, (2021).

23. B. D. Leinweber, G. Tsaprailis, T. J. Monks, S. S. Lau, Improved MALDI-TOF imaging yields increased protein signals at high molecular mass. *Journal of the American Society for Mass Spectrometry* **20**, 89-95 (2009).
24. S. Nicolardi, D. P. A. Kilgour, N. Dolezal, J. W. Drijfhout, M. Wuhrer, Y. E. M. van der Burgt, Evaluation of Sibling and Twin Fragment Ions Improves the Structural Characterization of Proteins by Top-Down MALDI In-Source Decay Mass Spectrometry. *Analytical Chemistry* **92**, 5871-5881 (2020).
25. J. S. Sampson, A. M. Hawkridge, D. C. Muddiman, Generation and Detection of Multiply-Charged Peptides and Proteins by Matrix-Assisted Laser Desorption Electrospray Ionization (MALDESI) Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **17**, 1712-1716 (2006).
26. J. Laskin, I. Lanekoff, Ambient Mass Spectrometry Imaging Using Direct Liquid Extraction Techniques. *Analytical Chemistry* **88**, 52-73 (2016).
27. K. Y. Garza, C. L. Feider, D. R. Klein, J. A. Rosenberg, J. S. Brodbelt, L. S. Eberlin, Desorption electrospray ionization mass spectrometry imaging of proteins directly from biological tissue sections. *Analytical chemistry* **90**, 7785-7789 (2018).
28. R. L. Griffiths, J. W. Hughes, S. E. Abbatiello, M. W. Belford, I. B. Styles, H. J. Cooper, Comprehensive LESA mass spectrometry imaging of intact proteins by integration of cylindrical FAIMS. *Analytical chemistry* **92**, 2885-2890 (2020).
29. H. H. M. Yang, P. Su, P. M. Thomas, J. M. Camarillo, J. B. Greer, B. P. Early, R. T. Fellers, N. L. Kelleher, J. Laskin, Proteoform-selective imaging of tissues using mass spectrometry. *ChemRxiv*, (2021).
30. J. Laskin, B. S. Heath, P. J. Roach, L. Cazares, O. J. Semmes, Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. *Analytical chemistry* **84**, 141-148 (2012).
31. R. Yin, K. E. Burnum-Johnson, X. Sun, S. K. Dey, J. Laskin, High spatial resolution imaging of biological tissues using nanospray desorption electrospray ionization mass spectrometry. *Nature Protocols* **14**, 3445-3470 (2019).
32. C.-L. Chen, T.-H. Kuo, H.-H. Chung, P. Huang, L.-E. Lin, C.-C. Hsu, Remodeling nanoDESI platform with ion mobility spectrometry to expand protein coverage in cancerous tissue. *Journal of the American Society for Mass Spectrometry* **32**, 653-660 (2021).
33. O. J. Hale, H. J. Cooper, Native Mass Spectrometry Imaging of Proteins and Protein Complexes by Nano-DESI. *Analytical chemistry* **93**, 4619-4627 (2021).
34. T. P. Wörner, J. Snijder, A. Bennett, M. Agbandje-McKenna, A. A. Makarov, A. J. Heck, Resolving heterogeneous macromolecular assemblies by Orbitrap-based single-particle charge detection mass spectrometry. *Nature methods* **17**, 395-398 (2020).
35. R. A. Zubarev, A. Makarov, Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288-5296 (2013).
36. J. P. McGee, R. D. Melani, P. F. Yip, M. W. Senko, P. D. Compton, J. O. Kafader, N. L. Kelleher, Isotopic Resolution of Protein Complexes up to 466 kDa Using Individual Ion Mass Spectrometry. *Analytical Chemistry* **93**, 2723-2727 (2021).
37. J. O. Kafader, R. D. Melani, K. R. Durbin, B. Ikwuagwu, B. P. Early, R. T. Fellers, S. C. Beu, V. Zabrouskov, A. A. Makarov, J. T. Maze, Multiplexed mass spectrometry of individual ions improves measurement of proteoforms and their complexes. *Nature methods* **17**, 391-394 (2020).
38. D. Wang, B. Eraslan, T. Wieland, B. Hallström, T. Hopf, D. P. Zolg, J. Zecha, A. Asplund, L.-h. Li, C. Meng, M. Frejno, T. Schmidt, K. Schnatbaum, M. Wilhelm, F. Ponten, M. Uhlen, J. Gagneur, H. Hahne, B. Kuster, A deep proteome and transcriptome

abundance atlas of 29 healthy human tissues. *Molecular Systems Biology* **15**, e8503 (2019).

39. J. O. Kafader, K. R. Durbin, R. D. Melani, B. J. Des Soye, L. F. Schachner, M. W. Senko, P. D. Compton, N. L. Kelleher, Individual Ion Mass Spectrometry Enhances the Sensitivity and Sequence Coverage of Top-Down Mass Spectrometry. *J. Proteome Res.* **19**, 1346-1350 (2020).
40. N. H. Patterson, E. K. Neumann, K. Sharman, J. Allen, R. Harris, A. B. Fogo, M. de Caestecker, R. M. Caprioli, R. Van de Plas, J. M. Spraggins, Autofluorescence microscopy as a label-free tool for renal histology and glomerular segmentation. *bioRxiv*, 2021.2007.2016.452703 (2021).
41. J. Liu, Y. Zhang, Q. Li, Y. Wang, Transgelins: Cytoskeletal Associated Proteins Implicated in the Metastasis of Colorectal Cancer. *Frontiers in Cell and Developmental Biology* **8**, (2020).
42. N. L. Kelleher, A Cell-Based Approach to the Human Proteome Project. *Journal of The American Society for Mass Spectrometry* **23**, 1617-1624 (2012).
43. K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G. R. Thomas, J. Vandekerckhove, Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat. Biotechnol.* **21**, 566-569 (2003).
44. J. Seo, J. Jeong, Y. M. Kim, N. Hwang, E. Paek, K. J. Lee, Strategy for comprehensive identification of post-translational modifications in cellular proteins, including low abundant modifications: application to glyceraldehyde-3-phosphate dehydrogenase. *J. Proteome Res.* **7**, 587-602 (2008).
45. X.-Y. Zhong, X.-M. Yuan, Y.-Y. Xu, M. Yin, W.-W. Yan, S.-W. Zou, L.-M. Wei, H.-J. Lu, Y.-P. Wang, Q.-Y. Lei, CARM1 Methylates GAPDH to Regulate Glucose Metabolism and Is Suppressed in Liver Cancer. *Cell Reports* **24**, 3207-3223 (2018).
46. K. Börner, S. A. Teichmann, E. M. Quardokus, J. C. Gee, K. Browne, D. Osumi-Sutherland, B. W. Herr, A. Bueckle, H. Paul, M. Haniffa, L. Jardine, A. Bernard, S.-L. Ding, J. A. Miller, S. Lin, M. K. Halushka, A. Boppana, T. A. Longacre, J. Hickey, Y. Lin, M. T. Valerius, Y. He, G. Pryhuber, X. Sun, M. Jorgensen, A. J. Radtke, C. Wasserfall, F. Ginty, J. Ho, J. Sunshine, R. T. Beuschel, M. Brusko, S. Lee, R. Malhotra, S. Jain, G. Weber, Anatomical structures, cell types and biomarkers of the Human Reference Atlas. *Nature Cell Biology* **23**, 1117-1128 (2021).
47. E. K. Neumann, N. H. Patterson, J. L. Allen, L. G. Migas, H. Yang, M. Brewer, D. M. Anderson, J. Harvey, D. B. Gutierrez, R. C. Harris, M. P. deCaestecker, A. B. Fogo, R. Van de Plas, R. M. Caprioli, J. M. Spraggins, Protocol for multimodal analysis of human kidney tissue by imaging mass spectrometry and CODEX multiplexed immunofluorescence. *STAR Protocols* **2**, 100747 (2021).
48. E. H. Seeley, S. R. Oppenheimer, D. Mi, P. Chaurand, R. M. Caprioli, Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *Journal of the American Society for Mass Spectrometry* **19**, 1069-1077 (2008).
49. I. Lanekoff, B. S. Heath, A. Liyu, M. Thomas, J. P. Carson, J. Laskin, Automated Platform for High-Resolution Tissue Imaging Using Nanospray Desorption Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **84**, 8351-8356 (2012).
50. J. O. Kafader, S. C. Beu, B. P. Early, R. D. Melani, K. R. Durbin, V. Zabrouskov, A. A. Makarov, J. T. Maze, D. L. Shinholt, P. F. Yip, STORI plots enable accurate tracking of individual ion signals. *Journal of The American Society for Mass Spectrometry* **30**, 2200-2203 (2019).

51. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner, S. K. Chanda, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* **10**, 1523 (2019).
52. J. O. Kafader, R. D. Melani, K. R. Durbin, B. Ikwuagwu, B. P. Early, R. T. Fellers, S. C. Beu, V. Zabrouskov, A. A. Makarov, J. T. Maze, D. L. Shinholt, P. F. Yip, D. Tullman-Ercek, M. W. Senko, P. D. Compton, N. L. Kelleher, Multiplexed mass spectrometry of individual ions improves measurement of proteoforms and their complexes. *Nat. Methods* **17**, 391-394 (2020).