Published in final edited form as:

J Proteome Res. 2018 January 05; 17(1): 635–646. doi:10.1021/acs.jproteome.7b00713.

Mapping Extracellular Matrix Proteins in Formalin-Fixed, Paraffin-embedded Tissues by MALDI Imaging Mass Spectrometry

Peggi M. Angel[1], **Susana Comte-Walters**[1], **Lauren E. Ball**[1], **Kacey Talbot**[1], **Anand Mehta**[1], **Kelvin G.M Brockbank**[2,3,4], and **Richard R. Drake**[1]

[1]Department of Cell and Molecular Pharmacology, Medical University of South Carolina, Charleston, SC

[2]Tissue Testing Technologies LLC, North Charleston, SC, USA

[3]Department of Bioengineering, Clemson University, Clemson, SC, USA

[4]Department of Regenerative Medicine, Medical University of South Carolina, Charleston, SC USA

Abstract

Collagens and elastin form the fundamental framework of all tissues and organs and their expression and post-translational processing are tightly regulated in disease and health. Due to their unique structural composition and properties, it is a recognized challenge to access these protein structures within the complex tissue microenvironment to understand how localized changes modulate tissue health. We describe a new workflow using a combination of matrix-assisted laser desorption /ionization imaging mass spectrometry (MALDI IMS) with collagenase and matrix metalloproteinase (MMP) enzymes to access and report on spatial localization of collagen and elastin sequences in FFPE tissues. The developed technology provides new access to collagens and elastin sequences localized to tissue features that were previously unattainable. This high throughput technological advance should be applicable to any tissue regardless of disease type, tissue origin, or disease status and is thus relevant to all research, basic, translational or clinical.

Keywords

imaging mass spectrometry; proteomics; peptide imaging; extracellular matrix; formalin-fixed; paraffin-embedded tissue imaging; tissue imaging; MALDI imaging mass spectrometry

Introduction

Collagens and elastin comprise the majority of extracellular matrix surrounding the eukaryotic cell, and form the basic components of all tissue scaffolding.(1, 2) A main

Corresponding Author: Peggi M. Angel, 173 Ashley Ave, BSB358, Charleston, SC 29425, angelp@musc.edu Phone: (843) 792-8410 Fax: (843) 792-0481.

function of these proteins is to provide both physiological and biomechanical support for tissue function, facilitated through matrix-matrix and cell-matrix interactions. Multiple biochemical changes to collagens and elastin influence quaternary and suprastructural architecture to direct the biological status of the tissue. Normal tissue turnover, remodeling and generation/regeneration are driven by matrix metalloproteinases (MMPs); these proteases specifically target collagens and elastin for degradation processes, producing expression patterns that vary qualitatively, e.g., collagen type, or quantitatively, by up or down regulation.(3, 4) A recognized challenge is identifying ways to spatially and quantitatively study the complexities of the biochemical changes in collagens and elastin that modulate tissue health; such technological advances are anticipated to contribute to new therapeutic routes.(3–5)

Collagens and elastin are complex polymeric protein structures characterized by multiple post-translational modifications (PTMs), crosslinking and suprastructural organization. Collagen proteins have at least one triple helical structure with a repeating unit of Gly-X-Y where X and Y are most frequently proline (28%) or 4-hydroxyproline (Hyp, 38%); these triple helices are organized as suprastructural networks through covalent crosslinking interactions involving disulfide bonds, lysyl-mediated crosslinks with lysine, hydroxyl lysines (HLys), and histidine.(6, 7) In vertebrates, there are 28 subtypes of collagens that are additionally modified by numerous PTMs involved in the synthesis, maturation and processing of these proteins. The major PTMs include hydroxylation of proline and lysine, N- and O-linked glycosylation, glycation, and sulfation of Tyr. (6, 8) Elastin, in contrast, is a single gene-copy protein with several proteoforms characterized in disease by multiple genetic insertions, mutations, and deletions.(9–11) Elastin is also heavily modified by PTMs including hydroxylation of lysine, allysine, hydroxylation of proline and glycation.(12) Although collagens and elastin are critical proteins of organogenesis, remodeling and repair, and disease, the abundant PTMs, genetic alterations and suprastructural arrangements through crosslinking have created significant challenges in defining the role of these changes within the tissue microenvironment.

Both antibody staining and conventional mass spectrometry-based proteomic analysis workflows, e.g., GeLC-MS(13) and Multidimensional Protein Identification Technology (MudPIT)(14) have been limited in the ability to interpret collagens and elastin sequence variation within the tissue microenvironment. Antibody staining can report localization of these target ECM proteins, but requires knowledge of an epitope and frequently very specific structural conformations.(15) This limits researchers to only studying epitopes for which there are available antibodies, and cannot account for variability in crosslinking or PTMs due to changes in biological status. Conventional proteomic methods using enzymes for sequencing collagens and elastin are limited due to the combination of PTMs and suprastructural organization. Common enzymes used in proteomic experiments, including trypsin, LysC, and ArgC, target arginines and lysines that are heavily modified in collagens and elastin and therefore limit proteolysis.(16) Suprastructural organization limits proteolytic access to cleavage sites that may not be modified; for collagens some have used cyanobromide to access peptides.(17) In addition, collagens and elastin are largely insoluble and most proteomic workflows involve homogenization that leaves behind collagens and elastin as insoluble fractions. Thus, antibodies are limited in reporting *in situ* biochemical

and structural modulation of collagens and elastin, while conventional proteomic workflows are limited in the ability to produce suitable targets for MS analysis.

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI IMS) is an imaging modality that uses mass spectrometry to map localization of proteins from thin histological tissue sections regardless of PTM or mutation.(18) In a workflow parallel to LC-MS/MS proteomic strategies, trypsin is sprayed onto the tissue to release tryptic peptides which remain localized on the tissue.(19, 20) After spraying the tissue with a chemical matrix that facilitates peptide ionization, a laser beam is used to desorb and ionize peptides from discrete locations on the tissue; ionized peptides are subsequently recorded by a mass spectrometer, resulting in thousands of peptide mass spectra from a single tissue section. Data is then assembled into heat maps across the tissue of peptide aligned to histological features. The result of a single MALDI IMS experiment is immediate visualization of hundreds to thousands of peptides localized to the tissue architecture. Currently, MALDI IMS uses trypsin to access peptide content, with LC-MS/MS proteomic experiments to identify peptides. This untargeted approach leaves behind most collagen and elastin sequences for the same reasons as LC-MS/MS based proteomics.

Here, we describe a MALDI IMS approach to localize collagen and elastin peptides within the tissue microenvironment using enzymes other than trypsin. Matrix metalloproteinase (MMPs) are enzymes that degrade major ECM components and are produced by many cell types including fibroblasts, inflammatory cells, and epithelial cells.(21) Most MMPs have high collagenase activity but have been reported to act on other ECM proteins such as fibronectin, fibrinogen, laminins, perlecan and vitronectin.(2) MMP12 (elastase) is an exception with activity directed primarily at elastin with very limited activity on aggrecan, fibronectin, osteonectin, laminin and nidogen.(2, 22) Mammalian MMPs have very specific collagenase activity to collagen sub-types, while bacterial MMPs have broader specificity to collagens and are referred to as Collagenase Types (COLases). COLases are available commercially in large scale and are used for viable cell isolation from all tissues and species to proteolytically degrade connective tissue. In the current study, we leverage COLases and MMP12 to digest collagens and elastin on thin tissue sections and detect products of the digestion with MALDI IMS. On similarly prepared adjacent tissue slides, LC-MS/MS proteomic methods were used to report sequence information from COLase/MMP peptides. The method allows *in situ* reporting ECM sequences localized to tissue features, yet 1) without prior knowledge of a specific epitope, 2) with increased sequence information compared to most antibody epitopes, and 3) unbiased as to tissue type. This method allows, for the first time, direct detection of localized biochemical information related to extracellular matrix collagens and elastin proteoforms and should be applicable to any mammalian tissue, organ, or system during health or disease.

Experimental

Materials

Acetonitrile, α-Cyano-4-hydroxycinnamic acid (CHCA), trifluoracetic acid (TFA), sodium citrate, and trizma base were purchased from Sigma-Aldrich (St. Louis, MO, USA). Collagenase type III (COLase3) (*C. histolyticum*) and IV and elastase (MMP12; porcine)

were purchased from Worthington Biochemicals (Lakewood, NJ), StemCell Technologies (Cambridge, MA), Thermo Scientific (Waltham, MA), Alfa Aesar (MMP12, Fisher Scientific, Pittsburgh, PA) or Abnova (Taipei City, Taiwan). Xylenes, 200 proof ethanol, methanol, citraconic anhydride were purchased from Fisher Scientific (Pittsburgh, PA).

Tissue Procurement

Tissue use was in accordance with protocols approved by the Medical University of South Carolina Institutional Review Board. Tissues were procured from the Biorepository & Tissue Analysis Shared Resource at the Hollings Cancer Center, Medical University of South Carolina. Tissues were human liver samples with hepatocellular carcinoma (HCC), colorectal cancer (CRC), and human aortic and aortic valve tissue. Porcine heart valve tissue was collected as fresh discarded tissues from large animal surgical training and are approved for research use by MUSC Institutional Animal Care and Use Committee.

Tissue Preparation

Formalin-fixed, paraffin-embedded tissue was sectioned at a 5 μm thickness, mounted onto standard microscope slides (Tissue Tack, Poly Sciences, Inc, Warrinton, PA) dehydrated overnight at 37°C, and stored in a cool, dry location prior to preparation. For imaging experiments, slides were heated at 60°C for one hour and dewaxed using xylenes (incubated 3 minutes in two fresh aliquots), 200 proof ethanol (incubated 1 minute in two fresh aliquots), 95% ethanol (incubated 1 minute), 70% ethanol (incubated 1 minute), and distilled water (incubated 3 minutes in two fresh aliquots). Heat induced epitope retrieval (HIER) was performed using a vegetable steamer (Rival Model CKRVSTLM2) maintaining a temperature of 95°C for 25 minutes. HIER buffers tested included 10 mM citraconic anhydride buffer, pH 3, 10 mM sodium citrate, pH 6; and 10 mM tris, pH 9. Enzymes were prepared in 25 mM ammonium bicarbonate and sprayed onto tissue at tested concentrations of 0.01, 0.1 and 0.5 μg/μL. Enzymes were sprayed with an automated sprayer (M3 TM-Sprayer, HTX-Imaging, Chapel Hill, NC) using parameters of 45°C, 10 psi, 25 μL/min, 1200 velocity, and 15 passes with a 2.5 mm offset. Cell culture dishes were prepared as humidity chambers for digestion using Wypall X 60 paper towels and two folded Kimwipes saturated with 5mL distilled water and pre-heated at 37°C for 15 minutes prior to placing samples in the chamber. Samples were incubated in the closed cell culture “humidity chambers” at 37.5 °C for 2 or 4 hours or overnight. Samples were removed and sprayed with CHCA matrix prepared as 7 mg/mL in 50% acetonitrile, 1% TFA. CHCA was sprayed onto tissues with the same automated sprayer (M3 TM-Sprayer, HTX-Imaging, Chapel Hill, NC) using parameters of 77°C, 10 psi, 100 μL/min, 1300 velocity, and 10 passes with a 3.0 mm offset. Coated slides were dipped for 1 second in cold 5 mM ammonium phosphate (monobasic), to reduce matrix clusters (23), and immediately dried in a desiccator.

Imaging Mass Spectrometry

Samples were analyzed by MALDI-FT-ICR (solariX Legacy 7.0 Tesla, Bruker Daltonics) in positive ion mode, collecting 300 laser shots per pixel using the SmartWalk feature to raster the laser in a 25 μm diameter. Transients of 1 megaword were acquired in broadband mode over m/z 700–5000, with a calculated on-tissue mass resolution at full width half maximum of 81,000 at m/z 1400. Data were visualized in fleximaging 4.0 (Bruker Daltonics, Bremen,

Germany) and analyzed by SCiLS Lab software 2017a (SCiLS Lab software, Bruker Daltonics, Bremen, Germany). All images are shown normalized to total ion current.

Staining and Immunohistochemistry

Tissues were stained with hematoxylin (Gill 2, Cancer Diagnostics, Durham, NC) and eosin (Cancer Diagnostics, Durham, NC) or Verhoeff Van Geison (American MasterTech, Lodi, CA) following manufacturer's protocols. Immunohistrochemistry utilized the DAKO Envision Chromophore Kits for mouse (K4006) and rabbit (K4010) following manufacturer's instructions. Primary antibodies for collagen 1A1 (Santa Cruz, SC-293182) and collagen 3A1 (Novus Biologicals, NB600–594) were applied at 1:100 ratios and incubated overnight at 4°C. The tissue overview was visualized using a high resolution document scanner (Epson America, Long Beach, CA).

Proteomics

Tissues were heated for 1 hr at 60°C, dewaxed and scraped off the slides into microcentrifuge tubes. Tissues were digested using collagenase type III or elastase (2 μg enzyme per tissue section, freshly prepared) at pH 7.2 in 25 mM ammonium bicarbonate with 3 mM calcium chloride. After 5 hours digestion, a new aliquot of enzyme was added to the solution and digestion continued overnight. For elastin from human aorta (E677, Sigma St Louis, MO), 100 μg of elastin was dissolved in 25 mM ammonium bicarbonate pH 7.2, and digested by 3 μg of elastase. A total of 4 μg peptides from each of the preparations were dried down, cleaned up by solid phase extraction using a C18 Ziptip (EMD Millipore, Damkstadt, Germany) following the manufacturer's protocol. Peptides were loaded onto a C18 PepMap trap column and separated on a. Peptides were loaded onto a C18 PepMap trap column and separated on a 75 μm × 30 cm fused-silica column (Reprosil-Pur 120, C18-AQ, 1.9 μm from Dr. Maisch GmbH, packed in house) thermostatted to 60°C, at a flow rate of 200 nl/min. The gradient was from 5 to 40% solvent B (98% acetonitrile and 0.2% formic acid) over 180 min, where solvent A was 2% acetonitrile, 0.2% formic acid. Peptides were analyzed by data dependent acquisition on an Orbitrap Elite mass spectrometer (Thermo Scientific) with a single FTMS survey scan acquired in the orbitrap followed by collision-induced dissociation MS/MS of the top 10 most intense ions in the ion trap. The survey scan was acquired over a mass range of m/z 400–1700 at a resolution of 60,000 at m/z 400 with an automatic gain control (AGC) target value of 10^6. Ions with an unassigned charge state were not selected for sequencing. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 30 sec, and exclusion duration of 180 sec. Tandem mass spectra were searched using both MASCOT (Version 2.4.01) and SEQUEST HT via Proteome Discoverer 1.4 against a subset of human protein sequences downloaded on May 7, 2017 from UniProtKB (SwissProt) containing 1,783 entries (keywords used: collagen, elastin, aggrecan, gelatin, osteonectin, perlecan, plasminogen, and fibronectin). Search parameters were unspecified enzyme, 2 missed cleavages, precursor mass tolerances of ± 20 ppm, and fragment mass tolerances ± 0.8 Da. Methionine oxidation, lysine hydroxylation, asparagine and glutamine deamidation were used as variable modifications. Data were uploaded into Scaffold (v4.8.1) and a peptide probability of 99% was used to report peptide identifications. For elastin, which has lower basic peptide content than collagens, singly charged peptides

from elastin were examined after filtered by scoring to those with Xcorr greater than 1.5 and delta CN of 0.1.

ProteinProspector MS-Fit (Baker, P.R. and Clauser, K.R. http://prospector.ucsf.edu) was used for peptide mass fingerprinting coupled to image alignment for select proteins identified by proteomics. Collagen lalphal (Col1A1, Uniprot P02452), Collagen 3alpha1 (Col3A1; Uniprot P02461), and elastin (P15502) were evaluated by *in silico* enzymatic digestion C-terminal to AILSTV and 1 missed cleavage to examine subsets of isobaric and near isobaric peptides for comparative purposes. For COL1A1 and COL3A1, the sequence was randomized using the Expasy RandSeq tool (http://web.expasy.org/randseq/) against the exact amino acid composition reported by Uniprot. *In silico* tryptic digestion against the forward sequences was used as a second reference for enzyme specificity. For elastin (ELN, Uniprot P15502), *in silico* digestion using enzymatic digestion elastase included potential modifications due crosslinking (uncleaved K). Peptide mass fingerprints were aligned with image data using a simple copy paste function in SCiLS Lab software. For PMF in image data, an image "hit" was considered positive if 1) a pattern was detectable, and 2) treated compared to untreated ratio was above 10. In most cases, a detectable image pattern yielded a treated/untreated ratio of 100 or greater.

Results

Overview of MALDI IMS method for mapping collagens and elastin

An overview of the optimized workflow using either COLase3 or MMP12 is shown in Figure 1. COLase3 was chosen as a focus due to prior studies in isolation and maintenance of pluripotent cell phenotypes from tissue including heart valve and stem cells, indicating that the enzyme has limited activity against cellular proteins. (24, 25) Hepatocellular carcinoma (HCC) tissue was used in initial phases of method development due to known and significant increases in many collagen types.(26, 27) COLase3 reportedly has higher proteolytic activity, and comparison of COLase3 to Collagen Type IV (COLase4) showed decreased peak intensities when comparing COLase4 to COLAse3 (Supplemental Figure 1). To test MMPs as an enzyme for MALDI IMS, we used MMP12. Porcine MMP12 is ideal for IMS experiments, as it is already obtainable in large quantities, has a rate constant 20 fold higher than human derived leukocyte elastase, is less sensitive to inhibition by chemicals, and is used routinely in dissociation of human tissues.(28) Optimization of activity for each enzyme for IMS was performed by investigating antigen retrieval conditions, applied concentrations, and digestion times. Thus far, the overall approach for IMS appears to be identical between the three tested enzymes.

Collagen mapping by MALDI IMS using Collagenase Type III (COLase3)

Figure 2 shows example results of the imaging experiment with significant signal achieved after optimization of COLase3 treatment across tissue types. Typically, it was observed that the total ion current for the non-collagenase treated sections showed intense matrix clusters which were reduced upon enzyme treatment. Furthermore, enzyme treatment produced multiple peaks spanning m/z 800–4,000 (Figure 2A). Distinct and intense images were produced from COLase3 treated tissue that were not observed in identically treated control

tissue where the enzyme was not applied (Figure 2B). Some images showed homogenous expression throughout the tissue, while many appeared localized to areas high in collagen as determined by a general collagen stain. Additional work showed that COLase3 also produced high intensity and distinctively localized images in a different HCC tissue, in CRC tissue and in porcine cardiac tissue (Figure 3). While many peaks were reproducible across different tissues, some peaks were detected only in certain tissue types, potentially due to collagen differences among tissue types. For example, m/z 860.4275 and m/z 1208.6558 mapped only to human HCC or CRC tissue, while m/z 1148.6577 maps only to CRC tissue and not HCC or pig cardiac tissue. Overall, the image data show that the optimized method for COLase3 generates high intensity image data only in enzyme treated tissue and suggests that the method will be applicable across different tissues and different species.

LC-MS/MS-Based Identification of Peptides Generated by Collagenase Type III (COLase3)

FFPE liver and colon cancer sections adjacent to those sections used for MALDI IMS were heated, dewaxed and scraped off the slide for in-solution digestion with COLAse3. The resulting peptides were analyzed by high resolution/accurate mass nLC-MS/MS yielding 127 peptides from 17 proteins when searched against either the subset protein or the human database. A total of 118 peptides (92%) (Supplemental Table 1) mapped to ECM proteins with 58 peptides (46%) from 10 collagens (COL1A1, COL1A2, COL3A1, COL4A2, COL5A1, COL6A1–2, COL14A1, COL18A1). The remaining 46% of the peptides from ECM proteins included fibronectin, vimentin, Emilin-1, tenascin-X. COLase3 digest of FFPE tissue thus results a high enrichment of collagens as well as ECM proteins with collagenous domains or collagen-binding activity. The identification of additional ECM proteins was expected as collagenases and MMPs are also active against other ECM proteins.(2) To directly compare COLase3 with trypsin, we performed in-solution digestion on an adjacent HCC tissue section. Proteomic analysis of this digestion yielded 454 proteins with three collagens and 18/3,617 peptides matching to collagens (0.49%). Searching these data against the smaller collagen protein database also identified three collagens represented by 13/249 spectra (5.2%), further demonstrating the limitations of on-tissue trypsin digestion of collagens. Evaluation of the identified peptide MS/MS by COLase3 digestion revealed that the most common C-terminal amino acid was arginine, likely due to more efficient ionization for peptides terminating in arginine (Supplemental Figure 2). The most common penultimate amino acid was glycine, expected due to high glycine content in collagen sequences. It is anticipated other fragmentation methods will improve the identification of peptides deficient in lysine, arginine, and histidine. Figure 4 shows example IMS data alignment with peptides found by HR/AM LC-MS/MS proteomics performed on collagenase digested tissue and illustrates localization of distinct sequences in the tissue. To summarize, the proteomics data confirms that COLase3 digestion on FFPE tissue sections enriches for collagen peptides, peptide MS/MS are interpretable, and MALDI IMS images can be matched to the proteomics data.

Peptide Mass Fingerprinting of Collagen Type III Peptides on Tissue

To further understand COLase3 activity on FFPE tissue sections, peptide mass fingerprinting (PMF) experiments were used to examine image data for collagen sequence coverage on collagens identified by HR/AM LC-MS/MS. The proteins COL1A1 and COL3A1 from

proteomics data were used as target protein candidates for PMF experiments (Figure 5). *In silico* digestion performed against neutral amino acids per collagenase type activity showed sequence coverage of COL1A1 at 30.9 ± 3.0% over three independent tissue imaging experiments of adjacent tissues sections (Data in Supplemental Table 2). An *in silico* tryptic digestion on the forward COL1A1 sequence resulted in 13.7 ± 2.6% sequence coverage. Randomizing the COL1A1 sequence and performing *in silico* digestion resulted in 9.9 ± 6.3% sequence coverage. Similar experiments were performed for COL3A1 resulted in 40.4 ± 6.5% sequence coverage by the image data (Data in Supplemental Table 3) with 10.9 ± 3.9% sequence coverage by *in silico* tryptic digestion on the forward COL3A1. Sequence coverage for randomized COL3A1 was 12.5 ± 3.7% and was not statistically different than trypsin digestion. Antibody staining for target collagens showed good correlation with most MALDI IMS image data, with the known caveat that 1) antibody staining may be altered by collagen suprastructure and PTMs and 2) there is potential the PMF peptides could be isobaric with other proteins (example images in Supplemental Figure 3). Low levels of staining and low intensity PMF image data were both observed for COL1A1 close to the tumor (Figure 5B), while Intense antibody staining and PMF image data were observed for COL 3A1 throughout the tissue with the exception of the tumor region (Figure 5C). PMF experiments thus show that the image data results in higher sequence coverage than randomized sequences or alternate enzyme activity and appears to map peptide sequences for target proteins to similar regions as detected by antibody staining.

MMP12 Mapping of Elastin Peptides by MALDI IMS

MMP12 activity was explored on human and porcine cardiac tissue using optimized protocols (Figure 6). Peptide mass fingerprinting on porcine cardiac tissue identified 34 peptides representing 31.3% coverage (Supplemental Table 4) with putative elastin peptides mapped to elastin rich areas. MMP12 applied to human aortic valve tissue mapped a total of 18 putative peptides representing 16.9% sequence coverage by PMF comparing to human elastin (Supplemental Table 5). PMF peptides correlated to regions of high elastin staining (Figure 7). In contrast, *in silico* digestion of elastin by trypsin with up to 3 missed cleavages generated 1 image hit with 2% sequence coverage. Adjacent tissue sections were elastase-digested, but produced limited results likely due to the small sample size of the aortic tissue (~2×2 mm) in comparison to the larger tissues (~3×3 cm) used in COLase studies. Elastin powder from human aorta was therefore digested in solution by MMP12 and analyzed by HR/AM LC-MS/MS proteomics at comparable amounts to the collagen tissues. This analysis produced 246 elastin peptides with 59% coverage of elastin isoform 2 (Supplemental Table 6). In this analysis, a total of 33/246 (13%) of the peptides were singly charged with significant scores, which we believe reflects the low basic amino acid levels in elastin. Figure 7B illustrates examples of single and double charged peptides from elastin with interpretable MS/MS patterns but generally with more b-ions than y-ions. Comparing with the elastin data, the HCC and CRC collagenase data did not produce single charged, significantly scored peptides. Overall, the data illustrates that the use of MMP12 produces informative images for MALDI IMS unbiased towards species type with a potential for significant elastin coverage given the appropriate sample amount.

Discussion

Historically, collagens and elastin have been inaccessible to proteomics and proteomics imaging due to insolubility, multiple PTMs and suprastructural features that block tryptic access to these proteins. In the current study, we show that collagenases and MMPs may be used to access collagens, elastin and potentially other ECM proteins from formalin-fixed, paraffin-embedded tissues and map localization of the produced peptides within the tissue microenvironment by MALDI IMS. Sequence data may be obtained by standard LC-MS/MS proteomics methods, although there is room for improvements. Combined, the methodology enables reporting localization of distinct collagen or elastin sequences that are not detected with tryptic enzyme applications for MALDI IMS.

The use of bacterial collagenase types and non-human derived MMP in an IMS workflow is an advantage since these enzymes are not biased towards species and are cost effective for large research studies. We show that these enzymes produce intense and localized image data by MALDI IMS, applicable across species. It is possible that some variation in the method may be needed for different tissues types, e.g., longer digestion times. However, consideration should be given that longer digestion times may result in smaller peptide fragments that produce ambiguous peptide tags due to high homology between proteins. For COLases, the known activity against ECM proteins other than collagens further expands the ability to localize multiple sequence information from on-tissue digestion that is not currently accessible in a targeted manner for IMS. Additional work and likely the use of other mass spectrometric approaches such as ion mobility will be needed to improve image detection of these species. For COLAse3, comparison of PMF data with antibody staining showed good matches in patterns, but also some differences. This is likely due to known problems with antibody binding dependent on suprastructural organization and presence of PTMs(15), and suggests that MALDI IMS combined with COLases may be able to access previously unattainable structures. Sequence coverage showed high reproducibility when applied to serial sections of the same tissue, but peaks varied between tissue types and species, suggesting that the method will be able to detect biological variability in tissues. To summarize, besides the use of a different enzyme, the treatment of tissue and coating with matrix is consistent with other MALDI IMS protocols that use enzymes(29, 30) to treat FFPE tissue; no unusual steps need to be taken.

An important aspect of the method is the ability to obtain sequence information for the MALDI-IMS data. Here, we evaluated digests using HR/AM LC-MS/MS proteomics using methods customarily used for tryptic peptides. Proteomics of collagenase produced peptides from FFPE showed that COLase3 digestion produced peptides enriched for arginine C-termini. In contrast, previous work reported that COLase3 (*C. histolyticum*) generated peptides with C-terminal nonpolar amino acids.(31) However, in those workflows, N-termini were biotinylated for affinity isolation which would change ionization properties. An additional observation in the current study was that the adjacent amino acid to the C-terminal cleavage site was frequently glycine, which is not surprising given the high glycine content of collagens. This does correspond with the aforementioned study, which also found that adjacent amino acids were frequently glycine in a COLase3 digest.(31) As anticipated due to targeting a protein subpopulation, lower numbers of proteins were found in COLase3

and MMP12 digests from FFPE tissues when compared to tryptic digestion of a serial section. However, both COLase3 and MMP12 reported significantly higher enrichment from collagens and elastin compared to tryptic digestion. Overall, HR/AM LC-MS/MS demonstrated that COLAse3 and MMP12 digestion produced complementary proteomic signatures enriched with appropriate ECM proteins when compared to tryptic digestion of serial sections.

In the current work, we have shown that the use of collagenases and MMPs will allow mapping of target ECM proteins from tissue, previously unattainable using standard trypsin-based MALDI IMS and LC-MS/MS methods. As with any new approach, many topics remain to be explored. Analogous to early work obtaining sequence information from intact proteins mapped by MALDI IMS(32), the sequence identification strategy requires additional method development to increase sequence information from the image data. This is a challenge as the peptides produced by collagenase types and MMPs generate essentially non-basic peptides either through modification or due to low basic amino acid content. However, many proteomic approaches targeting unexpected modifications or base-less peptides are available that may be used in furthering proteomics of collagenase or elastase digestion.(33–35) Over many years of study, electron transfer dissociation was identified to increase sequence information for multiply charged, non-tryptic peptides and intact proteins from imaging studies.(36, 37) It is anticipated that a similar paradigm, with ongoing and focused development of identification strategies, will increase sequence information from collagenase types and MMP digestion. Both imaging mass spectrometry and LC-MS/MS proteomics approaches will benefit from continued parallel exploration of the methods. Expansion to explore topics such as additional tissue sources and tissue types, different sources and preparations of the enzymes, and strategies to access and define post translational modifications for imaging studies will result in significant advances accessing these proteins for biomedical research. An exciting finding was the proteomic reporting of additional but expected ECM proteins that shared collagenous domains, or are known to be bound to collagens, that further improvement on the MALDI IMS side will improve image detection of these species. Exploring the use of alternate mass spectrometry instruments, including ion mobility and time of flight, will increase our understanding of the robustness of the method towards detection of collagenase and MMP peptides as biomarkers. We have highlighted use of two collagenases, but there are many different collagenase types used in tissue dissociation that may prove advantageous in the imaging of different types of tissue. Using the workflow described herein, other collagenase type enzymes and MMPs alone or in combination with different enzymes may grant access to different protein sequences in tissue, furthering access to localized data.

The ability to access ECM proteins by imaging mass spectrometry holds great promise to improve our understanding of development and disease. There is great potential that imaging mass spectrometry strategies will allow much higher sequence coverage than single antibodies, and will ultimately be able to report localized mutations, deletions, insertions, and post-translational modifications. As shown here among various tissue types from colon to heart valve, MALDI IMS in combination with COLases and MMPs appears to be able to access these large, previously inaccessible ECM proteins regardless of suprastructure, a clear advantage over antibody detection. Clinical applications are plentiful as a lack of tools to

access ECM superstructure has limited studies on how collagens and elastin contribute to disease.(3–5) For example, although collagen expression and patterning have been known for decades to be associated with altered prognosis dependent on cancer type,(38–41) deciphering the role of collagen in cancer has remained elusive due to limited tools capable of describing chemical features and localization of specific collagen isoforms in the complex tumor setting. Here, our methods were developed on human FFPE tissue and can be immediately applied to the vast amounts of stored FFPE tissues associated with clinical data. In addition, image data on localized changes from collagens or elastin in the tissue microenvironment will be useful tools to allow “tuning” of biological scaffolds in tissue engineering, or evaluating preservation processes in organ banking.(42, 43)

Conclusion

In summary, we report a first approach to mapping ECM protein sequences by collagenase type and MMPs digestion on formalin-fixed, paraffin-embedded tissue sections. The use of bacterial collagenases and highly active porcine pancreatic MMP12 with imaging mass spectrometry should allow image data collection research on any type of organ tissue and any mammalian tissue, shown here on human and porcine tissues. Combined with bottom-up proteomics, the method is capable of detecting collagen types and elastin isoforms, with further development of identification strategies expected to identify additional biochemical content. This method presents, for the first time, a new tool that allows sequenced mapping of target ECM proteoforms within the tissue microenvironment useful for basic and clinical research on studies spanning developmental biology to disease therapeutics, prognosis and diagnosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Kim Norris-Caneda and Matt Lee for assistance spraying matrix on samples. The authors appreciate the discussion on LC-MS/MS proteomic topics with Jennifer Bethard. The Orbitrap Elite used in this research was S10 D010731 to LEB. PMA is grateful for funding provided by the National Institute of General Medical Sciences (P20 GM103542), National Cancer Institute (1R21CA207779), and the American Heart Association (16GRNT31380005). Additional support was provided by the South Carolina Centers of Economic Excellence SmartState program to RRD. The MUSC Mass Spectrometry Facility is supported by the Office of the Provost and the South Carolina COBRE in Oxidants, Redox Balance and Stress Signaling (GM 103542).

Abbreviations

ECM	extracellular matrix protein
PTM	post-translational modification
MALDI IMS	matrix assisted laser desorption /ionization imaging mass spectrometry
HCC	hepatocellular carcinoma

CrC	colorectal cancer
HR/AM	high resolution/accurate mass
LC-MS/MS	liquid chromatography coupled to tandem mass spectrometry

References

1. Alberts B; Johnson A; Lewis J; Raff M; Roberts K; Walter P, The Extracellular Matrix of Animals In Molecular Biology of the Cell, Garland Science: New York, 2002; Vol. 4th edition.
2. Lu P; Takai K; Weaver VM; Werb Z, Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harbor Perspectives in Biology 2011, 3, (12), a005058. [PubMed: 21917992]
3. Cox TR; Erler JT, Remodeling and homeostasis of the extracellular matrix: implications for fibrotic diseases and cancer. Disease Models & Mechanisms 2011, 4, (2), 165–178. [PubMed: 21324931]
4. Bonnans C; Chou J; Werb Z, Remodelling the extracellular matrix in development and disease. Nature Reviews Molecular Cell Biology 2014, 15, (12), 786–801. [PubMed: 25415508]
5. Swinehart IT; Badylak SF, Extracellular matrix bioscaffolds in tissue remodeling and morphogenesis. Developmental Dynamics 2016, 245, (3), 351–360. [PubMed: 26699796]
6. Shoulders MD; Raines RT, Collagen structure and stability. Annual Review of Biochemistry 2009, 78, 929–958.
7. Ricard-Blum S, The collagen family. Cold Spring Harbor Perspectives in Biology 2011, 3, (1), a004978. [PubMed: 21421911]
8. Myllyharju J; Kivirikko KI, Collagens, modifying enzymes and their mutations in humans, flies and worms. TRENDS in Genetics 2004, 20, (1), 33–43. [PubMed: 14698617]
9. Bashir MM; Indik Z; Yeh H; Ornstein-Goldstein N; Rosenbloom JC; Abrams W; Fazio M; Uitto J; Rosenbloom J, Characterization of the complete human elastin gene. Delineation of unusual features in the 5'-flanking region. Journal of Biological Chemistry 1989, 264, (15), 8887–8891. [PubMed: 2722804]
10. Sugitani H; Hirano E; Knutsen RH; Shifren A; Wagenseil JE; Ciliberto C; Kozel BA; Urban Z; Davis EC; Broekelmann TJ, Alternative splicing and tissue-specific elastin misassembly act as biological modifiers of human elastin gene frameshift mutations associated with dominant cutis laxa. Journal of Biological Chemistry 2012, 287, (26), 22055–22067. [PubMed: 22573328]
11. Callewaert B; Renard M; Hucthagowder V; Albrecht B; Hausser I; Blair E; Dias C; Albino A; Wachi H; Sato F, New insights into the pathogenesis of autosomal-dominant cutis laxa with report of five ELN mutations. Human Mutation 2011, 32, (4), 445–455. [PubMed: 21309044]
12. Gallop PM; Paz MA, Posttranslational protein modifications, with special attention to collagen and elastin. Physiological Reviews 1975, 55, (3), 418–487. [PubMed: 50603]
13. Shevchenko A; Tomas H; Havli J; Olsen JV; Mann M, In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols 2006, 1, (6), 2856–2860. [PubMed: 17406544]
14. Delahunty CM; Yates JR Iii, MudPIT: multidimensional protein identification technology. Biotechniques 2007, 43, (5), 563–567. [PubMed: 18072585]
15. Bordeaux J; Welsh AW; Agarwal S; Killiam E; Baquero MT; Hanna JA; Anagnostou VK; Rimm DL, Antibody validation. Biotechniques 2010, 48, (3), 197–209. [PubMed: 20359301]
16. Zhang Y; Fonslow BR; Shan B; Baek M-C; Yates JR Iii, Protein analysis by shotgun/bottom-up proteomics. Chemical Reviews 2013, 113, (4), 2343–2394. [PubMed: 23438204]
17. Barsh GS; Peterson KE; Byers PH, Peptide mapping of collagen chains using CNBr cleavage of proteins within polyacrylamide gels. Collagen and Related Research 1981, 1, (6), 543–548. [PubMed: 7346234]
18. Angel PM; Caprioli RM, Matrix-assisted laser desorption ionization imaging mass spectrometry: in situ molecular mapping. Biochemistry 2013, 52, (22), 3818–3828. [PubMed: 23259809]

19. Groseclose MR; Andersson M; Hardesty WM; Caprioli RM, Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry. Journal of Mass Spectrometry 2007, 42, (2), 254–262. [PubMed: 17230433]
20. Casadonte R; Kriegsmann M; Zweynert F; Friedrich K; Bretton G; Otto M; Deininger SO; Paape R; Belau E; Suckau D, Imaging mass spectrometry to discriminate breast from pancreatic cancer metastasis in formalin-fixed paraffin-embedded tissues. Proteomics 2014, 14, (7–8), 956–964. [PubMed: 24482424]
21. Rydlova M; Holubec L; Ludvikova M; Kalfert D; Franekova J; Povysil C, Biological activity and clinical implications of the matrix metalloproteinases. Anticancer Research 2008, 28, (2B), 1389–1397. [PubMed: 18505085]
22. Visse R; Nagase H, Matrix metalloproteinases and tissue inhibitors of metalloproteinases. Circulation Research 2003, 92, (8), 827–839. [PubMed: 12730128]
23. Smirnov IP; Zhu X; Taylor T; Huang Y; Ross P; Papayanopoulos IA; Martin SA; Pappin DJ, Suppression of α-cyano-4-hydroxycinnamic acid matrix clusters and reduction of chemical noise in MALDI-TOF mass spectrometry. Analytical Chemistry 2004, 76, (10), 2958–2965. [PubMed: 15144210]
24. Stephens EH; Carroll JL; Grande-Allen KJ, The use of collagenase III for the isolation of porcine aortic valvular interstitial cells: rationale and optimization. Journal of Heart Valve Disease 2007, 16, (2), 175–183. [PubMed: 17484468]
25. Stephens EH; Shangkuan J; Kuo JJ; Carroll JL; Kearney DL; Carberry KE; Fraser CD Jr; Grande-Allen KJ, Extracellular matrix remodeling and cell phenotypic changes in dysplastic and hemodynamically altered semilunar human cardiac valves. Cardiovascular Pathology 2011, e157–e167. [PubMed: 20817569]
26. Fang M; Yuan J; Peng C; Li Y, Collagen as a double-edged sword in tumor progression. Tumor Biology 2014, 35, (4), 2871–2882. [PubMed: 24338768]
27. Baiocchini A; Montaldo C; Conigliaro A; Grimaldi A; Correani V; Mura F; Ciccosanti F; Rotiroti N; Brenna A; Montalbano M, Extracellular matrix molecular remodeling in human liver fibrosis evolution. PloS one 2016, 11, (3), e0151736. [PubMed: 26998606]
28. Bolognesi M; Djinovic-Carugo K; Ascenzi P, Molecular bases for human leucocyte elastase inhibition. Monaldi archives for chest disease= Archivio Monaldi per le malattie del torace/ Fondazione clinica del lavoro, IRCCS [and] Istituto di clinica tisiologica e malattie apparato respiratorio, Universita di Napoli, Secondo ateneo 1994, 49, (2), 144–149.
29. Groseclose MR; Massion PP; Chaurand P; Caprioli RM, High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. Proteomics 2008, 8, (18), 3715–24. [PubMed: 18712763]
30. Powers TW; Neely BA; Shao Y; Tang H; Troyer DA; Mehta AS; Haab BB; Drake RR, MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. PLoS One 2014, 9, (9), e106255, pp1–11.
31. Eckhard U; Huesgen PF; Brandstetter H; Overall CM, Proteomic protease specificity profiling of clostridial collagenases reveals their intrinsic nature as dedicated degraders of collagen. Journal of proteomics 2014, 100, 102–114. [PubMed: 24125730]
32. Caprioli RM; Farmer TB; Gile J, Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Analytical Chemistry 1997, 69, (23), 4751–4760. [PubMed: 9406525]
33. Biniossek ML; Schilling O, Enhanced identification of peptides lacking basic residues by LC-ESI-MS/MS analysis of singly charged peptides. Proteomics 2012, 12, (9), 1303–1309. [PubMed: 22589179]
34. Tanco S; Gevaert K; Damme P, C-terminomics: Targeted analysis of natural and posttranslationally modified protein and peptide C-termini. Proteomics 2015, 15, (5–6), 903–914. [PubMed: 25316308]
35. Li Q; Shortreed MR; Wenger CD; Frey BL; Schaffer LV; Scalf M; Smith LM, Global Post-Translational Modification Discovery. Journal of Proteome Research 2017, 16, 1383–1390. [PubMed: 28248113]

36. Schey KL; Anderson DM; Rose KL, Spatially-directed protein identification from tissue sections by top-down LC-MS/MS with electron transfer dissociation. Analytical Chemistry 2013, 85, (14), 6767–6774. [PubMed: 23718750]

37. Nicklay JJ; Harris GA; Schey KL; Caprioli RM, MALDI imaging and in situ identification of integral membrane proteins from rat brain tissue sections. Analytical Chemistry 2013, 85, (15), 7191–7196. [PubMed: 23829295]

38. Wolfe JN, Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976, 37, (5), 2486–2492. [PubMed: 1260729]

39. Hayashi M; Nomoto S; Hishida M; Inokawa Y; Kanda M; Okamura Y; Nishikawa Y; Tanaka C; Kobayashi D; Yamada S, Identification of the collagen type 1 alpha 1 gene (COL1A1) as a candidate survival-related factor associated with hepatocellular carcinoma. BMC Cancer 2014, 14, (1), 108. [PubMed: 24552139]

40. Calon A; Lonardo E; Berenguer-Llergo A; Espinet E; Hernando-Momblona X; Iglesias M; Sevillano M; Palomo-Ponce S; Tauriello DVF; Byrom D, Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. Nature Genetics 2015, 47, (4), 320–329. [PubMed: 25706628]

41. Drifka CR; Loeffler AG; Mathewson K; Keikhosravi A; Eickhoff JC; Liu Y; Weber SM; Kao WJ; Eliceiri KW, Highly aligned stromal collagen is a negative prognostic factor following pancreatic ductal adenocarcinoma resection. Oncotarget 2016, 7, (46), 76197. [PubMed: 27776346]

42. Cheung DY; Duan B; Butcher JT, Current progress in tissue engineering of heart valves: multiscale problems, multiscale solutions. Expert opinion on biological therapy 2015, 15, (8), 1155–1172. [PubMed: 26027436]

43. Lewis JK; Bischof JC; Braslavsky I; Brockbank KGM; Fahy GM; Fuller BJ; Rabin Y; Tocchio A; Woods EJ; Wowk BG, The Grand Challenges of Organ Banking: Proceedings from the first global summit on complex tissue cryopreservation. Cryobiology 2016, 72, (2), 169–182. [PubMed: 26687388]

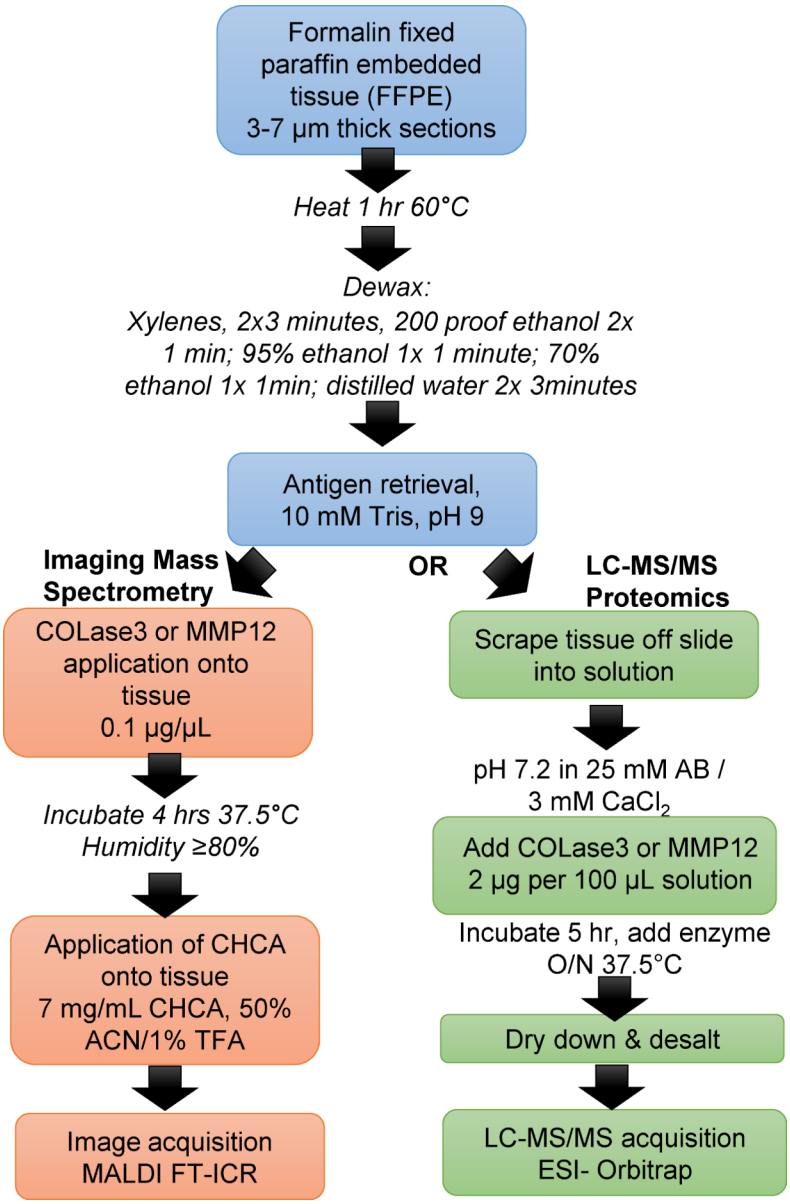

Figure 1.

MALDI IMS and parallel workflow for use of COLase3 or MMP12. Adjacent serial sections were used for in-solution proteomics.

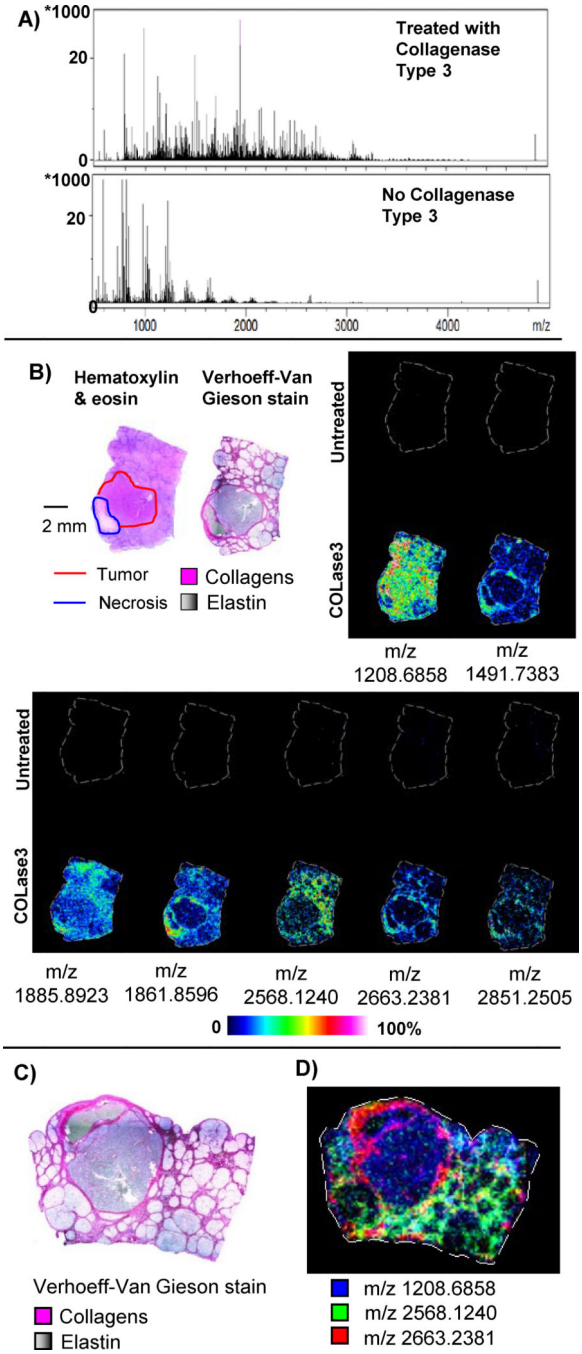

Figure 2.

Collagenase Type III Image Data. A) Representative total ion current across an entire section of liver tissue. Typical high intensity clusters of matrix peaks are observed in the control section that was identically treated except for the addition of COLase3. B) Representative images from select peaks. In each image, the top section is the control. Red line represents tumor region; blue line represents necrotic region. C) Verhoeff Van Gieson stain for collagens. Collagen is hot pink, elastin is black to gray. D) Combined ion image showing distinct unique patterns mapping to collagenous regions.

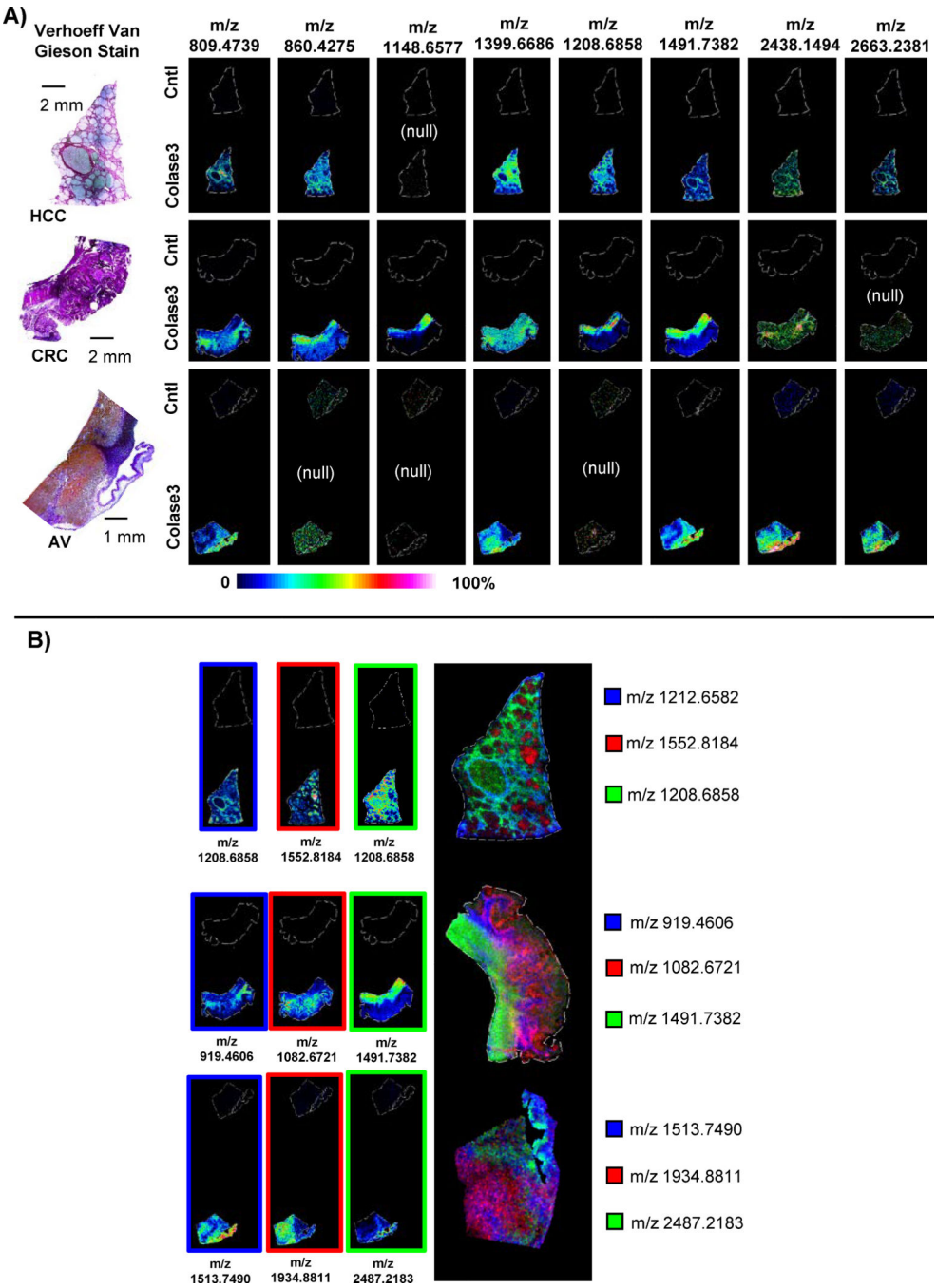

Figure 3.

Examples of COLase3 image data across tissue types. A) Comparison of same masses across tissues. Some peaks were not detected in certain tissue types (marked “null”), suggesting that distinct COLase3 produced peaks will be useful in assessing biological variability. B) Combined Images illustrating distinct localization of COLAse3 produced peptides across tissue types.

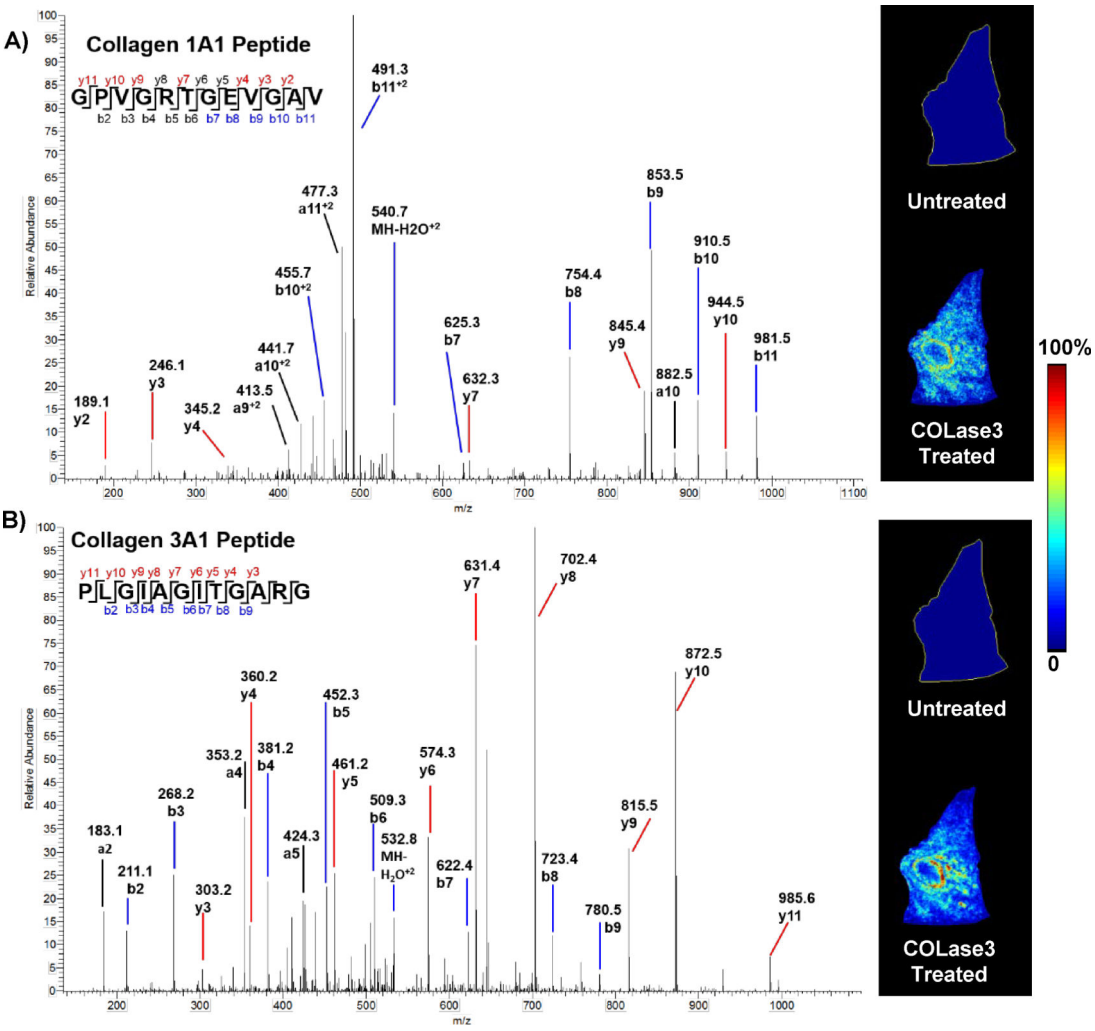

Figure 4.

Label-free HR/AM proteomic data from COLase3 produced digest on FFPE thin tissue sections and matching MALDI FT-ICR IMS images. A) Tandem mass spectra of COL1A1 peptide GPVGRTGEVGAV with MALDI IMS image; B) Tandem mass spectra of COL3A1 peptide PLGIAGITGARG with matching image in the same HCC section as COL1A1.

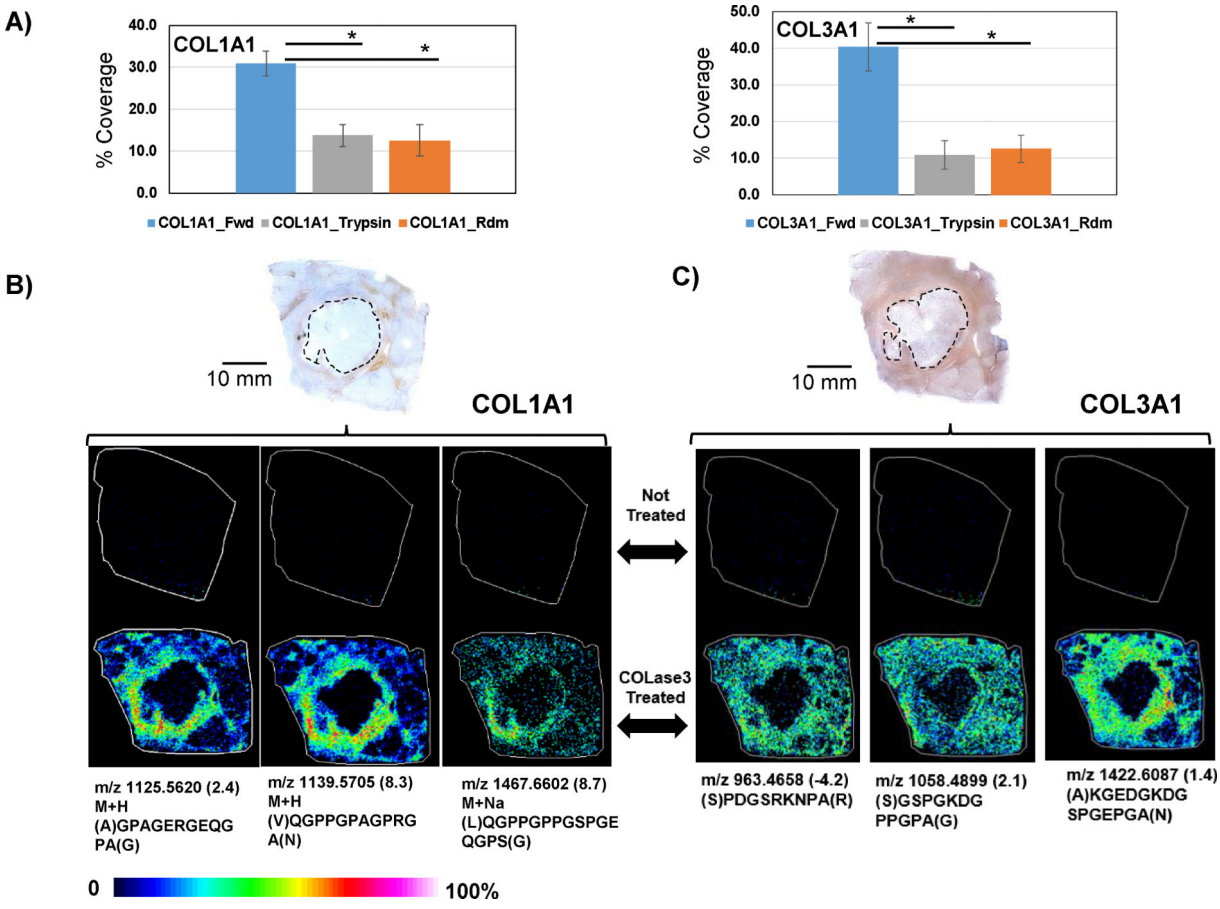

Figure 5.

Peptide mass fingerprint matches to collagens identified by proteomics. A) *In silico* digestion studies using image data coverage showing that COLase3 digestion produces significantly higher sequence matches to COL1A1 and COL3A1 compared to *in silico* generated tryptic peptides or random sequence collagen. B) COL1A1 antibody staining shows COL1A1 closely encircling the tumor (dashed line), which is repeated in the image data by peptides linked to COL1A1 through PMF. C) COL3A1 antibody staining shows high expression in most tissue surrounding the tumor (dashed line), which is also repeated in the image data by peptides linked to COL3A1 through PMF studies. Asterisk indicates student's t-test p-value ≤0.02.

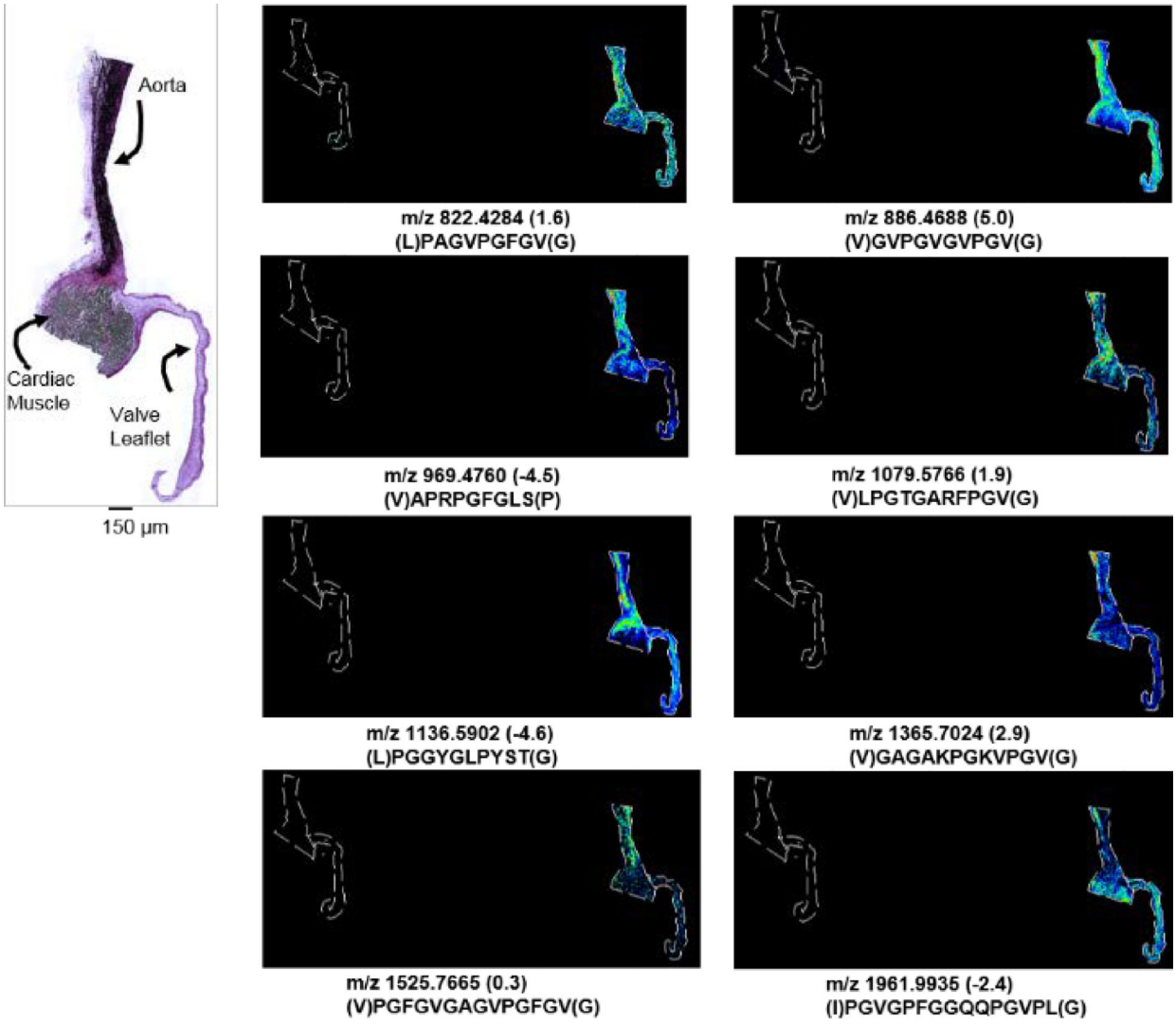

Figure 6.

MMP12 applied to porcine cardiac tissue. Example images mapping to putative porcine elastin sequences by high mass accuracy peptide mass fingerprint on porcine tissue.

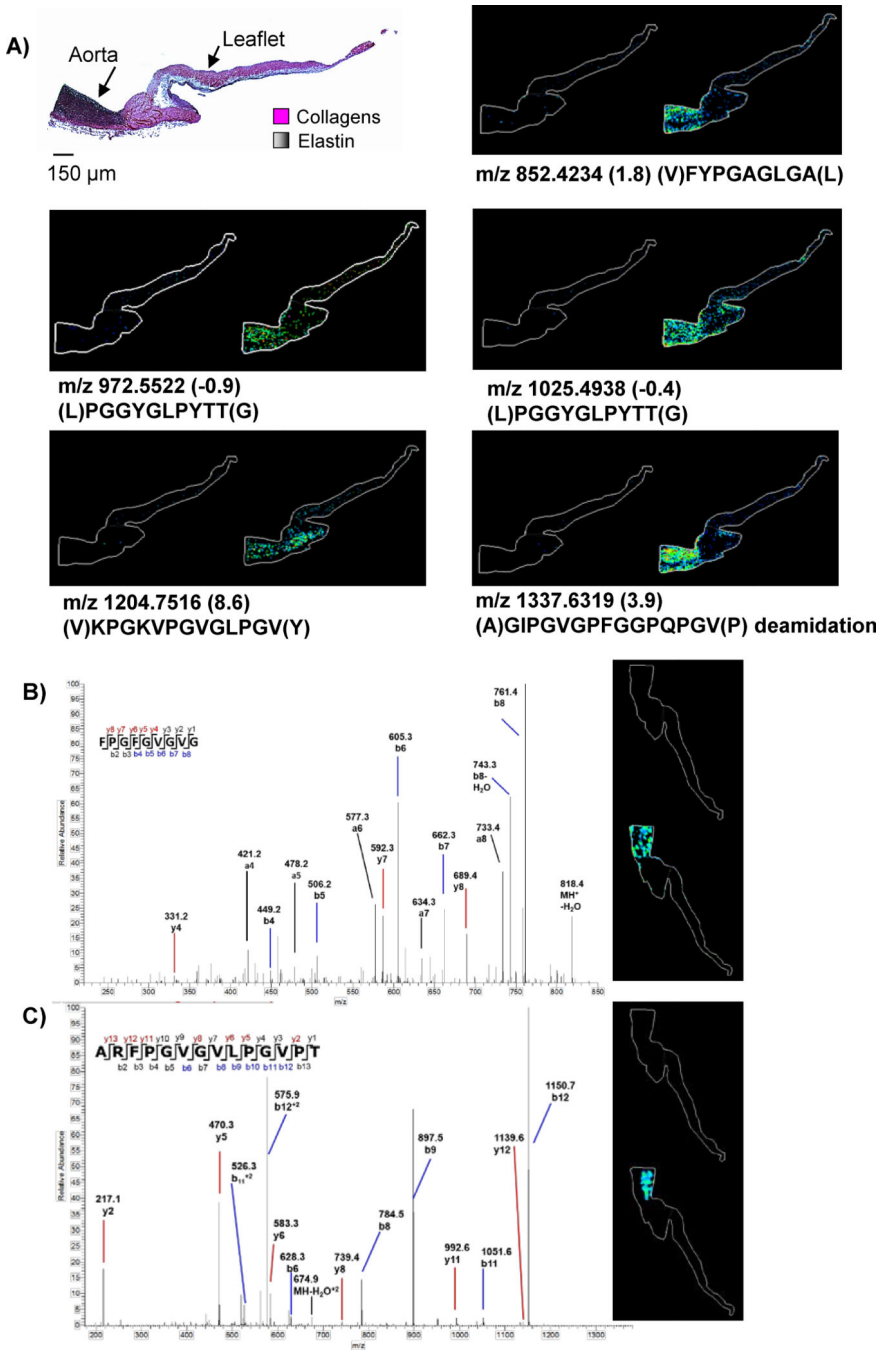

Figure 7.

MMP12 applied to human cardiac tissue. A) Example peptides mapped by peptide mass fingerprinting show putative elastin peptides found by PMF on human tropoelastin mapping to the aorta, stained black by Verhoeff Van Gieson stain. B) Example singly charged peptide from proteomics data with MALDI IMS image (M+H 836.4313 Δppm −1.4). C) Example doubly charged peptide from proteomics data matched to MALDI IMS image (M+H 1366.7720 Δppm-8.6).