

UC Davis

UC Davis Previously Published Works

Title

Highly multiplexed, label-free proteoform imaging of tissues by individual ion mass spectrometry.

Permalink

<https://escholarship.org/uc/item/52h8c846>

Journal

Science Advances, 8(32)

Authors

Su, Pei

McGee, John

Durbin, Kenneth

et al.

Publication Date

2022-08-12

DOI

10.1126/sciadv.abp9929

Peer reviewed

BIOCHEMISTRY

Highly multiplexed, label-free proteoform imaging of tissues by individual ion mass spectrometry

Pei Su¹, John P. McGee¹, Kenneth R. Durbin¹, Michael A. R. Hollas¹, Manxi Yang², Elizabeth K. Neumann³, Jamie L. Allen³, Bryon S. Drown¹, Fatma Ayaloglu Butun⁴, Joseph B. Greer¹, Bryan P. Early¹, Ryan T. Fellers¹, Jeffrey M. Spraggins^{3,5}, Julia Laskin², Jeannie M. Camarillo^{1,4}, Jared O. Kafader^{1,4}, Neil L. Kelleher^{1,4,6*}

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 fourfold compared to reported methods and reveals tissue localization of proteoforms at <80- μm spatial resolution. PiMS advances proteoform imaging by combining ambient nanospray desorption electrospray ionization with ion detection using individual ion mass spectrometry. We demonstrate highly multiplexed proteoform imaging of human kidney, annotating 169 of 400 proteoforms of <70 kDa using top-down MS and a database lookup of ~1000 kidney candidate proteoforms, including dozens of key enzymes in primary metabolism. PiMS images reveal distinct spatial localizations of proteoforms to both anatomical structures and cellular neighborhoods in the vasculature, medulla, and cortex regions of the human kidney. The benefits of PiMS are poised to increase proteome coverage for label-free protein imaging of tissues.

INTRODUCTION

Proteoforms are the protein-level products of gene expression and posttranslational modifications (PTMs) 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 use enzymes, metals, 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) have 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 and multiplexed ion beam imaging, use antibodies labeled with rare earth metals to detect protein localization for up to ~40 protein targets at once (10, 11). Despite the substantial advances in spatial resolution and sensitivity, antibody-based approaches require prior knowledge of the protein targets and do not provide 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) because of the broad mass range of proteome sampling (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 DESI (27), liquid extraction surface analysis (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 uses a subnanoliter 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 have been limited to <20-kDa species (27, 28, 32–34), with one report leading to the identification of 29-kDa Zn²⁺-bound carbonic anhydrase 2 (35). Here, we address this challenge using individual ion MS (I²MS). I²MS is a new Orbitrap-based charge detection technique (36–38) using individual ions. Compared to traditional ensemble MS techniques, I²MS workflow is compatible with 500 \times dilute samples and yields a 10- to 20-fold enhancement in resolving power (39). In particular, we combine nano-DESI (34) with I²MS (40) to create proteoform imaging MS (PiMS) for tissue imaging and direct identification of proteoforms up to ~70 kDa. We show ~400 isotopically resolved proteoform assignments from the human kidney and confidently identify 20 proteoforms in the 20- to 70-kDa range using tandem MS (MS/MS), illuminating differences in kidney architecture from the medulla, cortex, and vasculature. Incorporating I²MS, PiMS yielded 169 proteoform assignments/identifications based on

Copyright © 2022
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

¹Departments of Molecular Biosciences, Chemistry, and Chemical and Biological Engineering and the Feinberg School of Medicine, Northwestern University, Evanston, IL, USA. ²Department of Chemistry, Purdue University, West Lafayette, IN, USA. ³Department of Biochemistry and Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA. ⁴Proteomics Center of Excellence, Northwestern University, Evanston, IL, USA. ⁵Departments of Chemistry and Cell and Developmental Biology, Vanderbilt University, Nashville, TN, USA. ⁶Department of Biochemistry and Molecular Genetics, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.

*Corresponding author. Email: n-kelleher@northwestern.edu

MS/MS and intact mass matching [± 5 parts per million (ppm)] at 80- μm spatial resolution and demonstrates the potential to visualize the proteinaceous structures comprising human tissues.

RESULTS

Overview of PiMS workflow

PiMS illustrated in Fig. 1 combines nano-DESI imaging (34) with data acquisition and processing for individual ions (see Supplementary Text and fig. S1) (40). Specifically, we perform nano-DESI line scans on tissue, during which proteoforms are sampled as multiply charged ions distributed across multiple charge states (Fig. 1, top left). Instead of unresolved signals from overlapping charge states of protein mixtures typically observed for ensembles of ions, PiMS provides information of individual protein ions in the mass domain even when their charge states overlap in the mass-to-charge domain. Enabled by direct charge assignment to individual ion signals, PiMS generates mass-domain information of the individual proteoforms with resolution of their ^{13}C isotopic peaks in each pixel of the imaging data (fig. S1). This allows for confident assignment of proteoform masses with better than 2-ppm accuracy at 1σ in each pixel of the imaging data (Fig. 1, middle left). Beyond proteoform-specific images (Fig. 1, bottom left), molecular identification is achieved using direct top-down MS/MS off the tissue (Fig. 1, top right) supplemented by database searching from known intact proteoform 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. Figure 2A shows the full PiMS spectrum from 5 to 72 kDa from a sum of 16,500 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- to 80-kDa range was observed but not individually resolved because of limited resolution of a particularly dense set of proteoforms in this spectral range. Therefore, these were not included in either proteoform counts or annotation efforts in this study (fig. S2). Spectral attributes include a dynamic range of ~ 200 (using a signal-to-noise ratio of 3 as the limit of detection) and a mass resolution ($m/\Delta m$) of $\sim 100,000$. 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 on the results of 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 contains 1000 kidney candidate proteoforms and was constructed from the top 500 most abundant proteins identified in a bottom-up proteomics study of human kidney tissues (table S1) (41). To maximize the number of proteoform identification, we combined additional matches from top-down identification (details discussed in the next section) and manually inspected PTMs recorded in the Swiss-Prot database (41). 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 S2). Three technical replicates of PiMS on kidney tissue sections from the same patient show $>85\%$ overlap in the identified proteoforms

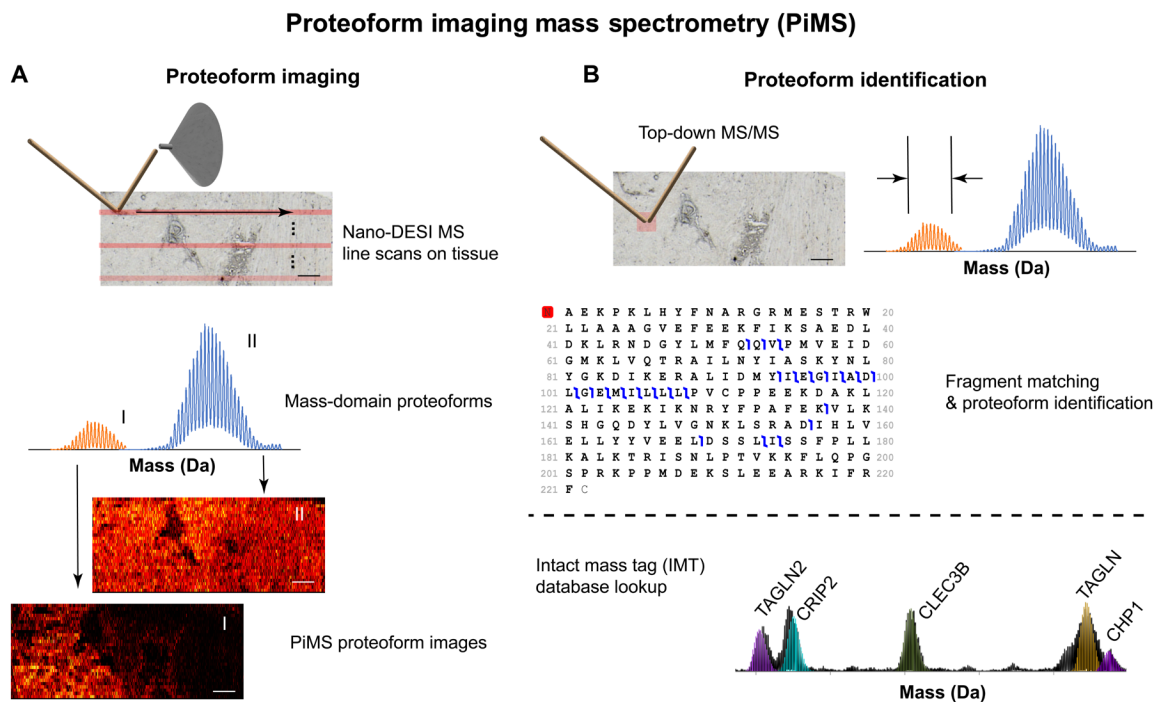


Fig. 1. Illustration of the PiMS workflow for proteoform imaging and identification. (A) Scanning approach (top), detection of proteoforms in the mass domain (middle), and image reconstruction (bottom). (B) Two approaches to identify proteoforms using either direct fragmentation of proteoform ions and spectral readout by individual ion MS/MS (top) or database lookup of accurate mass values (IMT, bottom). Scale bars, 1 mm.

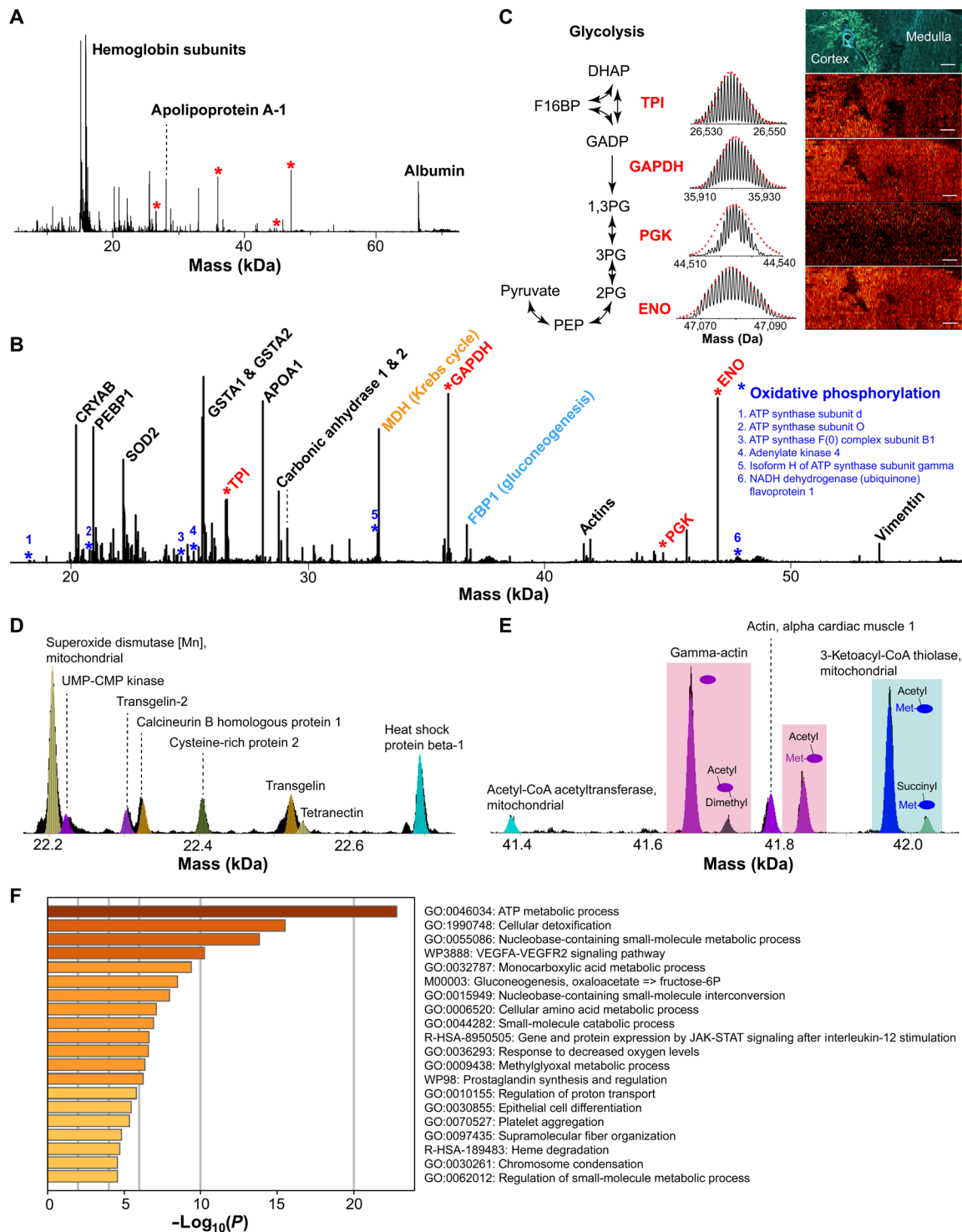


Fig. 2. Sum of mass-domain spectrum obtained from PiMS of human kidney. (A) Full-scale PiMS spectrum in 5 to 72 kDa summed from 16,500 MS scans; regions occupied by abundant blood proteins (hemoglobin subunits and albumin) are labeled in the spectrum; red asterisks denote the key glycolytic enzymes found in the spectrum. (B) PiMS full spectrum in the 18- to 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, and 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 the top. Scale bars, 1 mm. Selected mass range of 22.2 to 22.7 kDa (D) and 41.4 to 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.

(list of proteoforms shown in tables S3 and S4; Venn diagram of overlapping proteoforms shown in fig. S8). Figure 2 (D and E) shows two zoomed regions of the full PiMS spectrum with theoretical proteoform matches highlighted in color. Figure 2D shows various unmodified proteoforms in the mass range of 22.2 to 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 to 42.1 kDa shown in Fig. 2E contains proteoforms of γ -actin (highlighted in pink) and 3-ketoacyl-CoA thiolase (highlighted in blue) exhibiting diverse PTMs and their combinations. Aside from the monoacetylated and dimethylated proteoform of γ -actin identified by top-down MS/MS, other modified proteoforms in the displayed mass range were manually annotated. We note that the annotated proteoforms were manually confirmed by practitioners with the knowledge necessary to greatly suppress instances of false assignments. 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. 2 (A and B) to demonstrate the portion of the human kidney proteome captured by PiMS. Not unexpectedly, blood proteins (hemoglobin subunits, apolipoprotein A-1, and albumin; Fig. 2A) were found at highest abundance because of 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 [α -crystallin B chain (CRYAB)], signaling modulator [phosphatidylethanolamine-binding protein 1 (PEBP1)], proteins for cellular detoxification [mitochondrial manganese superoxide dismutase (SOD2) and glutathione S-transferase A1 and A2 (GSTA1 and GSTA2)] and homeostasis (carbonic anhydrase 1 and 2), and structural proteins (actins and vimentin). Moreover, proteins participating in central metabolic pathways are dominant. In particular, we found key enzymes in the 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 the 18- to 56-kDa mass range; others recorded in table S2). More intriguingly, four key enzymes in glycolysis [triosephosphate isomerase (TPI), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase (PGK), and α -enolase (ENO)] were imaged and identified (Fig. 2C and red asterisks in Fig. 2, A and B). PiMS images of the four detected glycolytic enzymes are 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 many of the kidney cell types. A more complete investigation of the biological pathways found for the 169 identified proteoforms is demonstrated by a Gene Ontology (GO) analysis shown in Fig. 2F, which indicates that cellular metabolic processes (glycolysis, the Krebs cycle, and oxidative phosphorylation) are the predominant biological pathways observed in the PiMS experiment.

Top-down characterizations of kidney proteoforms in PiMS

Human kidney proteoforms were characterized using MS/MS. Direct fragmentation of protein ions >20 kDa is challenging because of the low abundance of their fragment ions, particularly larger fragments (>15 kDa) that result from cleavage in the middle of the protein

sequence (42). To overcome this limitation, we used I²MS 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 (42). In particular, we selected >20-kDa proteoforms observed at >4% relative abundance for MS/MS, aiming for their most abundant locations in the tissue section. For each target proteoform, we selected a <0.8 *m/z* (mass/charge ratio) wide isolation window corresponding to the most abundant charge state of the proteoform obtained from PiMS data at MS¹ level (Fig. 1, right, and see Materials and Methods). By matching the originally observed intact mass with the subsequent fragmentation data, we confidently identified 20 proteoforms of >20 kDa (Table 1 and fig. S7, A to U).

In Fig. 3, we highlight two representative >20-kDa proteoforms confidently identified by MS/MS. The 11- to 14.5-kDa region of the fragmentation spectrum of monoacetylated GSTA1 (25,542 Da; Fig. 3A) contains abundant complementary fragments of a 14-amino acid-long sequence tag (B95-B108 and 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 shows that this proteoform is localized 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 the sequence fragments of vimentin (a few representative ones shown in Fig. 3B). PiMS image of vimentin (Fig. 3B) showing its localization to the vasculature also confirms the identity of the proteoform (43).

The identification of the >60-kDa proteoforms by MS/MS is increasingly challenging. We obtained 3% sequence coverage for a ~66.4-kDa proteoform putatively identified as albumin (fig. S7T). Poor sequence coverage obtained for this proteoform is likely due to the 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, which are not captured in the database. Despite these challenges, we were able to readily identify 20 human proteoforms ranging from 20 to 70 kDa in molecular mass using top-down MS/MS in PiMS.

Creation of a kidney proteoform map

PiMS images allow for direct visualization of submillimeter anatomical structures and functional tissue units of human kidney sections with proteoform-level precision. PiMS images of kidney tissue containing the cortex, medulla, and vasculature regions show 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) (44) and periodic acid-Schiff (PAS) staining histology (Fig. 4C).

PiMS images of proteoforms in the kidney show distinct localizations (Fig. 4, D to I). Apolipoprotein A-1 (Fig. 4D) and GSTA2 (Fig. 4E) were found to be enhanced in the medulla and cortex regions, respectively. Blood-abundant albumin (Fig. 4F) and α -crystallin B chain (Fig. 4G), a molecular chaperone, were expressed in multiple regions of the kidney, with particular enhancement in the inner and outer medulla, respectively. In addition, transgelin-2

Table 1. Proteoforms identified by on-tissue MS/MS. PFR, proteoform record number; N/A, not available; ATPase, adenosine triphosphatase.

Protein description	Gene	Accession	PFR	Observed mass	Theoretical mass	Mass accuracy (ppm)	Sequence coverage	Modifications
α -crystallin B chain	CRYAB	P02511	PFR00000325920	20,200.47	20,200.46	0.5	35%	Acetyl@N
Mitochondrial manganese superoxide dismutase	SOD2	P04179-1	PFR00000001332	22,204.25	22,204.27	0.9	16%	25-221
Membrane-spanning 4-domains subfamily A member 4A	M4A4A	Q96IQ5	N/A	25,431.57	25,431.50	2.8	4%	Met-off, acetyl@N, phosphoryl@S
Glutathione S-transferase A1	GSTA1	P08263-1	N/A	25,540.66	25,540.67	0.4	15%	Met-off, acetyl@N
Glutathione S-transferase A2	GSTA2	P09210	PFR00000432952	25,573.63	25,573.63	0.2	17%	Met-off, acetyl@N
Glutathione S-transferase A2	GSTA2	P09210	N/A	25,587.64	25,587.64	0.0	20%	Met-off, acetyl@N, 114 S→T
V-type proton ATPase subunit E1	VATE1	P36543-1	PFR00000053332	26,055.74	26,055.77	1.0	14%	2-225, α -amino acetyl@N
Apolipoprotein A-1	APOA1	P02647	N/A	28,078.50	28,078.53	1.0	17%	24-266
Adaptin ear-binding coat-associated protein 1	NECAP1	Q8NC96-1	N/A	29,764.03	29,764.21	6.1	18%	Omega-N-methyl@R180, methyl@N
N(G)/M(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1	O94760	PFR00000058601	31,031.95	31,031.95	0.0	16%	Met-off, acetyl@N
Malate dehydrogenase, mitochondrial	MDHM	P40926-1	PFR00000002776	32,999.46	32,999.49	0.9	18%	25-337
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001134	35,921.40	35,921.41	0.3	14%	Met-off
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001124	35,950.39	35,950.07	9.0	14%	Met-off, dimethyl@K65
Fructose-1,6-bisphosphatase 1	FBP1	P09467	PFR00006225388	36,724.76	36,724.82	1.6	22%	Met-off, acetyl@N, 217 R to K
Actin, cytoplasmic 1	ACTB	P60709	N/A	41,660.74	41,660.78	1.0	24%	Met-off, acetyl@N, methyl@H72
Actin, cytoplasmic 2	ACTG1	P63261	PFR00000001511	41,660.74	41,660.82	1.9	24%	Met-off
Actin, cytoplasmic 2	ACTG1	P63261	N/A	41,716.67	41,716.84	4.1	20%	Met-off, acetyl@N, methyl@H72/K83
α -enolase	ENOA	P06733	PFR00000001065	47,078.34	47,078.36	0.4	12%	Met-off, acetyl@N
Vimentin	VIME	P08670	PFR00000001800	53,561.13	53,561.13	0.0	16%	Met-off, acetyl@N
Albumin	ALBU	P02768	N/A	66,438.68	66,438.78	1.5	3%	25-608, 16 disulfide bonds, phosphoryl@S, acetyl@N, O@M
Mesothelin	MSLN	Q13421-3	N/A	70,900.52	67,980.42	N/A	7%	Met-off, acetyl@N

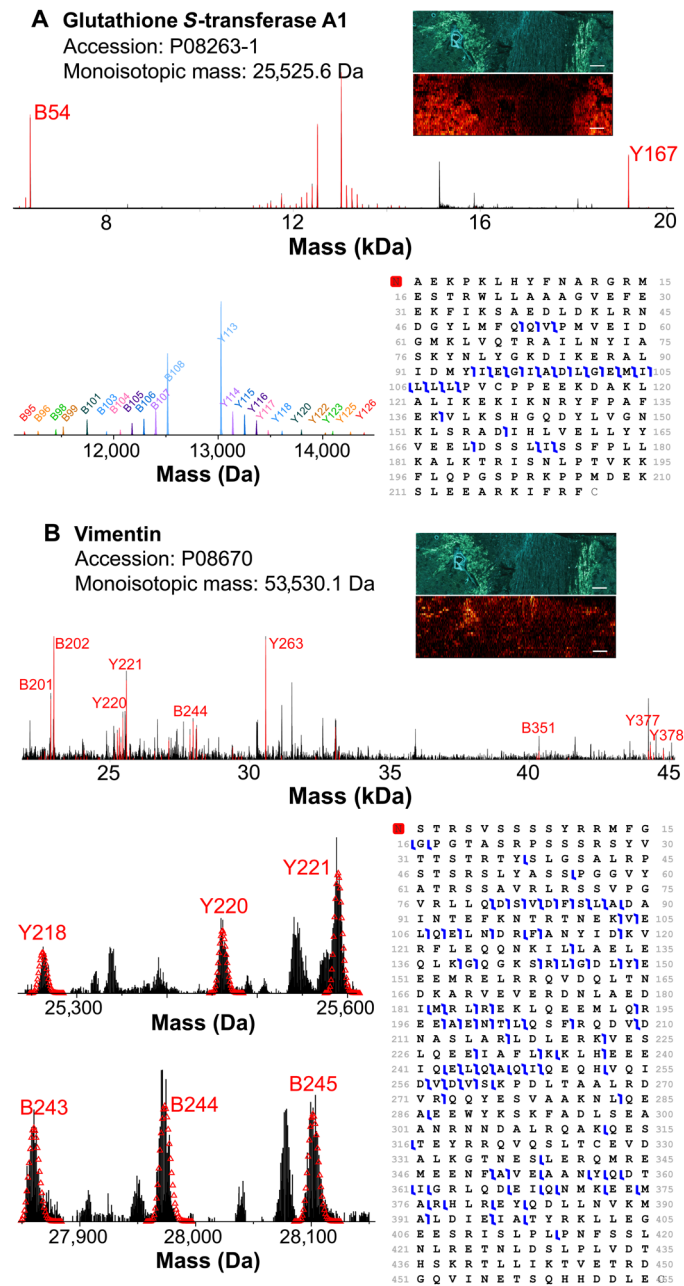


Fig. 3. On-tissue identification of proteoforms by MS/MS. Two representative human kidney proteoforms, GSTA1 (A) and vimentin (B). The PiMS images of the two proteoforms are shown on the top right of each panel along with an autofluorescence image of an adjacent section as a reference. On the bottom 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 bars, 1 mm.

(Fig. 4H) and vimentin (Fig. 4I) were 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. 4, J to L). 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. 4J, labeled as bulk vasculature), including the bulk vasculature and the inner medulla. In addition, the distinct points of vascularization in the cortex may be aligned with glomeruli (green). This is further supported by GSTA2 (red) localized to the tubules surrounding the glomeruli (3). Figure 4K is a composite PiMS image of transgelin-2 (Fig. 4H, green), GSTA2 (Fig. 4E, blue), and α -crystallin B chain (Fig. 4G, red). Using this image, we highlight the large artery via the specific localization of transgelin-2 (Fig. 4K, 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 composed of stromal cells (45). Last, the combination of the spatial distribution of vimentin (Fig. 4I, green), GSTA2 (Fig. 4E, blue), and α -crystallin B chain (Fig. 4G, red) provides the best outline of the vasculature in the tissue section (Fig. 4L). Vimentin is a filamentous protein found in most of the blood vessels and connective tissues. The localization of vimentin to the bulk blood vessel regions observed by PiMS is consistent with this. Most of the scattered vimentin spots observed 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.

Moreover, a 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. 5, B, D, and E). Figure 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 nonspecific biallelic tissue expression resulting from a cSNP (Fig. 5, D and E). Fragmentation data were able to localize the 14-Da mass shift to the region between Pro¹¹⁰ and Gln¹¹³ from the N terminus (regions highlighted in light red in Fig. 5, D and E, and fig. S7D). A UniProt search shows a Ser¹¹¹→Thr natural variant of GSTA2 (highlighted in red), 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 complementary to genomic and transcriptomic predictions.

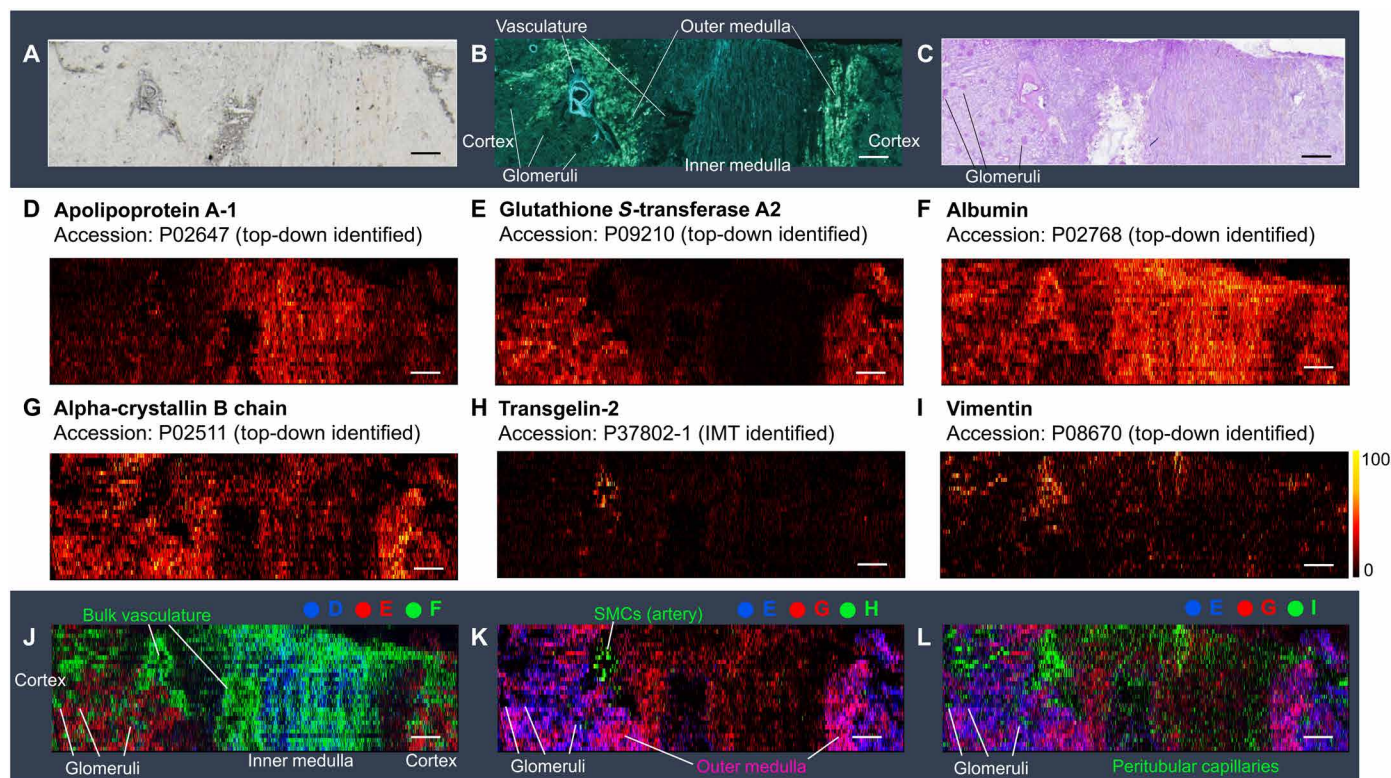


Fig. 4. Kidney proteoform maps. Optical (A) and autofluorescence (B) images of adjacent human kidney sections containing the cortex, medulla, and vasculature regions. (C) PAS staining of an adjacent section from the same kidney. (D to I) PiMS images of individual proteoforms that selectively illuminate 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 (J) apolipoprotein A-1 (blue) (D), GSTA2 (red) (E), and albumin (green) (F); (K) GSTA2 (blue) (E), α -crystallin B chain (red) (G), and transgelin-2 (green) (H); and (L) GSTA2 (blue) (E), α -crystallin B chain (red) (G), and vimentin (green) (I). Scale bars, 1 mm.

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, 46). The increased throughput of these approaches has provided an opportunity to develop comprehensive maps of proteins in human tissues at a rapid pace (4, 8, 9). However, antibody-based imaging approaches do not distinguish different proteoforms of a single protein. The development described in this study addresses this challenge and enables imaging of protein isoforms from different gene family members or allelic cSNPs in biological tissues.

PiMS is able to capture many abundant cytosolic and mitochondrial proteins in cells within one acquisition process, some of which have been previously investigated using antibody-based approaches (e.g., vimentin and GSTA2) (46). In an IMT search of the PiMS data against a database containing 100 most abundant proteins found in bottom-up proteomics of human kidney (i.e., 200 candidate proteoforms from each protein with and without the start methionine), 56 proteoform hits were asserted by PiMS (54 and 46 hits from the other two technical replicates). This provides an opportunity of using PiMS to delve into the major biological pathways in tissues in a spatially resolved manner. In the GO analysis result based on the 169 annotated proteoforms shown in Fig. 2F, a variety of central cellular metabolic pathways were found enriched by PiMS in many kidney cell types. 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 components of adenosine 5'-triphosphate (ATP) metabolic process, the highest enriched pathway.

Meanwhile, we acknowledge that nuclear and membrane proteins in kidney cells were not detected at high abundance in PiMS, suggesting that major sources of bias exist such as solubility, ion suppression, and limitations in detection at very high mass or low abundance. Despite the limitations of PiMS in its current form, these findings were absent in denaturing, global top-down MS experiments (47). In fact, a comparison of a global top-down MS and PiMS of human kidney tissues results in only 9 overlapping proteoforms, with 160 unique proteoforms annotated from PiMS (47). This lack of overlap is mainly due to the strong bias against the detection of >30-kDa proteins in global top-down MS, and more detailed comparison of biological pathways detectable in PiMS will await future studies with expanded proteome coverage. Clearly, uncovering the many additional proteoforms present in the human proteome will require another major leap forward in proteoform imaging technology (48). In this first iteration of the PiMS technology, only a small percentage of all proteoforms were revealed. While this limits the biological conclusions that can be drawn, future iterations of the technology can improve the detection rate of proteoforms, enabling deeper insights into biological function. Future work will optimize sample pretreatment and extraction solvent compositions to expand the detection of proteoforms localized to various subcellular compartments. Furthermore, low-abundance proteoforms may be captured by increasing the total number of ions collected in a single PiMS experiment.

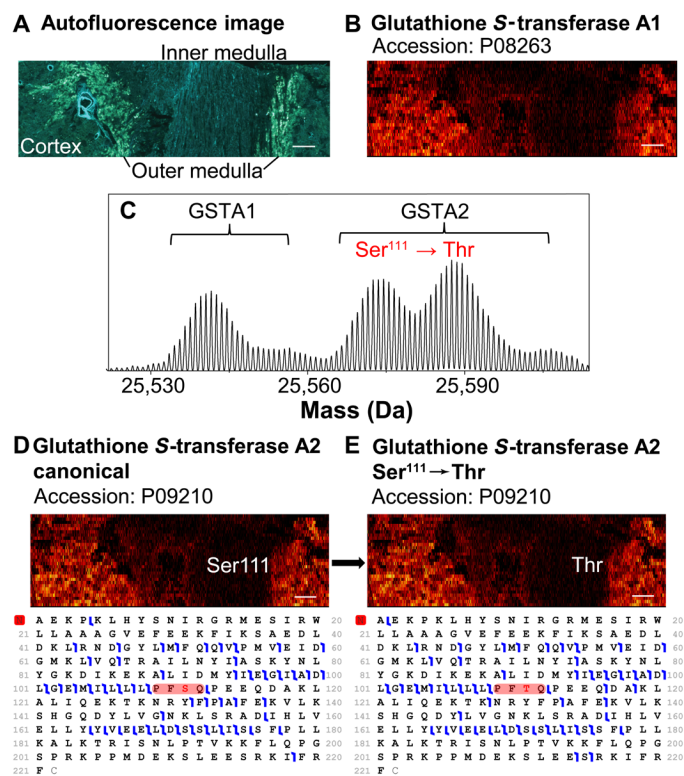


Fig. 5. Mass spectrometric detection of gene differences. (A) Autofluorescence and PiMS images of α GST enzymes of the human kidney section (B, D, and E). Mass-domain PiMS spectra of \sim 25.6-kDa range (C) shows GSTA1 and GSTA2 enzyme proteoforms. The GFMs of the two proteoforms of GSTA2 from a known biallelic cSNP are shown in (D) and (E) below the PiMS images. The sequence variations of the two cSNP GSTA2 proteoforms are highlighted in the GFMs in red. Scale bars, 1 mm.

PiMS expands the proteoform detection to 70-kDa mass range with $>10\times$ feature capacity compared to multiplexed imaging assays. More than 50% of the human proteins reside in the mass range between 20 and 70 kDa (48). The ability in measuring the proteome with a mass coverage of up to 70 kDa opens up the opportunities in revealing a myriad of biological pathways and molecular functions (i.e., the oxidative phosphorylation and mitochondrial proteins reported here), which were hidden in prior MS imaging techniques. High-mass proteoforms of >70 kDa with generally lower abundance and increased molecular heterogeneity (e.g., the idea that the larger the protein, the greater the number of high-stoichiometry PTMs) perhaps contribute to size limitations of proteoform imaging in its current form. Furthermore, high-mass proteoforms can experience poor desolvation and high decay rate during MS detection because of large cross sections. Future research effort to detect high-mass proteoforms will focus on these challenges.

Another advantage of whole-proteoform measurement is the robust detection of diverse types of PTMs with low bias and knowledge of their stoichiometry. For example, two proteoforms of GAPDH were identified in the kidney dataset, both missing the start methionine and one containing a dimethylation site at Lys⁶⁵. While both of these PTMs are known (49, 50), the dimethylation of Lys⁶⁵ was shown to be very low in abundance and inconsistently observed (50). Our data indicate that the dimethylated proteoform is about \sim 5% of the total

GAPDH in the human kidney. The function of the dimethylation is unknown, but its colocalization with the unmethylated proteoform suggests a possible role of the modified GAPDH proteoform in dimerization, which may contribute to its catalytic properties in glycolysis (51).

In conclusion, we present the PiMS approach that combines nano-DESI with I²MS technology, which enables the direct imaging and molecular identification of human kidney proteoforms of up to \sim 70 kDa. This approach increases observable proteoform masses by nearly 4-fold and resolving power by 10-fold compared to prior work (40). The development of PiMS opens up exciting opportunities to infuse proteoform knowledge into the multiomic approaches being evaluated for inclusion into the Human Reference Atlas (52). By providing spatial localization of proteoforms to anatomical regions, cell types, and functional tissue units, PiMS promises applications in molecular tissue mapping, biomarker discovery, and disease diagnostics.

MATERIALS AND METHODS

Tissue preparation

Human kidney tissue sections were prepared according to published protocols (53). Mouse brain tissue sections were sectioned at -21°C to a thickness of 12 μm using a CM1850 cryostat (Leica Microsystems, Wetzlar, Germany). Institutional Review Board and/or Institutional Animal Care and Use Committee guidelines were followed with human or animal subjects. 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 MS 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/H₂O solutions for 20 s each, delipidated by 99.8% chloroform for 60 s, 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, 54).

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 (29, 30). Briefly, the nano-DESI probe is composed of a primary [outer diameter (OD), 150 μm ; inside diameter (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 tubing (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 before

the imaging experiment (55). 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.

PiMS conditions and data acquisition

PiMS data acquisition was performed in the I^2MS mode on a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany), which has been described previously (fig. S1) (40). The source conditions on the mass spectrometer were set as follows: ESI voltage, 3 kV; in-source collision-induced dissociation, 15 eV; S-lens radio frequency level, 70%; capillary temperature, 325°C. 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 one 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, most of the ions in one detection period were 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 was performed at a reduced rate of 2.5 to 4 $\mu\text{m/s}$, which corresponds to a 5- to 8- μm sampling distance between adjacent pixels (fig. S3). 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 before an imaging experiment, and it may vary from 100 to 500 ms (figs. S4 to S6). For 10- μm -thin sections of human kidney tissue presented in this study, a 300-ms injection time was used 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.

Additional MS instrument conditions in the PiMS experiment are as follows: The Orbitrap central electrode voltage was adjusted to -1 kV to improve the ion survival rate under denatured conditions. Higher-energy collisional dissociation (HCD) pressure level was kept at 0.2 (ultrahigh vacuum pressure $< 2 \times 10^{-11}$ torr) to reduce collision-induced ion decay within the Orbitrap analyzer without substantial losses in trapping efficiency. Additional relevant data acquisition parameters were adjusted as follows: mass range, 400 to 2500 m/z ; automatic gain control 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 Ion (STORI) files (56).

PiMS data analysis and image generation

Ion images were generated using a MATLAB script developed in-house (available from the authors upon request). 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

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

Charge state (z) is obtained from the slope of induced image current determined by the STORI analysis (56). Accurate charge assignment of each ion was statistically evaluated by comparing the slopes of its isotopologs across different charge states from the entire tissue section. In particular, an iterative voting methodology was used 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 used a kernel density estimation 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 by 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.

IMT search and GO analysis

The summed mass-domain PiMS spectrum was converted to .mzML format and processed using a custom version of TDValidator (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- to 50-kDa mass range. A human protein database constructed from the top 500 most abundant proteins in a bottom-up proteomic 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.

GO analysis was performed using Metascape (<https://metascape.org/>) (57). 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.

Top-down proteomic data acquisition and analysis

Targeted MS/MS experiments were performed on a tissue section adjacent to the imaged section using HCD. In the first step, a target proteoform was selected in the mass-domain spectrum. We used 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 used for most of the targets; in special cases, 0.5 to 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 a scan rate of 2.5 to 4 $\mu\text{m/s}$. MS/MS data acquisition was conducted in the I^2MS mode with an Orbitrap detection period of 2 s (HCD pressure setting = 0.5) (42). HCD collision energy and injection

time were optimized to maximize the population of individual ion fragments. Typical ranges of collision energy and injection time used in this study were 7 to 14 eV and 200 to 1500 ms, respectively. Total data acquisition time for each target varied from 1 to 5 hours.

MS/MS data were first subjected to I²MS 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 MS² search function implemented in ProSight Native (Proteinaceous, Evanston, IL) to look for possible candidates from the entire Swiss-Prot human protein database (www.uniprot.org/proteomes/UP000005640) (58, 59). For each search, the top one to five candidates were manually validated using a custom version of TDValidator (Proteinaceous, Evanston, IL) to identify the best matching proteoform.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abp9929>

REFERENCES AND NOTES

- 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. Sahasrabudde, 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. Satischandra, J. T. Schroeder, R. Sirdeshmukh, A. Maitra, S. D. Leach, C. G. Drake, M. K. Halushka, T. S. K. Prasad, R. H. Brubaker, 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).
- N. L. Anderson, A. D. Matheson, S. Steiner, Proteomics: Applications in basic and applied biology. *Curr. Opin. Biotechnol.* **11**, 408–412 (2000).
- M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. M. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szijarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Pontén, Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
- 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).
- E. Lundberg, G. H. H. Borner, Spatial proteomics: A powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 285–302 (2019).
- 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.e15 (2018).
- 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 Int.* **101**, 137–143 (2021).
- 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. M. Hernandez, R. N. Germain, IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proc. Natl. Acad. Sci.* **117**, 33455–33465 (2020).
- 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. *Proc. Natl. Acad. Sci.* **110**, 11982–11987 (2013).
- C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schöffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther, B. Bodenmiller, Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
- L. Keren, M. Bosse, S. Thompson, T. Risom, K. Vijayaragavan, E. McCaffrey, D. Marquez, R. Angostari, N. F. Greenwald, H. Fienberg, J. Wang, N. Kambham, D. Kirkwood, G. Nolan, T. J. Montine, S. J. Galli, R. West, S. C. Bendall, M. Angelo, MIBI-TOF: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci. Adv.* **5**, eaax5851 (2019).
- L. M. Smith, N. L. Kelleher, Consortium for Top Down Proteomics, Proteoform: A single term describing protein complexity. *Nat. Methods* **10**, 186–187 (2013).
- 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. O. 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. Wohlschläger, V. H. Wysocki, N. A. Yates, N. L. Young, B. Zhang, How many human proteoforms are there? *Nat. Chem. Biol.* **14**, 206–214 (2018).
- N. Siuti, N. L. Kelleher, Decoding protein modifications using top-down mass spectrometry. *Nat. Methods* **4**, 817–821 (2007).
- 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. *Nat. Methods* **16**, 587–594 (2019).
- 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. Bantschkeff, A. Gerstmair, F. Faerber, B. Kuster, Mass-spectrometry-based draft of the human proteome. *Nature* **509**, 582–587 (2014).
- D. Unsihuay, D. M. Sanchez, J. Laskin, Quantitative mass spectrometry imaging of biological systems. *Annu. Rev. Phys. Chem.* **72**, 307–329 (2021).
- 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, B.-J. M. Webb-Robertson, S. K. Dey, R. T. Kelly, K. E. Burnum-Johnson, Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100- μ m spatial resolution. *Nat. Commun.* **11**, 8 (2020).
- 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).
- 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).
- A. Van Remoortere, R. J. Van Zeijl, N. Van den Oever, J. Franck, R. Longuespée, M. Wiszorski, M. Salzet, A. M. Deelder, I. Fournier, L. A. McDonnell, MALDI imaging and profiling MS of higher mass proteins from tissue. *J. Am. Soc. Mass Spectrom.* **21**, 1922–1929 (2010).
- H. Liu, M. Han, J. Li, L. Qin, L. Chen, Q. Hao, D. Jiang, D. Chen, Y. Ji, H. Han, C. Long, Y. Zhou, J. Feng, X. Wang, 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. *Anal. Chem.* **93**, 11920–11928 (2021).
- B. D. Leinweber, G. Tsaprailis, T. J. Monks, S. S. Lau, Improved MALDI-TOF imaging yields increased protein signals at high molecular mass. *J. Am. Soc. Mass Spectrom.* **20**, 89–95 (2009).
- 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. *Anal. Chem.* **92**, 5871–5881 (2020).
- J. S. Sampson, A. M. Hawkrige, 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. *J. Am. Soc. Mass Spectrom.* **17**, 1712–1716 (2006).
- J. Laskin, I. Lanekoff, Ambient mass spectrometry imaging using direct liquid extraction techniques. *Anal. Chem.* **88**, 52–73 (2016).
- 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. *Anal. Chem.* **90**, 7785–7789 (2018).
- 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. *Anal. Chem.* **92**, 2885–2890 (2020).
- J. Laskin, B. S. Heath, P. J. Roach, L. Cazares, O. J. Semmes, Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. *Anal. Chem.* **84**, 141–148 (2012).
- 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. *Nat. Protoc.* **14**, 3445–3470 (2019).
- R. Yin, J. Kyle, K. Burnum-Johnson, K. J. Bloodworth, L. Sussel, C. Ansong, J. Laskin, High spatial resolution imaging of mouse pancreatic islets using nanospray desorption electrospray ionization mass spectrometry. *Anal. Chem.* **90**, 6548–6555 (2018).

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. *J. Am. Soc. Mass Spectrom.* **32**, 653–660 (2021).
33. O. J. Hale, H. J. Cooper, Native mass spectrometry imaging of proteins and protein complexes by nano-DESI. *Anal. Chem.* **93**, 4619–4627 (2021).
34. M. Yang, H. Hu, 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. *Angew Chem. Int. Ed. Engl.* e202200721 (2022).
35. E. K. Sisley, O. J. Hale, I. B. Styles, H. J. Cooper, Native ambient mass spectrometry imaging of ligand-bound and metal-bound proteins in rat brain. *J. Am. Chem. Soc.* **144**, 2120–2128 (2022).
36. 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. *Nat. Methods* **17**, 395–398 (2020).
37. R. A. Zubarev, A. Makarov, Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).
38. 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. *Anal. Chem.* **93**, 2723–2727 (2021).
39. J. O. Kafader, R. D. Melani, M. W. Senko, A. A. Makarov, N. L. Kelleher, P. D. Compton, Measurement of individual ions sharply increases the resolution of orbitrap mass spectra of proteins. *Anal. Chem.* **91**, 2776–2783 (2019).
40. J. O. Kafader, R. D. Melani, K. R. Durbin, B. Ikwaugwu, 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).
41. D. Wang, B. Eraslan, T. Wieland, B. Hallström, T. Hopf, D. P. Zolg, J. Zecha, A. Asplund, L.-h. Li, C. Meng, M. Frejto, 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. *Mol. Syst. Biol.* **15**, e8503 (2019).
42. 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).
43. T. Yamasaki, N. Seki, Y. Yamada, H. Yoshino, H. Hidaka, T. Chiyomaru, N. Nohata, T. Kinoshita, M. Nakagawa, H. Enokida, Tumor suppressive microRNA-138 contributes to cell migration and invasion through its targeting of vimentin in renal cell carcinoma. *Int. J. Oncol.* **41**, 805–817 (2012).
44. 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).
45. J. Liu, Y. Zhang, Q. Li, Y. Wang, Transgelins: Cytoskeletal associated proteins implicated in the metastasis of colorectal cancer. *Front. Cell Dev. Biol.* **8**, 573859 (2020).
46. J. W. Hickey, E. K. Neumann, A. J. Radtke, J. M. Camarillo, R. T. Beuschel, A. Albanese, E. McDonough, J. Hatler, A. E. Wiblin, J. Fisher, J. Croteau, E. C. Small, A. Sood, R. M. Caprioli, R. M. Angelo, G. P. Nolan, K. Chung, S. M. Hewitt, R. N. Germain, J. M. Spraggins, E. Lundberg, M. P. Snyder, N. L. Kelleher, S. K. Saka, Spatial mapping of protein composition and tissue organization: A primer for multiplexed antibody-based imaging. *Nat. Methods* **19**, 284–295 (2022).
47. B. S. Drown, K. Joob, R. D. Melani, C. Lloyd-Jones, J. M. Camarillo, N. L. Kelleher, Mapping the proteoform landscape of five human tissues. *J. Proteome Res.* **21**, 1299–1310 (2022).
48. P. D. Compton, L. Zamborg, P. M. Thomas, N. L. Kelleher, On the scalability and requirements of whole protein mass spectrometry. *Anal. Chem.* **83**, 6868–6874 (2011).
49. 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).
50. 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).
51. 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 Rep.* **24**, 3207–3223 (2018).
52. 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. *Nat. Cell Biol.* **23**, 1117–1128 (2021).
53. 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 Protoc.* **2**, 100747 (2021).
54. 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. *J. Am. Soc. Mass Spectrom.* **19**, 1069–1077 (2008).
55. 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. *Anal. Chem.* **84**, 8351–8356 (2012).
56. 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, N. L. Kelleher, P. D. Compton, M. W. Senko, STORI plots enable accurate tracking of individual ion signals. *J. Am. Soc. Mass Spectrom.* **30**, 2200–2203 (2019).
57. 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. *Nat. Commun.* **10**, 1523 (2019).
58. F. Meng, B. J. Cargile, L. M. Miller, A. J. Forbes, J. R. Johnson, N. L. Kelleher, Informatics and multiplexing of intact protein identification in bacteria and the archaea. *Nat. Biotechnol.* **19**, 952–957 (2001).
59. L. F. Schachner, D. P. Tran, A. S. Lee, J. P. McGee, K. Jooss, K. R. Durbin, H. S. Seckler, L. Adams, E. N. Cline, R. D. Melani, A. N. Ives, B. D. Soye, N. L. Kelleher, S. M. Patrie, Reassembling protein complexes after controlled disassembly by top-down mass spectrometry in native mode. *Int. J. Mass Spectrom.* **465**, 116591 (2021).
60. C.-C. Hsu, P.-T. Chou, R. N. Zare, Imaging of proteins in tissue samples using nanospray desorption electrospray ionization mass spectrometry. *Anal. Chem.* **87**, 11171–11175 (2015).

Acknowledgments

Funding: This work was supported by NIH UH3 CA246635 (to N.L.K.), NIH P41 GM108569 (to N.L.K.), NIH P30 DA018310 (to N.L.K.), NIH P30 CA060553 (awarded to the Robert H. Lurie Comprehensive Cancer Center), NIH UH3CA255132 (to J.L.), NIH U54DK120058 (to J.M.S.), National Institute of Environmental Health Sciences T32ES007028 (to E.K.N.), NIH F32 CA246894 (to B.S.D.), and NIH National Cancer Institute 5 UM1 CA183727-08 (Cooperative Human Tissue Network at Vanderbilt University Medical Center). **Author contributions:** Conceptualization: P.S., J.M.C., J.O.K., and N.L.K. Methodology: P.S., M.Y., F.A.B., J.O.K., J.L., and N.L.K. Resources: M.Y., E.K.N., and J.L.A. Software: K.R.D., M.A.R.H., J.B.G., B.P.E., and R.T.F. Investigation: P.S., E.K.N., and F.A.B. Visualization: P.S., J.P.M., K.R.D., M.A.R.H., J.B.G., B.P.E., and J.O.K. Supervision: J.M.C., J.O.K., and N.L.K. Writing—original draft: P.S., J.M.C., and N.L.K. Writing—review and editing: P.S., J.P.M., K.R.D., M.A.R.H., M.Y., E.K.N., J.L.A., B.S.D., F.A.B., J.B.G., B.P.E., R.T.F., J.M.C., J.O.K., J.M.S., J.L., and N.L.K. **Competing interests:** N.L.K., K.R.D., and J.O.K. report a conflict of interest with I²MS technology, currently being commercialized by Thermo Fisher Scientific. All authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Custom compiled code used to process and create I²MS files is already available (40). The data can be provided by N.L.K. pending scientific review and a completed material transfer agreement. Requests for the presented material should be submitted to n-kelleher@northwestern.edu.

Submitted 10 March 2022

Accepted 24 June 2022

Published 10 August 2022

10.1126/sciadv.abp9929