scientific reports

Check for updates

OPEN

Targeted proteomic approach for quantification of collagen type I and type III in formalin-fixed paraffin-embedded tissue

Nicole L. Rosin[1], Tara M. L. Winstone[2], Margaret Kelley[2], Jeff Biernaskie[1,3,4,5], Antoine Dufour[6,7] & Dennis J. Orton[2]✉

Collagen is the most abundant protein in mammals and a major structural component of the extracellular matrix (ECM). Changes to ECM composition occur as a result of numerous physiological and pathophysiological causes, and a common means to evaluate these changes is the collagen 3 (Col3) to collagen 1 (Col1) ratio. Current methods to measure the Col3/1 ratio suffer from a lack of specificity and often under- or over-estimate collagen composition and quantity. This manuscript presents a targeted liquid chromatography tandem mass spectrometry (LC–MS/MS) method for quantification of Col3 and Col1 in FFPE tissues. Using surrogate peptides to generate calibration curves, Col3 and Col1 are readily quantified in FFPE tissue sections with high accuracy and precision. The method is applied to several tissue types from both human and reindeer sources, demonstrating its generalizability. In addition, the targeted LC–MS/MS method permits quantitation of the hydroxyprolinated form of Col3, which has significant implications for understanding not only the quantity of Col3 in tissue, but also understanding of the pathophysiology underlying many causes of ECM changes. This manuscript presents a straightforward, accurate, precise, and generalizable method for quantifying the Col3/1 ratio in a variety of tissue types and organisms.

Collagen is the most abundant protein in mammals, with 28 highly conserved subtypes[1]. Located primarily in the extracellular matrix (ECM), the fibrillar collagens 1 (Col1) and 3 (Col3) are the most abundant, making up approximately 70–80% and 5–20% of total collagen, respectively. Remodelling of the ECM occurs during several physiological and pathophysiological processes such as during development, wound healing, fibrosis, inflammation, and cancer, leading to changes in the relative abundance of Col1 and Col3 in a temporal fashion[2,3]. The ratio of Col3 to Col1 (Col3/1) in tissues is often used to understand these processes, however currently available methods to measure Col3 and Col1 abundances are limited to semi-quantitative or qualitative assessments such as gel electrophoresis or imaging[2,4–6]. Total hydroxyproline (Hyp) measurement kits may also be used but provide only an indirect estimate of collagen based on an assumed standard rate of hydroxyprolination in the ECM. Unfortunately, none of these methods provide reliable measurement of collagen components, presenting a clear need for direct quantitation strategies to better characterize ECM composition[7].

Targeted liquid chromatography-mass spectrometry (LC–MS/MS) is a powerful tool for quantification of proteins in a variety of matrices[8,9]. Previous studies have employed semi-quantitative proteomic analyses by untargeted LC–MS/MS to demonstrate relative changes in protein abundance, as well as to identify Hyp modification sites, however these provide largely qualitative or comparative results which are not useful when trying to calculate the Col3/1 ratio[10–12]. In contrast, targeted LC–MS/MS provides highly specific quantitative results and may be multiplexed to allow robust quantitation of dozens or hundreds of peptides in a single analysis[13].

In recent years, there has been renewed interest in understanding the role of the various collagen subtypes in wound healing and fibrosis. As molecular techniques such as scRNAseq have evolved, more questions have

[1]Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada. [2]Department of Pathology and Laboratory Medicine, University of Calgary, 3535 Research Rd NW, Room 1E-415, Calgary, AB T2I 2K8, Canada. [3]Department of Surgery, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada. [4]Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada. [5]Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB, Canada. [6]Department of Physiology and Pharmacology, University of Calgary, Calgary, AB, Canada. [7]McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, AB, Canada. ✉email: dennis.orton@ucalgary.ca

arisen regarding fibroblast gene expression profiles, how this relates to the composition of the ECM during scar formation, and how the properties of the ECM itself orchestrates immune and fibroblast responses to pathological conditions such as injury.

This manuscript presents a simple, reproducible targeted LC–MS/MS method for quantification of both Col1 and Col3 in formalin-fixed paraffin embedded (FFPE) tissues. As a proof of principle, the method is applied for calculation of the Col3/1 ratio on tissues from various sources. In addition to raw collagen quantitation, this method permits evaluation of the frequency of Hyp modification of collagens, providing additional ECM remodelling information not available by any traditional method.

Methods

Materials

Customized peptide standards corresponding to the tryptic peptide fragments (purity > 95%) and matched isotopic internal standards for Collagen 1A1 and 3A1 were purchased from Biosynth (supplemental data Table 1). LC–MS grade acetonitrile, formic acid (> 99%), and xylene were obtained from Fisher Scientific. Ethanol was obtained from Greenfield Inc. Trypsin, ammonium bicarbonate, and acetic acid were purchased from Sigma. Rat tail collagen standard was obtained from Corning Life Sciences.

FFPE tissue specimens

All experiments involving animals and/or human skin samples received before approval from the Health Sciences Animal Care Committee and the Conjoint Health Ethics Review Board at the University of Calgary, respectively and experiments were performed in accordance with relevant guidelines and regulations. The researchers were blinded to sample type for the duration.

Animals

Adult reindeer (*Rangifer tarandus*) were maintained in an outdoor enclosure at the University of Calgary Veterinary Sciences Research Station for the duration of these experiments. Experiments were performed during early antler growth in early spring for males, late spring for females since the timing of antler growth is unique to each sex. Sex was equally distributed across experimental groups, as we have previously noted that there is no significant no sexual dimorphism in the antler velvet or back skin ECM composition. Samples from a total of 13 animals were included in the study, with ages ranging from 4 to 8 years.

All experiments were done in accordance with the University Animal Welfare Committee, the Veterinary Sciences Animal Care Committee and Health Sciences Animal Care Committee at the University of Calgary, the Canadian Council on Animal Care and the Province of Alberta Animal Protection Act and Regulation, and ARRIVE guidelines[14]. Reindeer back skin and antler velvet samples were obtained under University of Calgary animal care protocol AC20-0030 as previously described[3]. In brief, animals were restrained in a hydraulic squeeze and anesthetized (Medetomidine, 0.07–0.15 mg/kg I.M. with Azaperone, 0.2 mg/kg I.M., if additional duration of anaesthesia required). Anaesthesia, heart and respiratory rate, and temperature were monitored throughout. For antler tissue collection, local anaesthetic was provided (2% Lidocaine–Wyeth) with a ring block and a tourniquet applied; an inverted "L" block was done along the dorsal back. Analgesia was provided (Meloxicam, 0.5 mg/kg S.Q.). Back skin and antler velvet were clipped to remove hair, and aseptically prepared using betadine surgical scrub and ethanol. Full-thickness excisional wounds were created using a 12 mm biopsy punch. Sedation was reversed (Atipamezole, IM at 5× the dose of medetomidine used). Tissue samples were fixed in 10% formalin in phosphate-buffered saline overnight at room temperature, then transferred to 70% ethanol, prior to embedding in paraffin.

Human samples

Explanted lungs from patients with idiopathic pulmonary fibrosis (IPF) undergoing transplantation were obtained from the University of Alberta under (REB15-1607). Informed consent was obtained from patients prior to any experimentation. Non-transplanted human lungs (controls) from normal donors were obtained from a tissue retrieval service (International Institute for the Advancement of Medicine, Edison, NJ). Ethical approval to receive and use lung tissues was obtained from the Conjoint Health Research Ethics Boards of the University of Calgary and the University of Alberta, and from the Internal Ethics Board of the International Institute for the Advancement of Medicine (REB15-0336). The lungs were sampled, fixed in 10% neutral buffered formalin for 24–48 h, and FFPE, blocks prepared.

Protein extraction and digestion

Collagen was extracted from FFPE tissues from either scrolls or on side using a method as described previously, with some modifications[15,16]. Tissue was obtained from the FFPE blocks as either 5 µm thick scrolls or sections on glass slides (4 µm sections for lung tissue). The same processing was used for all tissues, regardless of their source. Briefly, deparaffinization was performed by three washes with xylene for 10 min followed by centrifugation for 2 min at 20,000 × *g* and removal of supernatant (for scrolls) or by tipping the side on its side. Next, tissues were rehydrated by sequential addition of ethanol with increasing amounts of water (100% ethanol, 95% ethanol, 70% ethanol) with a 2-min incubation followed by centrifugation steps between each addition. The glass mounted sections were scraped off the slides using a razor blade and transferred to a 1.5 mL microcentrifuge tube following the addition of 5–10 µL of water. Tissue lysis and breakage of crosslinking related to formalin fixation was performed through the addition of 50% acetonitrile in 100 mM ammonium bicarbonate followed by sonication for 60 min in a sonicating water bath (Branson 1800). Samples were transferred to a heating block and incubated for 60 min at 90 °C. After each step, samples were subjected to centrifugation to ensure no sample loss. The lysis

buffer was then removed by evaporation at 40 °C under a flow of air and reconstituted directly into 100 mM ammonium bicarbonate buffer containing trypsin for digestion. At each time point, digestion was stopped by addition of formic acid to a final concentration of 0.1%. Optimization of proteolytic digestion conditions, extraction protocols, and chromatographic conditions was performed to minimize time required for processing and analysis. Results of optimization experiments are presented in Supplementary Tables 2–4.

The concentration of peptides in the mixture employed the BCA assay (Peirce) according to manufacturer's protocols. At this point, COL1A1 and COL3A1 internal standards were added and samples were subject to LC–MS/MS analysis. The workflow is summarized in Fig. 1.

Untargeted post-translational modification LC–MS/MS analysis

The untargeted LC–MS/MS analysis was carried out by the Southern Alberta Mass Spectrometry core facility at the University of Calgary, Canada. Analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, Mississauga, ON) operated with Xcalibur (version 4.0.21.10) and coupled to a Thermo Scientific Easy-nLC (nanoflow Liquid Chromatography) 1200 system. Tryptic peptides were loaded onto a C18 trap (75 µm × 2 cm; Acclaim PepMap 100, P/N 164946; Thermo Fisher Scientific) at a flow rate of 2 µL/min of solvent A (0.1% formic acid and 3% acetonitrile in LC-mass spectrometry grade water). Peptides were eluted using a 120 min gradient from 5 to 40% (5–28% in 105 min followed by an increase to 40% B in 15 min) of solvent B (0.1% formic acid in 80% LC-mass spectrometry grade acetonitrile) at a flow rate of 0.3% µL/min and separated on a C18 analytical column (75 µm × 50 cm; PepMap RSLC C18; P/N ES803; Thermo Fisher Scientific). Peptides were ionised by electrospray using 2.3 kV into the ion transfer tube (300 °C) of the Orbitrap Lumos operating in positive mode. The Orbitrap first performed a full MS scan at a resolution of 120,000 FWHM to detect the precursor ion having a mass-to-charge ratio (m/z) between 375 and 1575 and a + 2 to + 4 charge. The Orbitrap AGC (Auto Gain Control) and the maximum injection time were set at 4×10^5 and 50 ms, respectively. The Orbitrap was operated using the top speed mode with a 3 s cycle time for precursor selection. The most intense precursor ions presenting a peptidic isotopic profile and having an intensity threshold of at least 2×10^4 were isolated using the quadrupole (isolation window of m/z 0.7) and fragmented with HCD (38% collision energy) in the ion routing Multipole. The fragment ions (MS2) were analyzed in the Orbitrap at a resolution of 15,000. The AGC, the maximum injection time and the first mass were set at 1×10^5, 105 ms, and 100 ms, respectively. Dynamic exclusion was enabled for 45 s to avoid of the acquisition of the same precursor ion having a similar m/z (± 10 ppm).

Bioinformatic analysis and post-translational modification identification

Spectral data were matched to peptide sequences in the bovine UniProt protein database using the MaxQuant software package v.1.6.0.1, peptide-spectrum match false discovery rate (FDR) of < 0.01. Search parameters included a mass tolerance of 20 ppm for the parent ion, 0.05 Da for the fragment ion, carbamidomethylation of cysteine residues (+ 57.021464), variable N-terminal modification by acetylation (+ 42.010565 Da), variable methionine oxidation (+ 15.994915 Da) and variable hydroxyproline modification (+ 15.9949146221).

For the proteomics data, cleavage site specificity was set to Trypsin/P with up to two missed cleavages allowed. Database searches were limited to a maximal length of 40 residues per peptide. Peptide sequences matching reverse or contaminant entries were removed.

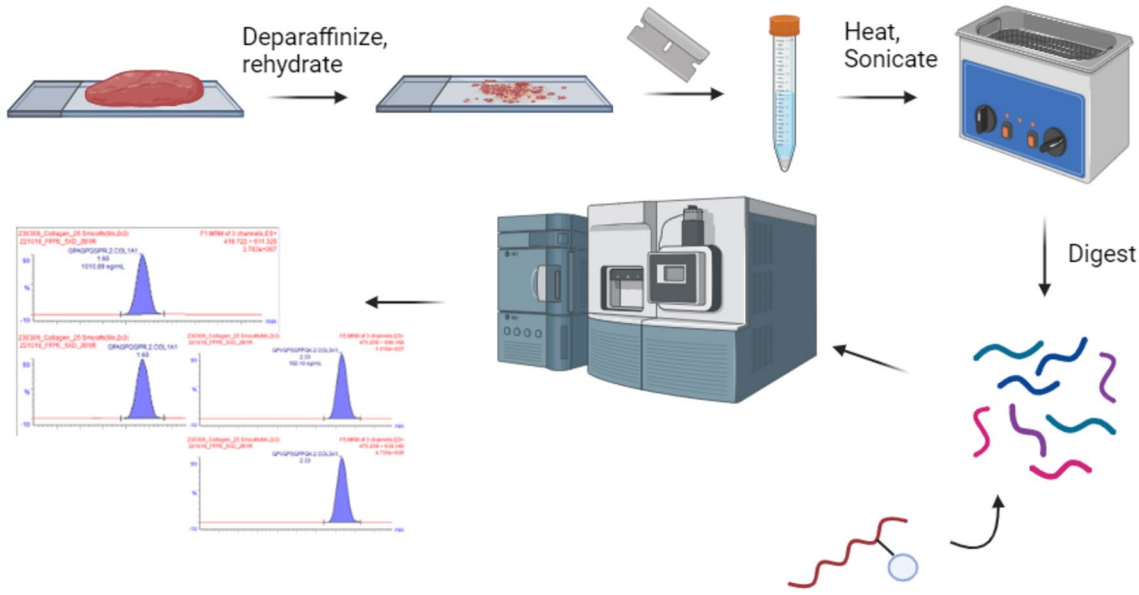

Figure 1. Finalized sample preparation workflow for the LC–MS/MS method.

Multiple reaction monitoring LC–MS/MS analysis

Targeted analysis was performed by multiple reaction monitoring (MRM) using surrogate marker peptides from COL1A1 and COL3A1, with and without hydroxyproline residues. Briefly, 5 µL of each sample was analyzed on a Waters Xevo TQ-XS Triple Quadrupole Mass Spectrometer equipped with an Acquity UPLC system. The stationary phase was a Waters Acquity C18 BEH 130 Å 1.7 µm, 2.1 mm × 50 mm column (Waters; PN 186002350) and mobile phase A consisted of clinical laboratory grade water with 0.1% Formic Acid (Fisher; PN A117-50) and B Acetonitrile (Fisher; PN A955-4) with 0.1% Formic Acid. Separation of peptides was achieved using the following gradient: 5 to 8% B from 0 to 0.5 min, up to 10%B from 0.5 to 2.0 min up to 95% B at 3.0 min. The total run time for each analysis was approximately 5.5 min from injection-to-injection. MRM parameters (e.g. precursor and fragment ion m/*z* values, collision energy, cone voltage) for each peptide were obtained from Skyline[18]. Initial testing began with at least five peptides and five MRM transitions per peptide prior to optimization. Each was tested for uniqueness, sensitivity, and reproducibility using a blank matrix of human serum albumin and FFPE tissues from various sources. Depending on fragment ion detectability in the MS, the final MRM method consisted of two or three transitions for reliable quantification for each peptide and internal standard included in analysis. Specific MRM parameters employed for this study are included in Supplemental Table 1.

Collagen 1 and 3 quantification

For calibration, a surrogate matrix of human serum albumin was chosen as it mimics the protein samples being analyzed without having any collagen in it. Using this method, this method was able to quantify COL1A1 and COL3A1 at concentrations ranging from 1.0 to 1000 µg/L and 0.5–300 µg/L respectively. Quality control materials were prepared as a bulk digest of a single tissue section which was aliquoted, with a single aliquot thawed prior to analysis of each batch. In addition, a rat tail collagen (Corning, 354236) material was included to monitor processing efficiency however contained only Col1. These materials were used to monitor batch-to-batch consistency and provide precision estimates.

Tissue immuno-staining and image analysis by SHG

5 µm thick sections from the FFPE blocks were immuno-stained with rabbit anti-collagen type 3 (600-401-105, Rockland) followed by goat anti-rabbit Alexa488 secondary antibody. Sections were imaged using a multiphoton confocal microscope (Leica SP8 confocal) and second harmonic generation (SHG). First, collagen 3 was imaged using the 488 laser, followed by total collagen imaged with SHG using a Ti:Sa Chameleon multiphoton tuneable laser (Coherent, Santa Clara, CA) at 800 nm, both with a 20× water immersion objective lens.

Single cell RNA sequencing data analysis

The scRNAseq data provided in Fig. 4 was extracted from a previously published dataset[3]. Briefly, a 12 mm biopsy of skin from the back and velvet from the antlers were harvested. The tissue was digested using collagenase IV and single cells isolated using fluorescence associated cell sorting (FACS). Single cells were processed for single cell 3′ gene expression using a 10× Genomics Controller with NextGEM V3.1 chemistry according to the manufacturer's protocol. cDNA libraries were sequenced using an Illumina NovaSeq. Data was processed using 10× Genomics CellRanger platform, followed by Seurat (version 4) in R for analysis[19].

Data analysis

The concentration of each surrogate peptide from COL1A1 and 3A1 were calculated by comparison to a calibration curve. From those values, the relative content of Col1 and Col3 levels were then calculated according to the following equations (calculations are included in Supplemental Table 6):

1. $\mathrm{Col1}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right) = \frac{\left[\mathrm{COL1A1\ peptide}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right)\right]}{2}$
2. $\mathrm{Col3}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right) = \frac{\left[\mathrm{COL3A1\ peptide}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right)\right]}{3}$
3. $Col3/1\ \mathrm{Ratio} = \mathrm{Col3}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right) \Big/ \mathrm{Col1}\left(\frac{\mathrm{mol}}{\mathrm{L}}\right)$

Results

Surrogate peptide selection and chromatographic separation for Col1 and Col3

Surrogate peptides to be used for quantitation of COL1A1 and COL3A1 were selected by 1) uniqueness to the protein of interest within the proteome, and 2) presence only once within the amino acid sequence of each collagen using Skyline[18]. To confirm the presence of the selected peptides an untargeted proteomic analysis was performed on FFPE tissue extracts from skin and antler tissues obtain from reindeer. The results from the untargeted analysis were loaded into Skyline, and both target peptides were identified. A single Hyp modification was noted in the selected COL3A1 peptide at position 8, within the expected Xaa-Yaa-Gly motif[20]. Notably, although Hyp sites were identified on the selected COL1A1 peptide in the untargeted analysis, these were found to be false identifications on closer examination within Skyline, thus Hyp-modified Col1 peptide is removed from further discussion.

Optimization of FFPE tissue extraction and LC–MS/MS conditions

Optimization of chromatographic conditions was performed to ensure good peak shape and adequate separation from interferences for the selected peptides (Fig. 2). Additional optimization experiments included trypsin enzyme quantity, proteolytic digestion time, and extraction efficiency by cycles of sonication with heat using two

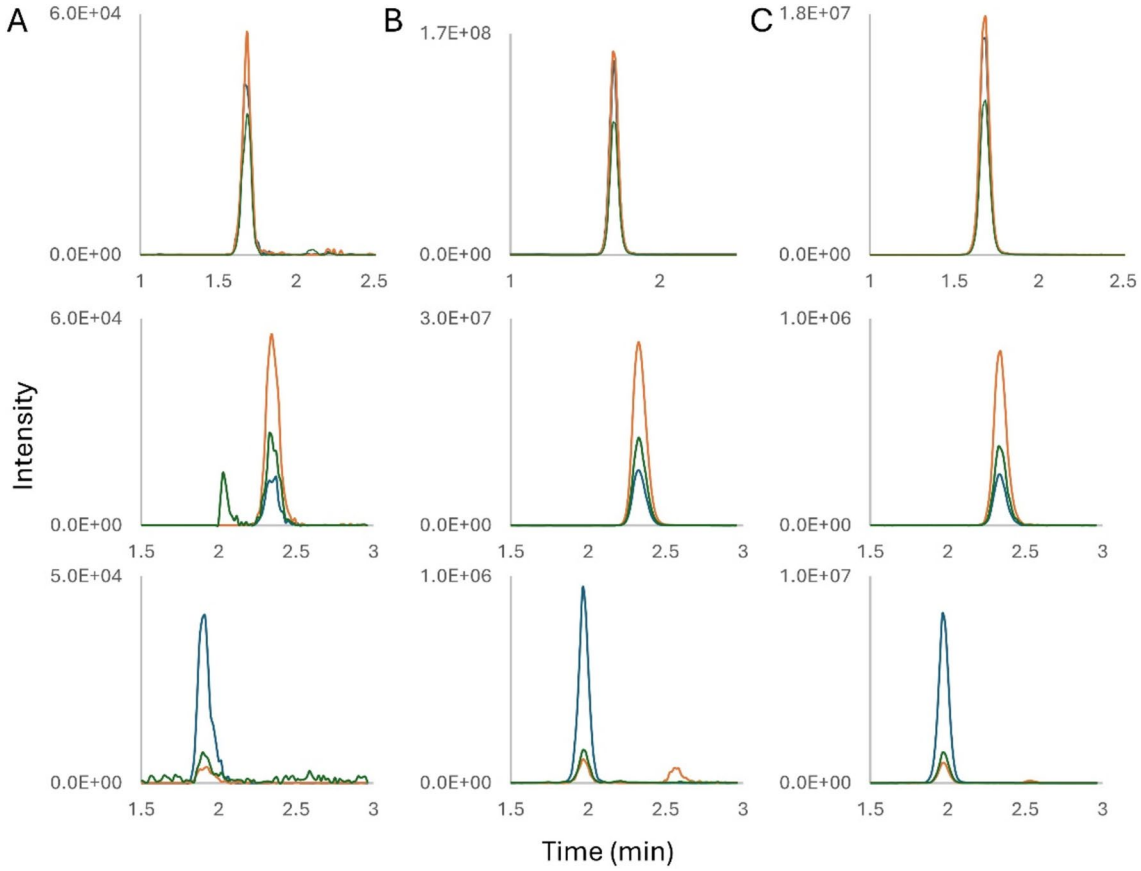

Figure 2. Representative chromatograms of peptide MRMs for the low calibrator (0.0025 ng on-column for each peptide) (**A**), a representative back skin sample (**B**) and a representative antler velvet sample (**C**) for COL1A1 (top row), COL3A1 (middle row) and hydroxyprolinated COL3A1 (bottom row). The calculated amount of each peptide was 9.5 ng, 1.5 ng, and 0.4 ng (back skin) and 0.7 ng, 0.4 ng, and 0.3 ng (antler velvet) for COL1A1, COL3A1, and COL3A1-Hyp, respectively. The three MRM transitions are included to visualize the background signal for each.

distinct tissue types from reindeer (antler velvet and back skin). Interestingly, while the COL1A1 peptide was rapidly and reproducibly recovered from FFPE tissues, the COL3A1 peptide required an overnight digestion to achieve high recovery (Fig. 3). Therefore, an overnight digestion protocol was chosen to permit equivalent and quantitative recovery of both peptides. Inclusion of an isotopically labelled internal standard for each peptide following digestion improved accuracy and precision while mitigating any deleterious matrix effects in the LC–MS/MS analysis. FFPE sections of various sized tissue samples were used to optimize trypsin digestion conditions. Dimensions of tissue were measured to the nearest millimetre to determine an approximate tissue area (mm^2) and the data is presented as nanomoles of collagen per mm^2. The finalized protocol is summarized in the methods section (Fig. 1).

Assay reproducibility and limit of detection

The precision and accuracy of the targeted LC–MS/MS method are summarized in Table 1 (experimental data is available in Supplemental Table 5). Overall, the method demonstrated excellent between-batch precision of 2.6% and 4.1% for COL3A1, and 3.5% and 3.6% for COL1A1 for two levels of quality control (QC) material, respectively. As these QCs were prepared from a single batch of FFPE tissue extracts, an additional QC sample was made from commercially available rat tail collagen. The rat tail QC contained only Col1, but provides a complete estimation of process imprecision, including crosslink breakage and digestion and showed good precision at 4.6%.

The size of tissue sections employed in this study ranged from approximately 20–80 mm^2 with 4 or 5 µm thickness and an average protein yield of 1.3 µg/mm^2. As Col1 is at much higher abundance in tissues than Col3, the assay was designed to cover as broad of a concentration range as possible to permit quantification of both analytes in a single run. Calibration curves were prepared using peptide standards for each peptide, ranging from 0.5 to 1000 µg/L for COL1A1 and 0.5 to 300 µg/L for both COL3A1 and Hyp COL3A1. As Col3 is generally on the order of 5–20% of the total collagen, the calibration ranges were successfully able to quantify Col3 and Col1 from all tissues analyzed in this study.

Importantly, this assay quantifies the amount of peptide within each extract. The peptide concentration must then be converted to a quantitative value for both Col1 and Col3 proteins. Peptide concentrations (in µg/L) were first converted to nmol/L by dividing the µg/L concentration by the molecular weight of the peptide, and then

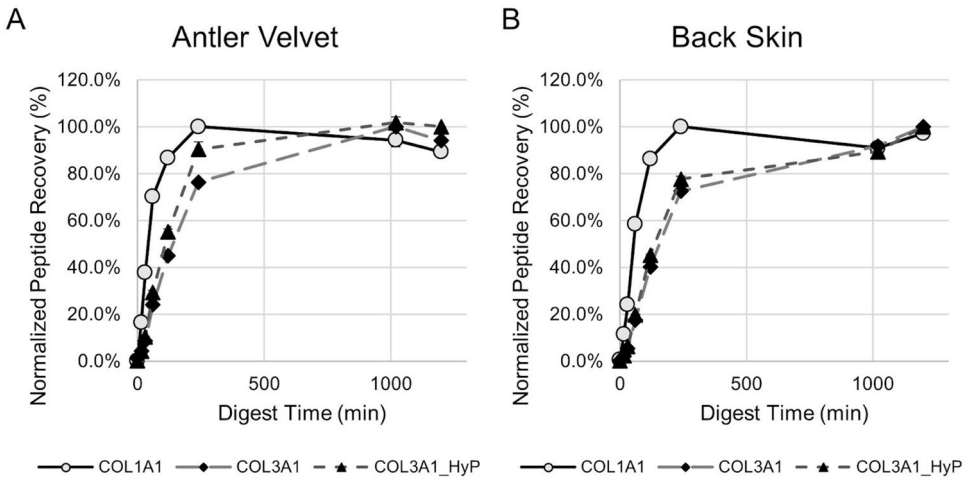

Figure 3. Time Course of trypsin digestion time versus peptide recovery for each surrogate peptide employed in the targeted LC–MS/MS method on tissue sections from (**A**) antler velvet and (**B**) back skin. The overnight digestion provides consistent recovery of all three peptides, thus was selected as optimum.

Protein	Peptide	Low QC			High QC			Rat Tail*			Measuring Range (µg/L)
		Mean (µg/L)	SD (µg/L)	CV (%)	Mean (µg/L)	SD (µg/L)	CV (%)	Mean (µg/L)	SD (µg/L)	CV (%)	
COL1A1	GPAGPQGPR	302.7	6.8	2.2	468.8	12.9	2.8	1479.3	196.9	13.3	0.5–1000
COL3A1	GPVGPSGPPGK	15.7	0.6	3.9	87.5	2.4	2.7				0.5–300
	GPVGPSG**P**^PGK	18.3	0.7	3.9	30.3	1.0	2.7				0.5–300

Table 1. Performance characteristics for the LC–MS/MS method established off 20 separate analyses of each QC sample. *Col3 not present in Rat Tail. ^Hydroxyproline residue.

multiplied by the number of mols of COL1A1 and COL3A1 present in Col1 and Col3, respectively. As Col1 is a heterotrimeric structure made of up two COL1A1 and one COL2A1 subunits, the COL1A1 peptide concentration in nmol/L was divided by two. Col3 is a homotrimer of three COL3A1 subunits, thus the COL3A1 peptide concentration was divided by three. Of note, the concentrations of the Hyp and non-Hyp modified peptides were summed prior to calculation of Col3. These formulae are included in the Materials and Methods. The Col3/1 ratio may then be calculated at this point or multiplied by the sample volume and normalized to the tissue section area in nmol/mm^2.

Assessment of LC–MS method accuracy

The accuracy of the LC–MS/MS method was thoroughly evaluated by comparing against two microscopic methods on adjacent tissue sections as well as an scRNAseq dataset collected on samples from a separate set of experiments. The overall structure of reindeer skin section is presented in Fig. 4A using a standard haematoxylin and eosin (H&E) stain. Figure 4B is a PolScope image, which infers collagen and ECM reorganization based on the relative light scattering. This is a qualitative approach known to provide structural changes in the ECM but does not provide quantitative values for ECM components. Figure 4C–E represent a SHG image (C), Col3 immunostaining in greed (D) and an overlaid image of these (E). The SHG image provides an estimate of total collagen (white) while the immunostain (green) shows the relative quantity of Col3. This method was applied to tissues from reindeer back skin (Fig. 4F) and compared to the results obtained from the LC–MS/MS method (Fig. 4G). Finally, the relative expression of the collagen genes in a scRNAseq dataset is also shown (previously published and publicly available at http://www.biernaskielab.ca/reindeer_atlas/[3]; Fig. 4H). Taken together, these results suggest the LC–MS/MS assay provides accurate quantitative information for Col1 and Col3 from FFPE tissue sections.

Effect post-translational modification quantitation for ECM

In a proof-of-principle experiment, the LC–MS/MS protocol was applied to quantify Col3/Col1 in FFPE tissues from reindeer back skin and antler velvet with a focus on Hyp modification (Fig. 5A). These tissues are known to express different quantities of collagen subtypes and display significant differences in their Col3 to total collagen ratio. Inclusion of the Hyp and non-Hyp modified peptides in the LC–MS/MS assay demonstrates the difference in Col3 abundance between tissues is primarily due to the non-hydroxyprolinated Col3, and that the relative quantity of hydroxyprolinated Col3 remains somewhat consistent between tissues.

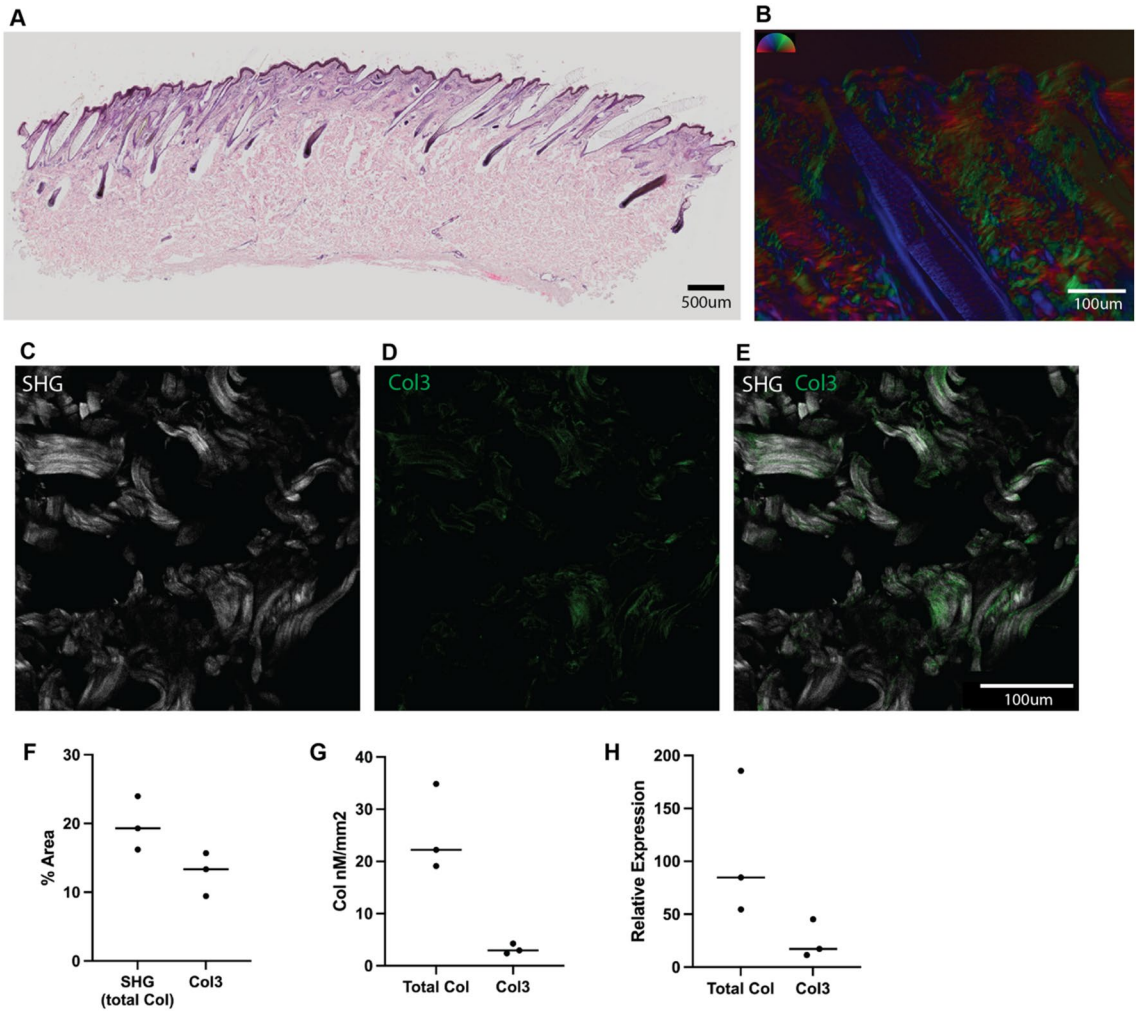

Figure 4. Accuracy and comparability of the proposed LC–MS/MS method versus alternate methodologies for measuring collagen in reindeer back skin. (**A**) Haematoxylin and eosin (H&E) stained representative image and (**B**) a PolScope image that visualizes the structure of the tissue and ECM in the section. (**C**) SHG representing total collagen is shown in white, while (**D**) provides the Col3 immunostaining in green and (**E**) provides the overlaid image of SHG and immunostaining. Results of these semi-quantitative analyses are presented for (**F**) SHG/immunostaining, (**G**) LC–MS/MS, and (**H**) scRNAseq. LC–MS/MS compared closest with scRNAseq, and imaging was less able to separate Col1 and Col3 quantities in the skin samples.

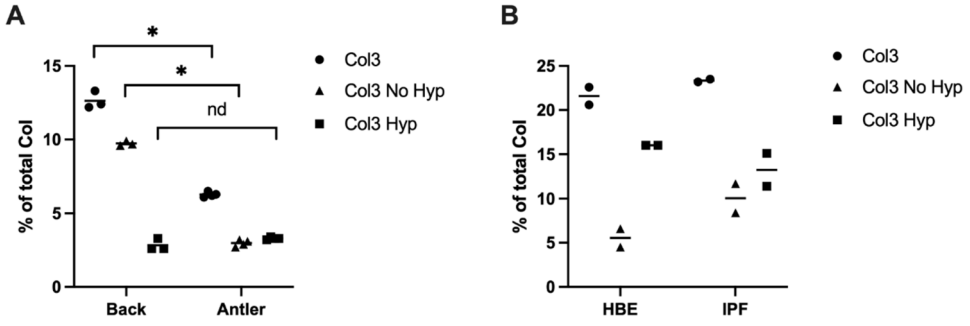

Figure 5. LC–MS/MS results for Col3-to-total collagen ratio for total, Hyp-modified, and non-hydroxyprolinated Col3 in (**A**) reindeer back skin versus antler velvet, and (**B**) healthy versus IPF human lung tissue.

In a separate experiment, the LC–MS/MS method was applied to FFPE tissues from human lung biopsies taken from healthy (HBE) and patients with interstitial pulmonary fibrosis (IPF). In diagnosis of IPF, tissues are known to demonstrate a marked increase in fibrosis and ECM remodelling, and monitoring turnover of Col3 and Col1 has been proposed as a biomarker for IPF disease progression[21]. Figure 5B demonstrates only a marginal change in the total Col3/1 ratio, but a statistically significant change in the Hyp-form of Col3, thus implicating the hydroxyprolinated form of Col3 as the primary driver of changes in the Col3/1 ratio in IPF.

Discussion

This manuscript presents a straightforward, accurate method for quantification of the Col3/1 ratio for FFPE tissue analysis. The method employs buffers and methodologies routinely available in most clinical or research laboratories and provides a template for researchers aiming to better quantify elements of the ECM in a broad spectrum of pathologies such as wound healing, fibrosis or cancer metastasis. Additionally, this method demonstrates the importance of quantifying post-translationally modifications in collagen to provide additional information into the pathophysiology of disease.

Considerations for method development

Development of a workflow for processing FFPE tissues required optimization of several parameters, including deparaffinization and rehydration of tissue sections followed by breaking crosslinks formed during formalin fixation by sonication and heat. Numerous studies have investigated optimal conditions for performing this step[15,16,22], however we found simple sonication and heat in a buffer containing ammonium bicarbonate and acetonitrile to be acceptable. Addition of detergents (e.g. sodium dodecyl sulphate, SDS) or chaotropic agents (e.g. urea) was deemed undesirable as these reagents are not routinely available in our clinical laboratory. Interestingly, we found sequential application of sonication (60 min in sonicating water bath) and heat (90 °C for 60 min) not only improved recovery from FFPE tissue sections, but also for the rat tail collagen standard. This effect is attributed to the sonication breaking crosslinks also found in normal collagen fibres as well as those found in FFPE processed tissues.

Selection of the surrogate peptides used in protein quantitation is an integral part of targeted proteomic method design. When selecting optimal peptides and MRM transitions for the final method, initial analysis included up to five peptides from each protein and five MRM transitions from each. Consideration was given to not only sensitivity and reproducibility in the LC–MS system, but also uniqueness to their respective collagen according to Skyline, a tool developed for establishing reproducible targeted proteomic methods[18]. In the end, three MRM transitions were chosen for each peptide included in the method. Specifically for collagen, inclusion of Hyp modifications was found to be imperative as these are common modifications which may vary under different pathophysiological conditions[20,23]. Additionally, chosen peptides were vetted by a literature search as well as an untargeted analysis using an untargeted proteomic approach. The chosen COL1A1 peptide has been identified in previous LC–MS studies with both Hyp form and unmodified[12], however this modification was never seen in our study. The peptides were further vetted for reproducibility between tissue types as well as sensitivity. Fortunately, Hyp modifications occur via enzymatic action of prolyl hydroxylase enzymes, which act at specific locations within the collagen sequence (Xaa-Yaa-Gly), permitting an informed decision when selecting modified peptides for analysis[20].

Effect of hydroxyproline on collagen quantification

Interestingly, differential expression of P4H, the primary enzyme responsible for Hyp formation in fibrillar collagen is known to be altered in hypoxemia, cancer progression, and exhibits temporal changes in expression during wound healing[12]. This presents a clear benefit to our targeted LC–MS/MS method as inclusion of Hyp residues allows quantification of the hydroxyprolination of collagen under different pathophysiological conditions. In fact, when applying our method to tissue sections from different sources we found that not only is the Col3/1 ratio different between tissue types, but also that the content of Hyp-modified Col3 is different between back skin and antler velvet Col3 (Fig. 5A). Previous work by Bernaskie et al.[3] has demonstrated similarities in cellular composition and gene expression profiles between reindeer skin and human skin and while investigation into the physiology of wound healing and skin regeneration is outside the scope of this manuscript, this application of targeted LC–MS/MS may be applied in understanding ECM makeup during wound healing and fibrosis. Few studies have focussed on the content of the ECM remodelling, instead using immunostaining or gene expression studies on tissues or indirect serum-based metabolomic or proteomic methods rather than quantitative analysis, likely owing to a lack of available methods to do so[21,24,25].

Notes on quantitation

This method takes advantage of the specificity provided by LC–MS/MS analysis to provide high confidence in collagen subtype quantitation in FFPE tissues. This method includes calibration curves for COL1A1 and COL3A1 peptides by spiking known concentrations of each target peptide into an artificial protein matrix made up of human serum albumin. This matrix mimics the environment of extracted FFPE tissues which are subsequently used for analysis, and is a common approach for establishing calibration curves in targeted proteomic methods. Additionally, inclusion of isotopically labelled internal standard peptides permits more reliable quantitation and verification of chromatographic performance. These peptides are spiked into all calibrators, QC samples, and FFPE extracts following trypsin digestion to control for matrix effects on the LC–MS/MS analysis. Although inclusion of internal standards improves confidence and assay precision, future methods may choose to include or exclude them based on the performance requirements of the assay. Additional peptides may also be added to

the targeted LC–MS/MS method to improve confidence in the relative differences in Hyp modification where required.

Limitations

The primary limitation of the LC–MS/MS approach is the loss of spatial information for distribution of Col3 and Col1 in tissues, as the proteolytic digestion and extraction steps are innately destructive. However, the use of staining and LC–MS/MS analysis of sequential tissue sections may help mitigate this effect, as only a single tissue section is required for analysis. Additionally, we found that although individual peptide recovery remains consistent for both COL1A1 and COL3A1, the overall amount of protein extracted from FFPE tissue sections can be highly variable between tissue sections. Effects of variable protein recovery are minimized here, as the main application is quantification of the Col3/1 ratio, thus equivalent recovery is a more important consideration than recovery of an absolute amount of protein from the tissue section.

Conclusions

This study presents a simple and robust method for quantifying the Col3/1 ratio, with additional information on important post-translational modifications. The approach can be easily multiplexed to include peptides corresponding to other proteins of interest such as other collagen subtypes or ECM proteins. It is noted that inclusion of additional peptides with hydroxyproline modification may improve the confidence in assessing post-translational modifications in ECM.

Data availability

The datasets generated and/or analysed during the current study are available in the PRIDE repository[17], http://www.ebi.ac.uk/pride/archive/projects/PXD050877 and PXD050877.

Received: 5 March 2024; Accepted: 23 July 2024
Published online: 01 August 2024

References

1. Myllyharju, J. & Kivirikko, K. I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **20**, 33–43. https://doi.org/10.1016/j.tig.2003.11.004 (2004).
2. Singh, D., Rai, V. & Agrawal, D. K. Regulation of collagen I and collagen III in tissue injury and regeneration. *Cardiol. Cardiovasc. Med.* **2023**, 7. https://doi.org/10.26502/fccm.92920302 (2023).
3. Sinha, S. *et al.* Fibroblast inflammatory priming determines regenerative versus fibrotic skin repair in reindeer. *Cell* **185**, 4717-4736. e25. https://doi.org/10.1016/j.cell.2022.11.004 (2022).
4. Song, K. *et al.* Collagen remodeling along cancer progression providing a novel opportunity for cancer diagnosis and treatment. *Int. J. Mol. Sci.* **2022**, 23. https://doi.org/10.3390/ijms231810509 (2022).
5. Stevenson, K., Kucich, U., Whitbeck, C., Levin, R. M. & Howard, P. S. Functional changes in bladder tissue from type III collagen-deficient mice. *Mol. Cell. Biochem.* **283**, 107–114. https://doi.org/10.1007/s11010-006-2388-1 (2006).
6. Pauschinger, M. *et al.* Dilated cardiomyopathy is associated with significant changes in collagen type I/III ratio (1999). http://www.circulationaha.org.
7. Qiu, B. *et al.* Measurement of hydroxyproline in collagen with three different methods. *Mol. Med. Rep.* **10**, 1157–1163. https://doi.org/10.3892/mmr.2014.2267 (2014).
8. Sprung, R. W. *et al.* Precision of multiple reaction monitoring mass spectrometry analysis of formalin-fixed, paraffin-embedded tissue. *J. Proteome Res.* **11**, 3498–3505. https://doi.org/10.1021/pr300130t (2012).
9. Steiner, C. *et al.* Applications of mass spectrometry for quantitative protein analysis in formalin-fixed paraffin-embedded tissues. *Proteomics* **14**, 441–451. https://doi.org/10.1002/pmic.201300311 (2014).
10. van Huizen, N. A., Ijzermans, J. N. M., Burgers, P. C. & Luider, T. M. Collagen analysis with mass spectrometry. *Mass Spectrom. Rev.* **39**, 309–335. https://doi.org/10.1002/mas.21600 (2020).
11. Pataridis, S., Eckhardt, A., Mikulíková, K., Sedláková, P. & Mikšík, I. Identification of collagen types in tissues using HPLC-MS/MS. *J. Sep. Sci.* **31**, 3483–3488. https://doi.org/10.1002/jssc.200800351 (2008).
12. Montgomery, H. *et al.* Proteomic profiling of breast tissue collagens and site-specific characterization of hydroxyproline residues of collagen alpha-1-(I). *J. Proteome Res.* **11**, 5890–5902. https://doi.org/10.1021/pr300656r (2012).
13. Gaither, C., Popp, R., Zahedi, R. P. & Borchers, C. H. Multiple reaction monitoring-mass spectrometry enables robust quantitation of plasma proteins regardless of whole blood processing delays that may occur in the clinic. *Mol. Cell. Proteom.* **21**, 10012. https://doi.org/10.1016/j.mcpro.2022.100212 (2022).
14. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLOS Biol.* **8**, e1000412. https://doi.org/10.1371/journal.pbio.1000412 (2010).
15. Coscia, F. *et al.* A streamlined mass spectrometry–based proteomics workflow for large-scale FFPE tissue analysis. *J. Pathol.* **251**, 100–112. https://doi.org/10.1002/path.5420 (2020).
16. García-Vence, M. *et al.* Protein extraction from FFPE kidney tissue samples: A review of the literature and characterization of techniques. *Front. Med.* **2021**, 8. https://doi.org/10.3389/fmed.2021.657313 (2021).
17. Deutsch, E. W. *et al.* The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res.* **51**(2023), D1539–D1548. https://doi.org/10.1093/nar/gkac1040 (2023).
18. MacLean, B. *et al.* Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968. https://doi.org/10.1093/bioinformatics/btq054 (2010).
19. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048 (2021).
20. Gorres, K. L. & Raines, R. T. Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* **45**, 106–124. https://doi.org/10.3109/10409231003627991 (2010).
21. Jessen, H. *et al.* Turnover of type I and III collagen predicts progression of idiopathic pulmonary fibrosis. *Respir. Res.* **22**, 1. https://doi.org/10.1186/s12931-021-01801-0 (2021).
22. Buczak, K. *et al.* Spatially resolved analysis of FFPE tissue proteomes by quantitative mass spectrometry. *Nat. Protoc.* **15**, 2956–2979. https://doi.org/10.1038/s41596-020-0356-y (2020).
23. Gilkes, D. M. *et al.* Collagen prolyl hydroxylases are essential for breast cancer metastasis. *Cancer Res.* **73**, 3285–3296. https://doi.org/10.1158/0008-5472.CAN-12-3963 (2013).

24. Tsitoura, E. *et al.* Collagen 1a1 expression by airway macrophages increases in fibrotic ILDs and is associated with FVC decline and increased mortality. *Front. Immunol.* **12**, 586. https://doi.org/10.3389/fimmu.2021.645548 (2021).
25. Kang, Y. P. *et al.* Metabolic profiling regarding pathogenesis of idiopathic pulmonary fibrosis. *J. Proteome Res.* **15**, 1724. https://doi.org/10.1021/acs.jproteome.6b00156 (2016).

Author contributions

N.R. and D.O. wrote the manuscript and designed all experiments. T.W. performed all development work related to the targeted LC–MS/MS assay and contributed to experimental design. M.K. provided lung tissue samples as well as aiding in interpretation of IPF and non-IPF regions of tissue sections. J.B. provided reindeer antler and back skin tissue sections as well as scRNA seq data to aid in interpretation of the LC–MS/MS method. A.D. performed the untargeted proteomic analysis, isolating target peptides for Skyline analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-68377-9.

Correspondence and requests for materials should be addressed to D.J.O.