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

¹Departments of Molecular Biosciences, Chemistry, 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, Evanston, IL, USA

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

⁶Department 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

26 **Abstract**

27 Imaging of proteoforms in human tissues is hindered by low molecular specificity and
 28 limited proteome coverage. Here, we introduce proteoform imaging mass spectrometry (PiMS),
 29 which increases the size limit for proteoform detection and identification by 4-fold compared to
 30 reported methods, and reveals tissue localization of proteoforms at <80 μm spatial resolution.
 31 PiMS advances proteoform imaging by combining ambient nanospray desorption electrospray
 32 ionization (nano-DESI) with ion detection using individual ion mass spectrometry (I^2MS). We
 33 demonstrate the first proteoform imaging of human kidney, identifying 169 of 400 proteoforms
 34 <70 kDa using top-down mass spectrometry and database lookup from the human proteoform
 35 atlas, including dozens of key enzymes in primary metabolism. PiMS images reveal distinct
 36 spatial localizations of proteoforms to both anatomical structures and cellular neighborhoods in
 37 the vasculature, medulla, and cortex regions of the human kidney. The benefits of PiMS are
 38 poised to increase proteome coverage for label-free protein imaging of tissues.

39

40

41 **Teaser**

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

44

45 Introduction

46 Proteoforms are the protein-level products of gene expression and post-translational
 47 modifications functioning as key effectors in human health and disease (1, 2). In addition to
 48 understanding of their molecular compositions, interactions, and biological function,
 49 comprehensive characterization of the human proteoform landscape also requires mapping of their
 50 spatial distributions in human tissues and organs (3). Protein-level imaging using antibody-based
 51 optical microscopy has revealed distinct cell types, functional tissue units, and subcellular structures
 52 (4). These techniques employ enzymes, metals, and fluorophores as reporters to obtain high-
 53 resolution maps of protein targets in tissues (5). In recent years, highly multiplexed antibody-based
 54 imaging assays such as CODEX (6, 7), IBEX (8), and Cell-DIVE (9) have drastically increased the
 55 number of protein targets that can be probed in a single experiment. Alternatively, mass
 56 spectrometry (MS)-based imaging assays, including imaging mass cytometry (IMC) and
 57 multiplexed ion beam imaging (MIBI), utilize antibodies labeled with rare earth metals to detect
 58 protein localization for up to ~60 protein targets at once (10, 11). Despite the significant advances
 59 in spatial resolution and sensitivity, antibody-based approaches require prior knowledge of the
 60 protein targets and do not provide proteoform-level information (12, 13).

61 MS-based top-down proteomics (14, 15) has been widely used for proteoform
 62 characterizations (16). Modern MS instrumentation has reached the sensitivity for spatially-
 63 resolved top-down proteomics suitable for imaging experiments (17, 18). Matrix-assisted laser
 64 desorption/ionization (MALDI) is widely used for protein imaging (19, 20) due to the broad mass
 65 range of proteome sampling (21-23). However, MALDI predominantly generates singly-charged
 66 ions, which gives limited fragment information for direct top-down identification of intact proteins
 67 (24). This challenge may be addressed using matrix-assisted laser desorption electrospray
 68 ionization (MALDESI), which combines MALDI with extractive ESI to generate multiply-charged
 69 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, 33), with one report leading to the identification of subunits from a 43 kDa trimeric protein complex (34). Here, we address this challenge using individual ion mass spectrometry ($I^2\text{MS}$). $I^2\text{MS}$ is a new Orbitrap-based charge detection technique (35-37) utilizing individual ions. Using this workflow yields a 10-20-fold enhancement in resolving power and 500 $\times$ dilute sample compatibility compared to traditional ensemble MS techniques (38). In particular, we combine nano-DESI ionization (33) with $I^2\text{MS}$ (39) to create proteoform imaging mass spectrometry (PiMS) for tissue imaging and direct identification of proteoforms up to ~70 kDa. We show ~400 isotopically-resolved proteoform assignments from human kidney and confidently identify proteoforms up to 53 kDa using MS/MS, illuminating differences in kidney architecture from the medulla, cortex, and vasculature. Incorporating $I^2\text{MS}$, PiMS yielded 169 proteoform assignments/identifications at 80 μm spatial resolution and demonstrates the potential to visualize the proteinaceous structures comprising human tissues.

92

93 **Results**

94 *Overview of PiMS workflow*

95 Proteoform imaging mass spectrometry (PiMS) illustrated in Fig. 1 combines nanospray
 96 desorption electrospray ionization (nano-DESI) imaging (33) with data acquisition and processing
 97 for individual ions (39). Specifically, we perform nano-DESI line scans on tissue, during which
 98 proteoforms are sampled as multiply-charged ions distributed across multiple charge states (Fig. 1,
 99 top left). Instead of unresolved protein signals typically observed for ensembles of ions, we detect
 100 charge-assigned individual protein ions as a function of location on the tissue section thereby
 101 generating mass spectral features of the individual proteoforms with resolution of their ^{13}C isotopic
 102 peaks in each pixel of the imaging data (fig. S1). This allows for confident assignment of
 103 proteoform masses with better than 2 parts-per-million accuracy at one sigma in each pixel of the
 104 imaging data (Fig. 1, middle left). Beyond proteoform-specific images (Fig. 1, bottom left),
 105 molecular identification is achieved using either direct top-down MS/MS off the tissue (Fig. 1, top
 106 right) or database searching from known proteoforms of the intact masses (Fig. 1, bottom right).

108 *Human Kidney Proteoforms Detected by PiMS*

109 We used PiMS to examine the proteoforms and their localizations in a 10 μm -thick human
 110 kidney tissue section. Encouragingly, we immediately expanded the mass detection range for
 111 proteoform imaging to >70 kDa. Fig. 2A shows the full PiMS spectrum from 5-72 kDa from a sum
 112 of 16500 MS scans (~ 8 million single ions). This spectrum contains ~ 400 proteoform masses above
 113 0.1% relative abundance that are isotopically resolved. A complex group of proteoforms in the 68-
 114 80 kDa range were observed, but not individually resolved (fig. S2). Spectral attributes include a
 115 dynamic range of ~ 200 (using S/N 3 as the limit of detection) and a mass resolution ($m/\Delta m$) of
 116 $\sim 100,000$. PiMS images can be constructed for any of the 242 proteoform masses with relative
 117 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 was constructed from the top 500 most abundant proteins identified in a bottom-up proteomics study of human kidney tissues (40). 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 (40). 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 gamma-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 gamma-actin 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

143 observed in PiMS. From low to high mass range shown in Fig. 2B, we found molecular chaperone
 144 (alpha-crystallin B chain, CRYAB), signaling modulator (phosphatidylethanolamine-binding
 145 protein 1, PEBP1), proteins for cellular detoxification (superoxide dismutase [Mn], mitochondrial,
 146 SOD2, glutathione S-transferase A1&A2, GSTA1&GSTA2) and homeostasis (carbonic anhydrase
 147 1&2), and structural proteins (actins, vimentin). Moreover, proteins participating in central
 148 metabolic pathways are dominant. In particular, we found key enzymes in Krebs cycle (malate
 149 dehydrogenase, MDH) and gluconeogenesis (fructose-1,6-bisphosphatase 1, FBP1), and 27
 150 subunits of protein complexes in the electron transport chain of oxidative phosphorylation (blue
 151 asterisks for 6 subunits in 18-56 kDa mass range, others recorded in table S1). More intriguingly,
 152 four key enzymes in glycolysis (triosephosphate isomerase, TPI, glyceraldehyde-3-phosphate
 153 dehydrogenase, GAPDH, phosphoglycerate kinase, PGK, and alpha-enolase, ENO) were imaged
 154 and identified (Fig. 2C, red asterisks in Fig. 2A and 2B). PiMS images of the four detected
 155 glycolytic enzymes were shown in Fig. 2C with largely even distributions across the entire tissue
 156 section spanning from the kidney medulla to the cortex, confirming the presence of glycolysis in
 157 many of the kidney cell types. A more complete investigation of the biological pathways found for
 158 the 169 identified proteoforms is demonstrated by a Gene Ontology (GO) analysis shown in Fig.
 159 2F, which indicates that cellular metabolic processes is the predominant biological pathway
 160 observed in the PiMS experiment.

161

162 *Top-Down Characterizations of Kidney Proteoforms in PiMS*

163 Human kidney proteoforms were characterized using tandem MS (MS/MS). Direct
 164 fragmentation of protein ions >20 kDa is challenging due to the low abundance of their fragment
 165 ions, particularly larger fragments (>15 kDa) that result from cleavage in the middle of the protein
 166 sequence (41). To overcome this limitation, we employed I²MS for the readout of top-down
 167 fragmentation spectra to capture those large fragment ions typically buried under the noise level in

ensemble MS/MS experiments (41). 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 MS¹-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 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 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 (42).

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 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,

193 which are not captured in the database. Despite these challenges, we were able to readily identify
194 21 human proteoforms ranging from 20 to 70 kDa in molecular mass using top-down MS in PiMS.

195
196 *Creation of a Kidney Proteoform Map*

197 PiMS images allow for direct visualization of sub-mm anatomical structures and functional
198 tissue units of human kidney sections with proteoform-level precision. PiMS of kidney tissue
199 containing cortex, medulla, and vasculature regions shows distinct differences in the distribution of
200 proteoforms across these vastly different anatomical regions (Fig. 4, optical image shown in Fig.
201 4A). The identification of kidney internal structures was supported by autofluorescence microscopy
202 (Fig. 4B) (43) and periodic acid-Schiff (PAS) staining histology (Fig. 4C).

203 PiMS images of proteoforms in the kidney show distinct localizations (Fig. 4D-J).
204 Apolipoprotein A-1 (Fig. 4D) and GSTA2 (Fig. 4E) were found to be enhanced in the medulla and
205 cortex regions, respectively. Blood-abundant albumin (Fig. 4F) and alpha-crystallin B chain (Fig.
206 4G), a molecular chaperone, were ubiquitously expressed in most of the regions of the kidney.
207 Additionally, transgelin-2 (Fig. 4H) and vimentin (Fig. 4J) were abundant only in highly-focused
208 regions near the artery.

209
210 *Assessing Proteoform Biology: Differences in Space and Molecular Composition*

211 Next, we created composite PiMS proteoform images to enable more efficient readout of
212 anatomical regions and functional tissue units in the kidney. Using the abovementioned single PiMS
213 images, we created a series of tricolor composite PiMS images (Fig. 4K-M). By combining images
214 of apolipoprotein A-1 (Fig. 4D, blue), GSTA2 (Fig. 4E, red) and albumin (Fig. 4F, green), we
215 highlighted vascular regions within the kidney (Fig. 4K, labeled as bulk vasculature). This confirms
216 the highly vascularized inner medulla, along with distinct points of vascularization in the cortex
217 aligning with glomeruli. Fig. 4L is a composite PiMS image of transgelin-2 (Fig. 4H), GSTA2 (Fig.

218 4E, blue) and alpha-crystallin B chain (Fig. 4G, red). Using this image, we highlight the large artery
 219 via the specific localization of transgelin-2 (Fig. 4L, green). Previous studies have shown the
 220 specific localization of transgelin-2 to the smooth muscle cells (SMCs). This observation is
 221 consistent with the abundance of SMC in arteries whereas veins are mainly comprised of stromal
 222 cells (44). Finally, the combination of the spatial distribution of vimentin (Fig. 4J, green), GSTA2
 223 (Fig. 4E, blue), and alpha-crystallin B chain (Fig. 4G, red) provides the best outline of the
 224 vasculature in the tissue section. Vimentin is a filamentous protein found in most of the blood
 225 vessels and connective tissues. The localization of vimentin to the bulk blood vessel regions
 226 observed by PiMS is consistent with this. A majority of the scattered vimentin spots observed in
 227 the cortex fit well into the dark spots corresponding to glomeruli. Moreover, vimentin is found in
 228 many scattered locations in the inner medulla region. These illuminated spots correspond to the
 229 spatially-dispersed peritubular capillaries, which form a complex three-dimensional network in the
 230 medulla and become dispersed when the tissue is cross-sectioned.

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

239 Three major proteoforms of N-terminal acetylated GST subunits, GSTA1 and GSTA2, were
 240 observed in kidney tissue with localizations to the cortex region (Fig. 5B, 5D, and 5E). Fig. 5C
 241 shows the mass domain spectrum of GSTA1 and GSTA2 proteoforms. A single proteoform was
 242 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 the 14 Da mass shift to the region between Pro110 and Gln113 from the N-terminus (regions highlighted in light red, Fig. 5D and 5E 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.

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 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 protein isoforms from different gene family members or allelic cSNPs in biological tissues.

PiMS presents low-bias to the protein masses compared to multiplexed 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 opportunity of using PiMS to delve into the major biological pathways in tissues in a spatially-resolved manner. In the GO analysis result shown in Fig. 2F, a variety of central

cellular metabolic pathways were found enriched by 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 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 (45, 46), the dimethylation of Lys65 was shown to be very low in abundance and inconsistently observed (46). Our data indicates that the dimethylated proteoform is about ~5% of the total GAPDH in 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 (47).

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 up to ~70 kDa. This approach increases observable proteoform masses by nearly 4-fold and resolving power by 10-fold compared to prior work (48). The development of PiMS opens up exciting opportunities to infuse proteoform knowledge into the multi-omic approaches being evaluated for inclusion into the Human Reference Atlas (49). 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

293 *Tissue Preparation*

294 Human kidney tissue sections were prepared according to published protocols (50). Mouse
295 brain tissue sections were sectioned at -21°C to a $12\text{ }\mu\text{m}$ thickness using a CM1850 Cryostat (Leica
296 Microsystems, Wetzlar, Germany). Tissue sections were thaw mounted onto glass microscope
297 slides (IMEB, Inc Tek-Select Gold Series Microscope Slides, Clear Glass, Positive Charged) and
298 stored at -80°C before mass spectrometry imaging analysis.

299 Human kidney tissue sections were thawed under slight vacuum at room temperature, fixed
300 and desalted via successive immersion in 70%/30%, 90%/10%, and 100%/0% ethanol/ H_2O
301 solutions for 20 s each, delipidated by 99.8% chloroform for 60s, and dried under slight vacuum
302 right before nano-DESI imaging experiments. These sample preparation steps allow for the *in situ*
303 precipitation of proteins and removal of lipids avoiding suppression of protein signals upon
304 ionization into the mass spectrometer (19, 51).

305 *Nano-DESI Ion Source*

306 A custom-designed nano-DESI source was used for all data acquisition. The experimental
307 details of nano-DESI MSI have been described elsewhere (29, 30). Briefly, the nano-DESI probe
308 is comprised of a primary (OD $150\text{ }\mu\text{m}$, ID $20\text{ }\mu\text{m}$) and a nanospray capillary (OD $150\text{ }\mu\text{m}$, ID $40\text{ }\mu\text{m}$)
309 with the spray side of the nanospray capillary positioned close to the MS inlet. The probe was
310 fabricated using fused silica capillary tubing (Molex, Thief River Falls, MN). A liquid bridge
311 formed at the location where the two capillaries meet is brought into contact with the tissue section
312 for analyte extraction. The liquid bridge is dynamically maintained by solvent propulsion from the
313 primary capillary and instantaneous vacuum aspiration through the nanospray capillary. The
314 extracted analytes are continuously transferred to a mass spectrometer inlet and ionized by ESI.
315 Imaging experiments are performed by moving the sample under the nano-DESI probe in lines. The
316 optimal scan rate is discussed in the next section. The strip step between the line scans was set to
317 $150\text{ }\mu\text{m}$ 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 (52). 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 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) (48). 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 °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 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 $\mu\text{m/s}$, which corresponds to a 5-8 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 prior to an imaging experiment, and it may vary from 100 to 500 ms (fig. S4-fig. S6). 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.

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 substantial losses in trapping efficiency. 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 (53).

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 (53). 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 $\sim 10 \mu\text{m} \times 150 \mu\text{m}$ 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 MS¹ 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.

GO analysis was performed using Metascape (<https://metascape.org/>) (54). 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.

391 *Top-Down Proteomics Data Acquisition & Analysis*

392 Targeted MS/MS experiments were performed on a tissue section adjacent to the imaged
 393 section using higher-energy collisional dissociation (HCD). In the first step, a target proteoform
 394 was selected in the mass-domain spectrum. We utilized the PiMS image of the target proteoform to
 395 select a target area on the section where the proteoform abundance is enhanced. For the selected
 396 area, the mass-domain PiMS spectrum was convoluted back to m/z domain, from which a proper
 397 isolation window that contains predominantly the target proteoform was selected. A 0.8 m/z
 398 isolation window was typically employed for most of the targets; in special cases, 0.5-0.6 m/z
 399 window was used to avoid overlapping signal. MS/MS experiments were performed by scanning
 400 the nano-DESI probe over the selected region with the selected isolation window at 2.5-4 $\mu\text{m/s}$ scan
 401 rate. MS/MS data acquisition was conducted in the I²MS mode with an Orbitrap detection period
 402 of 2 s (HCD pressure setting = 0.5) (41). HCD collision energy and injection time was optimized
 403 to maximize the population of individual ion fragments. Typical ranges of collision energy and
 404 injection time used in this study were 7-14 eV and 200-1500 ms, respectively. Total data acquisition
 405 time for each target varied from 1-5 hours.

406 MS/MS data was first subjected to I²MS processing for fragment ion charge assignment and
 407 mass-domain spectrum construction following the same procedure as described above. Mass-
 408 domain spectrum was converted to *.mzML* format subjected to MS² search function implemented
 409 in ProSight Native (Proteinaceous, Evanston, IL) to look for possible candidates from the entire
 410 human protein database. For each search, the top 1-5 candidates were manually validated using a
 411 custom version of TDValidator (Proteinaceous, Evanston, IL) to identify the best matching
 412 proteoform. Proteoform E-values were obtained from ProSight Native and TD Validator reports.

413 **Acknowledgments**

414 **Funding:**

415 National Institutes of Health UH3 CA246635 (NLK)

416 National Institutes of Health P41 GM108569 (NLK)

417 National Institutes of Health P30 DA018310 (NLK)

418 National Institutes of Health P30 CA060553 (awarded to the Robert H. Lurie

419 Comprehensive Cancer Center)

420 National Institutes of Health UH3CA255132 (JL)

421 National Institutes of Health U54DK120058 (JMS)

422 National Institute of Environmental Health Sciences T32ES007028 (EKN).

423 NIH National Cancer Institute 5 UM1 CA183727-08 (Cooperative Human Tissue

424 Network at Vanderbilt University Medical Center).

425 **Author contributions:**

426 Conceptualization: PS, JMC, JOK, NLK

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

428 Resources: MY, EKN, JLA

429 Software: KRD, MARH, JBG, BPE, RTF

430 Investigation: PS, EKN, FAB

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

432 Supervision: JMC, JOK, NLK

433 Writing—original draft: PS, JMC, NLK

434 Writing—review & editing: PS, JPM, KRB, MARH, MY, EKN, JLA, BSD, FAB,

435 JBG, BPE, RTF, JMC, JOK, JMS, JL, NLK

436 **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.

Data and materials availability:

Custom compiled code used to process and create I²MS files is already available (39). 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 16500 MS scans; regions occupied by abundant blood

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

468
 469

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, d), glutathione S-transferase A2 (red, e), and albumin (green, f); (l) glutathione S-transferase A2 (blue, e), alpha-crystallin B chain (red, g), and Transgelin-2 (green, h); and (m) glutathione S-transferase A2 (blue, e), alpha-crystallin B chain (red, g), and vimentin (green, j). Scale bar: 1 mm.

494

495

496

497

498

499

500

501

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 GFM 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 red.

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

Protein description	Gene	Accession	PFR	Observed mass	Theoretical mass
Alpha-crystallin B chain	CRYAB	P02511	PFR00000325920	20200.47	20200.46
Superoxide dismutase [Mn], mitochondrial	SOD2	P04179-1	PFR00000001332	22204.25	22204.27
Membrane-spanning 4-domains subfamily A member 4A	M4A4A	Q96JQ5	N/A	25431.57	25431.5
Glutathione S-transferase A1	GSTA1	P08263-1	N/A	25540.66	25540.67
Glutathione S-transferase A2	GSTA2	P09210	PFR00000432952	25573.626	25573.63
Glutathione S-transferase A2	GSTA2	P09210	N/A	25587.64	25587.64
V-type proton ATPase subunit E 1	VATE1	P36543-1	PFR00000053332	26055.743	26055.77
Apolipoprotein A-1	APOA1	P02647	N/A	28078.5	28078.53
Adaptin ear-binding coat-associated protein 1	NECAP	Q8NC96-1	N/A	29764.03	29764.21
N(G),N(G)-dimethylarginine dimethylaminohydrolase 1	DDAH1	O94760	PFR00000058601	31031.95	31031.95
Malate dehydrogenase, mitochondrial	MDHM	P40926-1	PFR00000002776	32999.46	32999.49
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001134	35921.4	35921.41
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406-1	PFR00000001124	35950.39	35950.07
Fructose-1,6-bisphosphatase 1	FBP1	P09467	PFR000006225388	36724.76	36724.82
Actin, cytoplasmic 2	ACTG1	P63261	PFR00000001511	41660.74	41660.82
Actin, cytoplasmic 1	ACTB	P60709	N/A	41660.74	41660.78
Actin, cytoplasmic 2	ACTG1	P63261	N/A	41716.67	41716.84
Alpha-enolase	ENOA	P06733	PFR00000001065	47078.34	47078.36
Vimentin	VIME	P08670	PFR00000001800	53561.13	53561.13
Albumin	ALBU	P02768	N/A	66438.68	66438.78
Mesothelin	MSLN	Q13421-3	N/A	70900.52	67980.42

Mass accuracy, ppm	Sequence coverage	E-value	Modifications
0.50	35%	1.06E-131	Acetyl@N
0.90	16%	2.08E-58	Truncation 1-24
2.75	4%	2.66E-12	Met-off, acetyl@N, phosphoryl@S
0.39	15%	2.50E-75	Met-off, acetyl@N
0.16	17%	1.38E-57	Met-off, acetyl@N
0.00	20%	8.64E-87	Met-off, acetyl@N, 114 S to T
1.04	14%	4.16E-54	2-225, alpha-amino acetyl@N
1.07	17%	1.70E-65	Truncation 24-266
6.05	18%	3.52E-75	Omega-N-methyl@R180, methyl@N
0.00	16%	1.28E-74	Met-off, acetyl@N
0.91	18%	6.08E-94	Truncation 1-24
0.28	14%	4.48E-76	Met-off
8.90	14%	8.64E-63	Met-off, dimethyl@K65
1.63	22%	1.12E-121	Met-off, acetyl@N, 217 R to K
1.92	24%	1.12E-154	Met-off
0.96	24%	3.84E-160	Met-off, acetyl@N, methyl@H72
4.08	20%	1.73E-115	Met-off, acetyl@N, methyl@H72/K83
0.42	12%	1.60E-74	Met-off, acetyl@N
0.00	16%	4.80E-106	Met-off, acetyl@N
1.51	3%	8.32E-22	Truncation 1-24, 16 disulfide bonds, phosphoryl, acetyl, O@M
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. Sahasrabudhe, 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. Wohlschläger, 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. Unsuhay, 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. Salzert, 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. Wührer, 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. 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. *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. 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).
30. 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).
31. R. Yin, J. Kyle, K. Burnum-Johnson, K. J. Bloodsworth, L. Sussel, C. Ansong, J. Laskin, High Spatial Resolution Imaging of Mouse Pancreatic Islets Using Nanospray Desorption Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **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. *Journal of the American Society for Mass Spectrometry* **32**, 653-660 (2021).
33. 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. *ChemRxiv*, (2021).
34. O. J. Hale, H. J. Cooper, Native Mass Spectrometry Imaging of Proteins and Protein Complexes by Nano-DESI. *Analytical chemistry* **93**, 4619-4627 (2021).
35. 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).
36. R. A. Zubarev, A. Makarov, Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288-5296 (2013).
37. 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).
38. 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. *Analytical Chemistry* **91**, 2776-2783 (2019).

39. 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).
40. 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).
41. 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).
42. 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).
43. 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).
44. 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).
45. 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).
46. 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).
47. 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).
48. 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).
49. 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).
50. 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).
51. 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).
52. 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).
53. 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. *Journal of the American Society for Mass Spectrometry* **30**, 2200-2203 (2019).
54. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner, S. K. Chanda, Metascope provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* **10**, 1523 (2019).