

Identification of novel collagen breakdown products by human osteoclasts in vitro and in vivo

Brendan P. Norman^{1,*} , Jane P. Dillon¹, Sahem M. Alkharabsheh², Francoise Congues¹, Peter J.M. Wilson¹, Andrew S. Davison³ , Elinor Chapman⁴, James Baker⁴, Séamus Coyle^{5,6}, Chris Probert⁴, Paul E. Beaulé⁷, Lakshminarayan R. Ranganath^{1,3}, Jeremy M. Wilkinson⁸, James A. Gallagher¹

¹Department of Musculoskeletal and Ageing Science, Institute of Life Course and Medical Sciences, University of Liverpool, West Derby Street, Liverpool, L7 8TX, United Kingdom

²Department of Medical Laboratory Sciences, Faculty of Allied Medical Sciences, Mutah University, Alkarak 61710, Jordan

³Department of Clinical Biochemistry and Metabolic Medicine, Liverpool Clinical Laboratories, Royal Liverpool University Hospitals Trust, Mount Vernon Street, Liverpool, L7 8YE, United Kingdom

⁴Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Crown Street, Liverpool, L69 7BE, United Kingdom

⁵Liverpool Head and Neck Cancer Centre, University of Liverpool, West Derby Street, Liverpool, L7 8TX, United Kingdom

⁶Department of Palliative Medicine, Clatterbridge Cancer Centre, Pembroke Place, Liverpool, L7 8YA, United Kingdom

⁷Department of Surgery, The Ottawa Hospital, Smyth Road, Ottawa, Ontario K1H 8L6, Canada

⁸Department of Oncology and Metabolism, University of Sheffield, Beech Hill Road, Sheffield, S10 2RX, United Kingdom

*Corresponding author: Brendan P. Norman, Department of Musculoskeletal and Ageing Science, Institute of Life Course & Medical Sciences, University of Liverpool, William Henry Duncan Building, 6 West Derby Street, Liverpool L7 8TX, United Kingdom (bnorman@liverpool.ac.uk).

Abstract

Bone resorption involves dissolution of minerals and enzymatic degradation of bone matrix. The primary enzyme is cathepsin K but other proteases including matrix metalloproteinases are involved. Some cathepsin K cleavage products have been partially identified, including cross-linked telopeptides of type I collagen. Here, we aimed to characterize the entire complement of bone breakdown products resulting from osteoclast action under controlled conditions in vitro. We analyzed extracellular media from human osteoclasts cultured on dentin substrates, using untargeted liquid chromatography mass spectrometry. We discovered 22 breakdown products resulting from osteoclastic action. These products were peptide fragment sequences that mapped to various collagen proteins present in bone and dentin. Nine peptide fragments mapped exclusively to collagen I alpha-1 chain (COL1A1), the most abundant protein in bone. We subsequently detected 21 of the fragment products, initially observed in vitro, in human serum and/or urine. Consistent positive correlations were observed between the COL1A1-specific peptide fragments and established bone biochemical markers in serum and urine. Ten urine fragments and two serum fragments markedly increased ($p < .05$) following total hip arthroplasty, capturing the transient local peri-prosthetic osteolysis observed in these patients (serum, $n = 86$ patients; urine, $n = 83$ patients). Among these candidate osteolytic markers, four (two COL1A1-specific products) showed decreases from baseline ($p < .05$) in patients on denosumab ($n = 10$ patients). Additionally, two fragment peptides were higher ($p < .05$, fold change > 2) in urine from patients with bone metastasis (24 out of 112) among a lung cancer cohort. The range of collagen peptide fragments we discovered as a direct result of osteoclast activity indicates a complexity of bone resorption pathways not previously known, extending beyond the known proteolytic cleavage events in bone collagen proteins. Monitoring biofluid concentrations of these novel bone markers has the potential to capture multiple pathways of bone resorption activity beyond the existing assays based on cathepsin K.

Keywords: osteoclasts, resorption, peptide fragments, collagen

Lay Summary

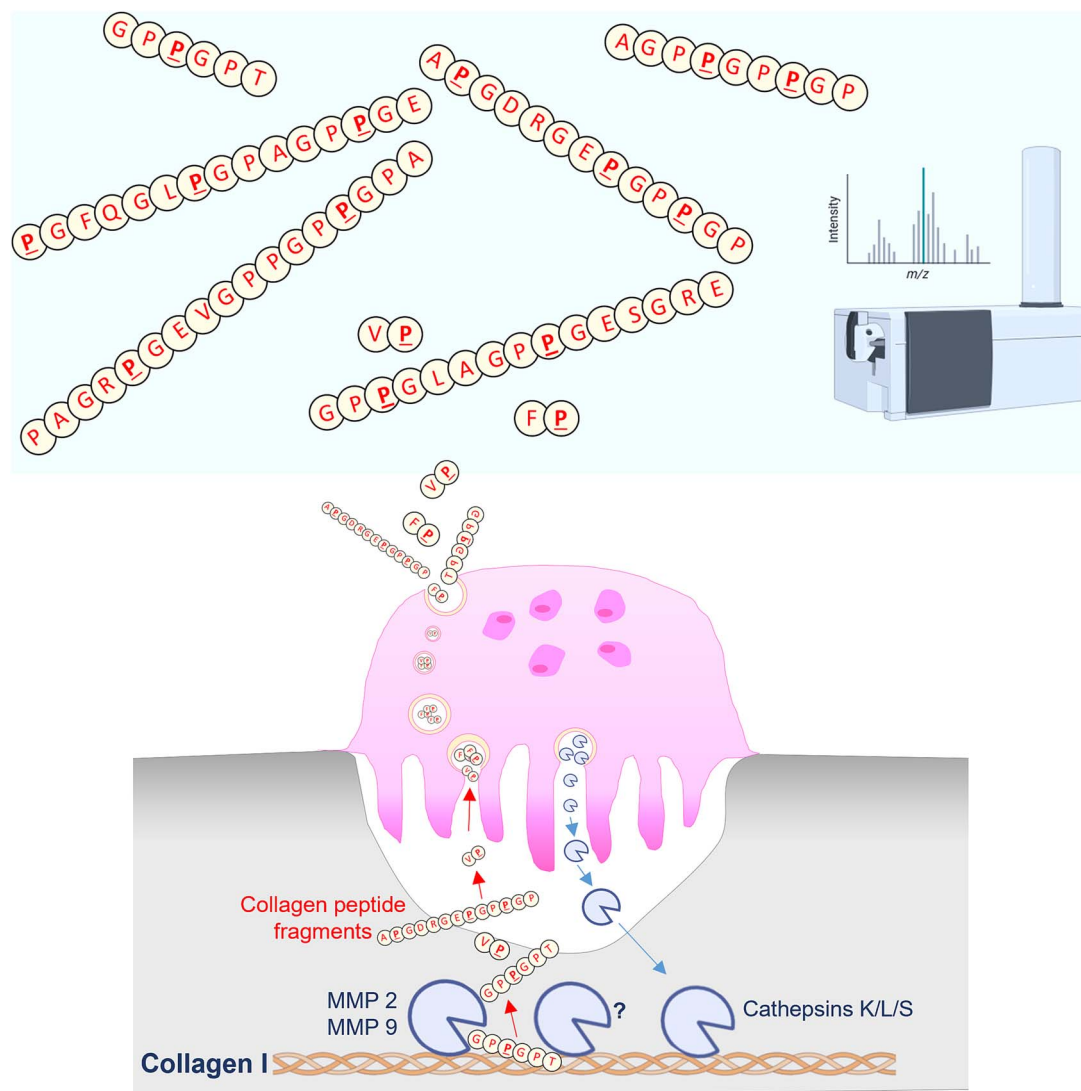
Here, we aimed to characterize the entire complement of bone breakdown products produced by specialized bone-degrading cells (osteoclasts). Analysis of osteoclasts cultured on bone-like wafers in the laboratory revealed 22 breakdown products, including newly described fragment molecules specific to collagen-I, the most abundant protein in bone. We subsequently showed that 21 out of 22 fragment molecules were present in human blood and/or urine. A subset of these molecules was higher in urine from patients with known bone loss following hip replacement or bone cancer. These findings enhance the understanding of bone resorption and suggest new methods for monitoring bone disorders through measuring specific protein fragments.

Received: January 10, 2025. Revised: August 27, 2025. Accepted: September 25, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of the American Society for Bone and Mineral Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical Abstract



Introduction

Osteoclasts are the principal cells that break down bone. Osteoclast action is fundamental to bone homeostasis, with an essential role in maintenance, repair, and remodeling of bone tissue. Osteoclasts resorb bone by producing various proteolytic enzymes and secreting hydrogen ions into the sealed extracellular compartment between the osteoclast and bone tissue matrix where resorption occurs.¹ Bone resorption results in the breakdown of bone matrix proteins. The predominant protein in bone is type I collagen, which constitutes 90% of the organic component of bone matrix. Self-assembly of the 2 type I collagen chains, the alpha-1 (I) and alpha-2 (I) molecules, results in a protein structure comprising a triple helix with a coiled conformation.²

Specific degradation products of type I collagen produced during bone resorption by osteoclasts are well-established markers of bone turnover. Following release of these products from the tissue matrix, they enter the circulation and are then excreted in urine. Quantification of type I collagen breakdown products in biofluid compartments provides a noninvasive readout of bone resorption activity. This is the

basis of several biochemical assays employed routinely in clinical management of bone disorders. The existing biochemical bone resorption markers produced by the enzyme cathepsin K do not represent the entirety of bone breakdown pathways by osteoclasts. Multiple studies have shown that even in healthy subjects, human serum and urine contain a range of peptide products derived from collagens, particularly collagen I alpha-1 chain (COL1A1).^{3–8} The variety of urine collagen fragments indicates that the protease cleavage sites in type I collagen I are extensive and not limited to the well-characterized cathepsin K cleavage sites.⁵ It is likely that some type I collagen fragments are the products of other proteolytic enzymes. For example, a group of non-cysteine protease enzymes expressed in bone and known to act on collagen is the matrix metalloproteinases (MMPs).⁹

A recent study demonstrated the potential of an approach based on proteolysis of bone under controlled conditions in the laboratory, combined with high-resolution mass spectrometry (MS), for characterizing the global breakdown pathways of collagens.¹⁰ Digestion of human type I collagen with purified cathepsin K resulted in the production of more than 100

peptide products originating from each of the COL1A1 and COL1A2 proteins. These products included peptides from known cathepsin K cleavage sites, in addition to a range of previously uncharacterized peptides. This indicates new cleavage sites not only limited to the cross-linking telopeptide domains but also including the helical portion of these proteins. The degree to which these products reflect osteoclastic resorption as it occurs *in vivo*, or if they are present in human blood or urine, is not known.

To provide a more physiologically representative view of osteoclast resorption by proteases not limited to cathepsin K, we studied the resorption products generated by activated osteoclasts cultured on dentin substrates. The organic matrix component of dentin, like bone, is comprised predominantly of type I collagen.¹¹ We then established the extent to which these identified breakdown products were observed in human serum and urine. Understanding the complexity of bone resorption pathways is a fundamental knowledge in basic bone biology and has potential to inform new approaches for monitoring and treatment of disorders associated with dysregulation of bone remodeling.

Materials and methods

Chemicals and reagents

Human osteoclast precursors, culture medium, and medium supplements were from Lonza. Base medium was OCP basal medium. Supplements and growth factors were from the OCP growth medium SingleQuots kit. Final growth medium composition was 10% fetal bovine serum, 2 mM L-glutamine, 100 units/mL penicillin/streptomycin, 33 ng/mL macrophage colony-stimulating factor (M-CSF), and 66 ng/mL receptor activator of nuclear factor kappa-beta ligand (RANKL).

For liquid chromatography mass spectrometry (LC/MS) analysis, deionized water was purified in-house by DIRECT-Q 3UV water purification system (Millipore). Methanol, isopropanol (Sigma-Aldrich), formic acid (Biosolve), and ammonium formate (Fisher Scientific) were LC/MS grade. Reference mass correction solution was prepared in 95:5 methanol:water containing 5 mmol/L purine (CAS No. 120-73-0) and 2.5 mmol/L hexakis (1H, 1H, 3H-tetrafluoropropoxy)phosphazine (HP-0921, CAS No. 58943-98-9) (Agilent).

In vitro dentin resorption assay

Dentin disks were prepared from hippo and walrus tusks. Disks were placed into wells of a 96-well plate. Human osteoclast precursor cells were cultured and differentiated into osteoclasts according to manufacturer guidelines. Osteoclast precursors were seeded into 96-well culture plates onto hippo dentin, walrus dentin, or plastic at a density of 10 000 precursor cells per well in 100 μ L growth medium containing M-CSF and RANKL for osteoclast generation. At day 7 and 14 of culture, 50 μ L (50%) and 100 μ L (100%) of medium, respectively, from each well was replaced with fresh growth medium. After 19 d, culture medium was harvested and stored at -80°C until analysis. Dentin wafers were fixed in phosphate buffered saline containing 10% formalin, stained with toluidine blue, and examined using reflected light microscopy to quantify resorption lacunae. Total resorption area for each dentin disk was quantified using an Olympus BH2 microscope ($\times 10$ objective) fitted with a source of incident light and a drawing tube to enable

projection of a grid over the visual field. Resorption area was determined with aid of the projected grid using the point counting technique previously described.¹² Non-resorptive control conditions were osteoclast precursors cultured onto plastic disks (+RANKL/–dentin; $N=6$) or without RANKL (–RANKL/+dentin; $N=6$, 3 walrus disks, 3 dentin disks), or onto plastic and without RANKL (–RANKL/–dentin; $N=3$). Additional +RANKL/+dentin cultures were treated with 10 fM–100 μ M zoledronic acid (6 concentrations, $N=3$ per concentration) at day 2, to examine potential effects of anti-resorptive treatment.

LC/MS analysis

Extracellular media samples were thawed at room temperature then diluted 1:3 media: deionized water (v/v) in individual 1.5 mL microcentrifuge tubes. Sample tubes were vortexed for 10 s then centrifuged at $1500 \times g$ for 5 min. Fifty microliter of supernatant was transferred to a 150 μ L 96-well plate (Agilent) for analysis. Pooled samples were created for quality control (QC) purposes for each treatment condition in addition to an overall QC by pooling equal volumes of individual samples. Pooled samples were prepared in an identical way to individual samples.

Human serum and urine samples (see “Human metabolomic data” section below) were stored at -80°C until preparation. Urine samples were analyzed following 1:3 (v/v) dilution urine: deionized water in 700 μ L 96-well polypropylene sample collection plates (Waters). Serum samples were analyzed following deproteinization with ice-cold methanol. Two-hundred microliters of serum was added to 600 μ L methanol contained in 1.5 mL microcentrifuge tubes on ice, then vortexed for 10 s and centrifuged at $16\,000 \times g$, 4°C to precipitate protein. One-hundred microliters of supernatant extract was transferred to a 700 μ L 96-well polypropylene sample collection plate (Waters) then methanol extracts were dried at room temperature under compressed air (<1 h). Dried extracts were stored at -80°C until analysis, at which point they were reconstituted in 200 μ L deionized water and agitated on a plate shaker (MTS 2/4 m IKA, Germany) at 800 rpm for 10 min before analysis. Pooled samples were created for human samples in each study for QC purposes, as described above.

Sample analysis was performed using a published liquid chromatography quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) acquisition method, which employed a 1290 Infinity II HPLC coupled to a 6550 QTOF-MS equipped with dual AJS electrospray ionization source (Agilent).¹³ Sample injection volume was 10 μ L for culture medium, 1 μ L for urine, and 2 μ L for serum. Reversed-phase LC was performed on an Atlantis dC18 column (3×100 mm, 3 μ m, Waters) maintained at 60°C with mobile phase flow rate 0.4 mL/min. Mobile phase composition was (A) water and (B) methanol, both with 5 mmol/L ammonium formate and 0.1% formic acid. The elution gradient started at 5% B 0–1 min and increased linearly to 100% B by 12 min, held at 100% B until 14 min, then at 5% B for a further 5 min. MS data acquisition was performed in positive ionization polarity with mass range 50–1700 in 2 GHz mode with acquisition rate at 3 spectra/s. A reference mass correction solution was continually infused at a flow rate of 0.5 mL/min via an external isocratic pump (Agilent) for constant mass correction (see preparation of reference mass correction solution). Capillary and fragmentor voltages were 400 and 380 V, respectively. Desolvation gas

temperature was 200 °C with flow rate at 15 L/min. Sheath gas temperature was 300 °C with flow rate at 12 L/min, and nebulizer pressure was 40 psi and nozzle voltage 1000 V. Reference ions monitored were: purine (m/z 121.0509) and HP-0921 (m/z 922.0098).

The LC/MS analytical sequence in each study was designed following published guidance.¹⁴ Injection order of individual samples was randomized computationally using Microsoft Excel. Pooled QC samples were interspersed throughout each analytical sequence, every 10th injection. Samples from each study were analyzed in one analytical batch except for the total hip arthroplasty study; sample preparation and analysis split over two representative batches.

Compound fragmentation spectra for resulting analytes of interest from the in vitro study were obtained from MS2 analysis of media samples with accurate mass precursor ion targets detailed in Table 1; no more than 6 compounds targets per injection. Multiple fixed collision energies were applied; 10, 20, and 40 eV. Acquisition rates were 6 spectra/s in MS1 and 4 spectra/s in MS2.

Human metabolomic data

Human serum and urine untargeted metabolomic datasets were acquired under identical analytical conditions to those described above for in vitro experiment samples.

Serum and urine data were acquired on samples obtained from 2 clinical trials conducted in patients undergoing total hip arthroplasty. One study compared 2 femoral prosthesis designs in patients undergoing total hip arthroplasty for hip osteoarthritis at The Ottawa Hospital, Ottawa, Canada between 2013 and 2017.¹⁵ The conventional prosthesis was associated with a smaller reduction in peri-prosthetic bone mineral density loss compared with a Tri-Lock Bone Preservation Stem femoral prosthesis. No statistically significant difference was observed between the two prosthesis designs in the profiles of established biochemical bone turnover markers serum CTX and P1NP. Serum and urine (total serum samples, $n = 334$; total urine samples, $n = 374$) were collected at five time points across the study: baseline (pre-surgery) then at 3, 6, 12, and 24 mo post-surgery. Serum samples were from 86 patients (40 male, 46 female, mean age [$\pm$ SD] 59 ± 11 yr) and urine samples were from 83 patients (39 male, 44 female, mean age 59 ± 11 yr). The second study was a phase 2 randomized, double-blind, placebo-controlled trial carried out at Sheffield Teaching Hospitals, Sheffield, UK between 2012 and 2018 to investigate the effect of the monoclonal antibody denosumab on osteolytic activity in patients listed for revision total hip arthroplasty surgery. The denosumab group showed 83% fewer osteoclasts at the osteolysis membrane-bone interface compared with the placebo group.¹⁶ Serum and urine (total serum samples, $n = 89$; total urine samples, $n = 94$) were collected from patients immediately before a single subcutaneous injection of 60 mg denosumab or placebo control (study visit 2), then again at 4 (study visit 3) and 8 wk (study visit 4) on treatment. Serum samples were from 21 patients: 10 in the denosumab group (6 male, 4 female, mean age 72 ± 12 yr) and 11 in the placebo group (9 male, 2 female, mean age 71 ± 9 yr). Urine samples were from 22 patients: 10 in the denosumab group (6 male, 4 female, mean age 72 ± 12 yr) and 12 in the placebo group (10 male, 2 female, mean age 72 ± 9 yr). Serum and urine from both total hip arthroplasty sample sets were collected following complete overnight fast. In the UK study, patients had not received

bisphosphonates (no oral bisphosphonate therapy within 12 mo of the study, no intravenous bisphosphonates within 5 yr of the study), and in the Canadian study, patients had no past or present use of drugs known to affect bone metabolism.

Additional urine metabolomic data were from a study investigating the metabolic changes that occur during the dying process in patients with incurable lung cancer (patients $n = 112$).¹⁷ This dataset is available at the NIH Common Fund's National Metabolomics Data Repository website, the Metabolomics Workbench (<https://www.metabolomicsworkbench.org/>); Project ID PR001322, directly accessible via Project DOI: [10.21228/M8TT44](https://doi.org/10.21228/M8TT44). Serial urine samples were collected from patients at various hospitals/hospices in northwest England between 2016 and 2018 at various time points leading up to death (>12 to <1 wk before death); total samples $n = 233$. Among the 112 patients in the urine study, 24 patients (12 male, 12 female, mean age 69 ± 9 yr) had bone metastasis confirmed by CT, PET, or bone scintigraphy; the other 88 patients (55 male, 33 female, mean age 70 ± 10 yr) had lung cancer but without known bone metastasis.

Research on these human datasets was performed under the following ethical approvals. Total hip arthroplasty datasets: approval by Ottawa Hospital Institutional Review Board (OHSN-REB 2010913-01H) and registered with clinicaltrials.gov (NCT01558752); for the denosumab trial, approval by National Research Ethics Service Committee Yorkshire and the Humber-Leeds West (REC reference 11/YH/0252) and registered with the EU Clinical Trials Register (EudraCT 2011-000541-20). Biology of dying urine dataset: approval provided by North Wales (West) Research Ethics Committee (REC reference 15/WA/0464).

Data processing and statistical analysis

Raw LC/MS data were mined for compound features in Masshunter Profinder (build 10.00, Agilent). Data from in vitro media samples were mined by an untargeted approach using the batch recursive (small molecules/peptides) feature extraction function. Untargeted feature extraction parameters: peak height >5000 counts, allowed ion species +H, +Na, +K, and +NH₄, alignment retention time (RT) tolerance ± 0.3 min, alignment mass tolerance ± 20 ppm, molecular feature extraction score >70 , and Agile 2 peak integration method. Compounds not detected in all individual samples across at least one treatment condition were removed from the analysis.

Human metabolomic data were mined for the oligopeptide features identified from the in vitro experiment using the targeted feature extraction function in Profinder. Feature extraction parameters were accurate mass match window ± 5 ppm and RT window ± 0.3 min. Allowed ion species were +H, +Na, +K and +NH₄. "Find by formula" filters were: score >60 in at least 60% of samples from at least one sample group.

Extracted peak area intensity data were exported in .csv file format and imported into Mass Profiler Professional (MPP; build 15.1, Agilent). In MPP, all data were log₂ transformed and pareto scaled. Urine data from the total hip arthroplasty studies were normalized to total peak area, ascertained from untargeted metabolite feature extraction (Mass Profiler; version B.07.02, build 123.0, Agilent). Urine data from the biology of dying study were normalized to urine creatinine concentration. QC of extracted features was performed on all datasets as previously described, based on

reproducibility of signals across replicate injections of each pooled sample; features present in 100% of injections for at least one pooled sample, and with coefficient of variation (CV) <30% across replicate injections were retained in the analysis.¹³ Statistical comparisons between in vitro experiment groups were performed in IBM SPSS (version 31) by Kruskal-Wallis test and post-hoc analysis by Dunn's test, both with Benjamini-Hochberg false-discovery rate (FDR) multiple testing correction. Potential longitudinal differences in the abundance of urine and serum compounds in the total hip arthroplasty datasets were assessed in MPP by repeated-measures ANOVA and post-hoc analysis by pairwise repeated measures *t*-tests (versus baseline visit), both with Benjamini-Hochberg FDR adjustment. In the lung cancer urine dataset, volcano plot analysis combining fold-change (FC) and *T*-test analyses were performed in MetaboAnalyst (version 6.0)¹⁸ to examine potential differences in urine compounds between patients with versus without bone metastasis.

Compound structure identification

Compound identification was performed using multiple chemical properties; accurate mass and MS2 fragmentation spectra of extracted features, in addition to the known amino acid sequences of the major extracellular matrix proteins present in dentin. Extracellular proteins mapped to were collagens 1, 2, 3, and 5; COL1A1, COL1A2, COL2A1, COL3A1, COL5A1, and COL5A2. Reference amino acid sequences for these proteins, with hydroxyproline and hydroxylysine post-translational modifications, were obtained from the NCBI protein database (National Library of Medicine; National Institutes of Health; U.S. Department of Health and Human Services) with the following accession numbers: NP_000079.2 (COL1A1), NP_000080.2 (COL1A2), NP_001835.3 (COL2A1), NP_000081.2 (COL3A1), NP_000084.3 (COL5A1), NP_000384.2 (COL5A2). The theoretical monoisotopic neutral accurate masses of all possible oligopeptide chain sequences 1-20 amino acids long for these proteins were mapped out in Microsoft Excel. Matching of calculated neutral masses of observed features was performed against these sequence mass databases generated for each candidate protein (mass match window ± 5 ppm).

Results

Resorption of dentin disks in vitro

Extensive resorption of dentin disks was observed in cell cultures grown on dentin substrate with activation of osteoclast activity by RANKL supplementation in growth medium (+dentin/+RANKL). In these cultures, resorption pits covered >60% of the dentin surface area (Figure 1A). No difference was observed in extent of resorption between hippo and walrus dentin. Resorption of dentin substrates was not observed in control cultures without RANKL supplementation.

Oligopeptide markers of in vitro dentin resorption in extracellular media

Untargeted feature extraction of LC/MS data acquired from extracellular media resulted in 153 positively charged ion features across all samples. QC filtering resulted in 148 features present in all replicate injections for at least one pooled group sample; 147 of these features passed additional variability filtering, with CV <25% across replicate injections

of each pooled sample. Thirty-one features had increased abundance ($p < .05$, fold change >2) in resorption-positive +dentin/+RANKL osteoclast cultures compared with all control cultures. Twenty-two of these features were identified as potential resorption markers following additional QC checks on raw data for visualization of feature integration quality and removal of duplicate extracted features (Table 1). Among these chemical entities, 21 compounds were exclusively present in +dentin/+RANKL cultures (Figure 1B and F), except for glycyl-prolyl-hydroxyproline, which was also detected in control cultures at lower abundance; raw peak area mean ($\pm$ SD) = 364 326 ($\pm$ 73 573) in +dentin/+RANKL, mean ($\pm$ SD) = 45 755 ($\pm$ 15 310) across control cultures (Figure S1).

The abundance of 21 out of 22 resorption compounds showed a clear effect of zoledronic acid treatment in +dentin/+RANKL cultures; $p < .05$. There was strong concordance between the profiles of these compounds, with a consistent sharp decline at zoledronic acid concentration $\geq 1 \mu\text{M}$ (Figure 1C and D). Post-hoc analysis confirmed this trend, with significant differences ($p < .05$ or $< .01$) between the 100 μM zoledronic acid groups compared with untreated control for 21 out of 21 zoledronic acid responsive compounds. Significantly lower signal ($p < .05$) was also observed at 1 μM zoledronic acid, versus control, for 11 of these 21 compounds. Strong positive correlations were observed between peptide fragment abundance and the quantified resorption area of dentin disks across the zoledronic acid treatment groups (mean $\pm$ SD, $R^2 = 0.72 \pm 0.11$; Figure 1E). Inhibition with zoledronic acid treatment was not observed for the compound identified as the dipeptide His-Hyp (or Hyp-His; m/z 251.1142) despite not being detected in the -dentin/+RANKL, +dentin/-RANKL, or -dentin/-RANKL controls (Figure S2).

Chemical identification and mapping oligopeptide fragments to matrix proteins

With exception of one compound (compound #13, Table 1), all compounds mapped to one or more oligopeptide amino acid sequence present in the extracellular matrix proteins COL1A1, COL1A2, COL2A1, COL3A1, COL5A1, or COL5A2 (Table 1, Table S1). Eleven of the 22 compounds were identified as various di-peptides and mapped to multiple sequences across the proteins examined, except for three dipeptides that were unique to COL1A1. Nine of the 22 compounds mapped exclusively to the COL1A1 protein. These COL1A1-specific oligopeptides ranged from 2 (Val-Hyp and Phe-Hyp) to 17 amino acids. The five highest MW products were all COL1A1 specific and mapped uniquely to single sequences within this protein (Figure 2). These COL1A1-specific compounds were doubly charged and contained a minimum of one Gly-Pro-Hyp sequence. Four compounds were identified as products of dipeptides minus the accurate mass of H_2O . These molecules are unlikely to be technical artifacts caused by MS in-source fragmentation, as the chromatographic RT was different to their derivative complete dipeptides where these were observed.

Comparison of COL1A1 fragment peptide sequence locations to known COL1A1 cleavage sites

Known proteolytic cleavage sites in the COL1A1 protein were investigated to identify candidate collagenolytic enzymes with the potential to produce the COL1A1 peptide

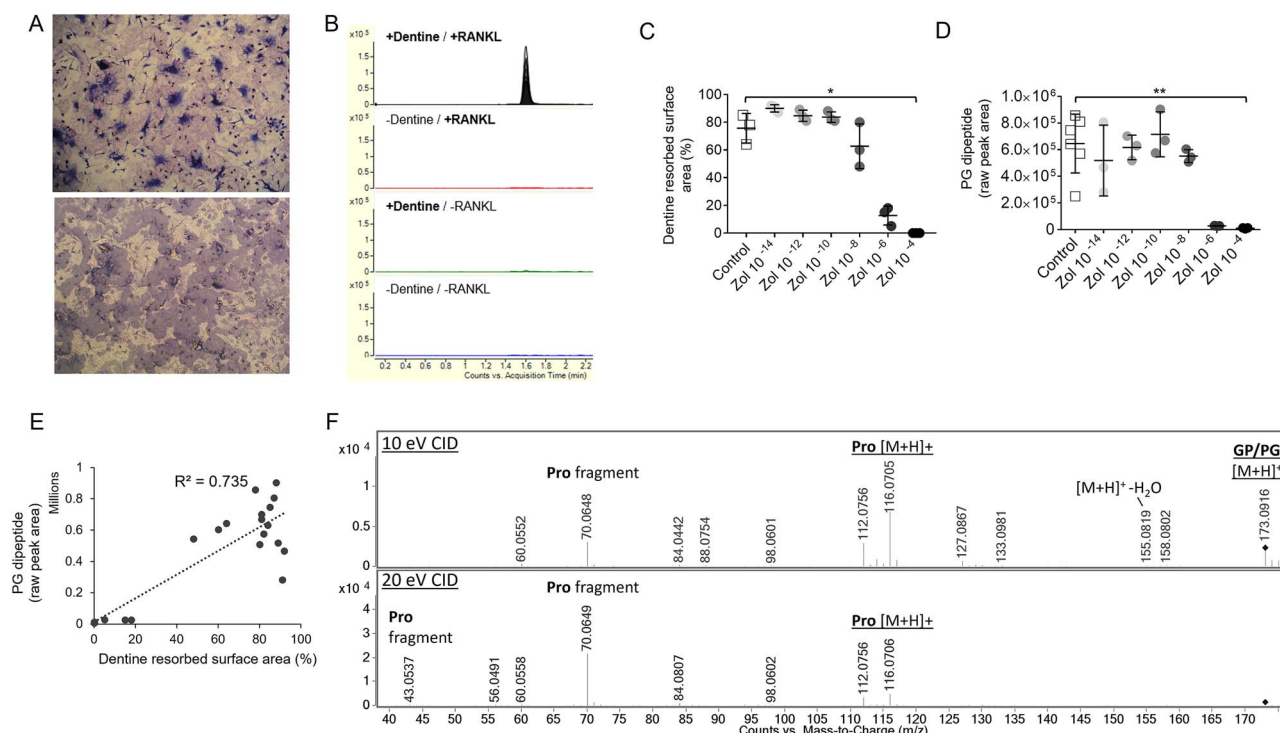


Figure 1. Observation of bone resorption in the dentin slice assay and identification of oligopeptide biochemical markers of bone matrix breakdown. (A, upper panel) Transmitted light photomicrograph showing osteoclasts resorbing dentin substrate. (Lower panel) reflected light photomicrograph showing resorption trails. (B) Representative chromatogram of an oligopeptide bone resorption marker from extracellular media sampled from the dentin slice assay. Note the exclusive presence of signal in cultures of RANKL-activated osteoclast precursors grown on dentin but no signal detected in any of the control conditions. Traces from individual samples are overlaid within each treatment group. (C, D) Inhibition of dentin resorption and peptide fragment release into extracellular media with zoledronic acid treatment 10^{-14} M (10 fM)– 10^{-4} M (100 μ M) in activated osteoclast cultures. Dot plots show inhibition of osteoclast resorption between 10^{-8} and 10^{-6} M zoledronic acid concentrations, indicated by sharp reduction in the total resorbed surface area of dentin disks (C) and abundance of oligopeptide fragments detected in media (D). * $p < .05$; ** $p < .01$. (E) Positive correlation between oligopeptide fragment signal and total resorbed surface area of dentin slices across zoledronic acid treated cultures 10^{-14} M (10 fM)– 10^{-4} M (100 μ M). (F) Example MS/MS fragmentation spectra for peptide fragment at multiple collision energies, showing the intact molecular $[M + H]^+$ cation and MS peaks representing constituent substructures, including the individual amino acid proline. The representative peptide fragment product shown in sections (B–F) is the dipeptide comprised of proline and glycine. Abbreviation: CID = collision-induced dissociation.

fragments discovered. Using the MEROPS database of proteolytic enzymes,¹⁹ COL1A1 was selected as a substrate protein and all reported cleavage sites across all reported COL1A1-targeting peptidases were identified. These peptidases included cathepsins D, L, K, and S and the MMPs. All COL1A1 cleavage sites reported from any form of evidence in this database were considered. All our identified COL1A1 peptide fragments greater than 3 amino acid residues were considered in this analysis, except for the di-peptides Phe-Hyp and Val-Hyp which were also considered. These fragments, and particularly the longer sequences, were generally more specific to COL1A1.

Table 1 shows the predicted proteases associated with each peptide fragment based on overlap between their COL1A1 mapping positions and all known COL1A1 protease cleavage sites provided in MEROPS. The COL1A1 protein sequence positions of four of our peptide fragments overlapped with a known COL1A1 proteolytic cleavage site. These known cleavage sites are targeted by the collagenolytic enzymes cathepsins K, L, and S, MMP2 and MMP9. Figure 2 shows the COL1A1 sequence positions of the 9 peptide fragments that mapped exclusively to COL1A1. Only 1 out of 9 COL1A1-specific peptides overlapped with a known cleavage site; a 15-amino acid residue peptide fragment which overlapped with one cleavage site reported for MMP9 (compound #20;

see Table 1). None of our other observed peptide fragments greater than 5 amino acid residues overlapped a known COL1A1 cleavage site.

Novel oligopeptide markers in human serum and urine

Serum and urine untargeted metabolomic datasets acquired in our laboratory were mined for the aforementioned oligopeptide products. The total hip arthroplasty datasets provided the opportunity to assess longitudinal profiles of the oligopeptides in serum and urine among patient cohorts with characteristic alteration in established biochemical bone turnover markers resulting from (1) acute post-surgical periprosthetic osteolysis¹⁵ (Figure 3A) and (2) attenuation of local osteolytic lesions with denosumab¹⁶ (Figure 3B). We also mined urine data from a study designed to characterize the urine metabolite changes occurring in terminally ill lung cancer patients at various time points leading up to death.¹⁷ Among this cohort, 24 out of 112 patients had known bone metastasis, providing the opportunity to examine oligopeptide profiles in this bone disorder.

Twenty-one of the 22 peptide fragments, we discovered from in vitro experiments, were detected in human serum and/or urine; 11 fragments in serum, 20 fragments in urine

Table 1. Summary of bone resorption products observed from LC/MS analysis.

Compound ID	Proposed compound identification	Amino acid sequence	Proposed empirical formula	Human biofluids detected in	Possible originating proteins	Predicted proteases
1	Gly-Pro/Pro-Gly	GP/PG	C ₇ H ₁₂ N ₂ O ₃	Serum, urine	COL1A1, COL1A2, COL2A1, COL3A1 COL5A1, COL5A2	
2	Val-Gly/Gly-Val	VG/GV	C ₇ H ₁₄ N ₂ O ₃	Serum, urine	COL1A1, COL1A2, COL2A1, COL3A1 COL5A1, COL5A2	
3	Hyp-Ala(Ala)-Hyp (-H ₂ O)	PA/AP	C ₈ H ₁₂ N ₂ O ₃	Serum, urine	COL1A1, COL1A2, COL2A1, COL3A1, COL5A2	
4	Pro-Ser/Ser-Pro/Ala-Hyp	PS/SP/AP	C ₈ H ₁₄ N ₂ O ₄	Serum, urine	COL1A1, COL1A2, COL2A1, COL3A1 COL5A1, COL5A2	
5	Leu-Hyp/Hyp-Leu Ile-Hyp/Hyp-Ile (-H ₂ O)	LP/PL/IP/PI	C ₁₁ H ₁₈ N ₂ O ₃	Urine	COL1A1, COL2A1 COL3A1, COL5A1	
6	Val-Hyp	VP	C ₁₀ H ₁₈ N ₂ O ₄	Urine	COL1A1	
7	Leu-Hyp/Hyp-Leu	LP/PL/IP/PI	C ₁₁ H ₂₀ N ₂ O ₄	Serum, urine	COL1A1, COL2A1 COL3A1, COL5A2	
8	Ile-Hyp/Hyp-Ile Hyp-His/His-Hyp (-H ₂ O)	PH/HP	C ₁₁ H ₁₄ N ₄ O ₃	Urine	COL3A1, COL5A1	
9	Phe-Hyp (-H ₂ O)	FP	C ₁₄ H ₁₆ N ₂ O ₃	Serum	COL1A1	
10	Phe-Hyp (-H ₂ O)	FP	C ₁₄ H ₁₈ N ₂ O ₄	Serum, urine	COL1A1	
11	Gly-Pro-Hyp	GPP ^a	C ₁₂ H ₁₉ N ₃ O ₅	Serum, urine	COL1A1, COL1A2, COL2A1	Cathepsins K/L/S, MMP2, MMP9
12	Arg-Hyp	RP	C ₁₁ H ₂₁ N ₅ O ₄	Urine	COL1A1, COL1A2,	
13	Unknown (MW: 307)	-	-	Urine		
14	Gly-Pro-Hyp-Gly	GPPG ^a	C ₁₄ H ₂₂ N ₄ O ₆	Serum, urine	COL1A1, COL1A2, COL2A1	Cathepsins K/L/S, MMP2, MMP9
15	Multiple candidates	PGPPG/GPPGP/PGGPP/PGPPG ^a	C ₁₉ H ₂₉ N ₅ O ₇	Urine	COL1A1, COL1A2, COL2A1	
16	Gly-Pro-Hyp-Gly-Pro-Thr	GPPGPT	C ₂₃ H ₃₆ N ₆ O ₉		COL1A1	
17	Multiple candidates	PPGPPGP/GPPGPPG	C ₂₆ H ₃₉ N ₇ O ₁₀	Urine	COL1A1, COL1A2	
18	Multiple candidates ^c	AGPPGPPGP/GPPGPPGPA/PPGPPGPAG	C ₃₄ H ₅₁ N ₉ O ₁₂	Urine	COL1A1	
19	Multiple candidates ^c	APGDRGEPGPPGP/PGDRGEPPGP ^b	C ₅₁ H ₇₈ N ₁₆ O ₂₁	Serum, urine	COL1A1	
20	Multiple candidates ^c	GPPGLAGPPGESGRE/ ^a PPGLAGPPGESGREG ^a	C ₅₈ H ₉₂ N ₁₈ O ₂₃	Urine	COL1A1	MMP9
21	Hyp-Gly-Phe-Gln-Gly-Leu-Hyp-Gly-Pro-Ala-Gly-Pro-Hyp-Gly-Glu	PGFQGLPGAGPPGE ^a	C ₆₃ H ₉₂ N ₁₆ O ₂₂	Urine	COL1A1	
22	Pro-Ala-Gly-Arg-Hyp-Gly-Glu-Val-Gly-Pro-Gly-Pro-Hyp-Gly-Pro-Ala	PAGRPGEVGP ^a PPGPA	C ₆₇ H ₁₀₄ N ₂₀ O ₂₂	Serum, urine	COL1A1	

Proposed identifications are based on mapping of observed peptide fragments against all possible oligopeptide products obtained from amino acid sequences for mature collagen proteins present in dentin. Predicted proteases are those proposed to produce the observed oligopeptides based on overlap between our markers and known protease cleavage sites in the COL1A1 protein; octapeptide cleavage sites obtained for cathepsin proteases and matrix metalloproteinases (MMPs) from the MEROPS¹⁹ database. MW (neutral monoisotopic mass). Abbreviation: P = hydroxyproline. ^aCompound overlaps sequence position of a previously-reported urine COL1A1 peptide fragment⁵ (for fragments of >3 amino acid residues). ^bCompound exactly reproduces a previously-reported urine COL1A1 peptide fragment. ^cMultiple possible positional isomers originating from the same singular amino acid sequence in COL1A1 (see Figure 2).

P hydroxyproline post-translational modification
K hydroxylysine post-translational modification

Correlation analysis was performed on the denosumab dataset (data from all visits), as data from established bone

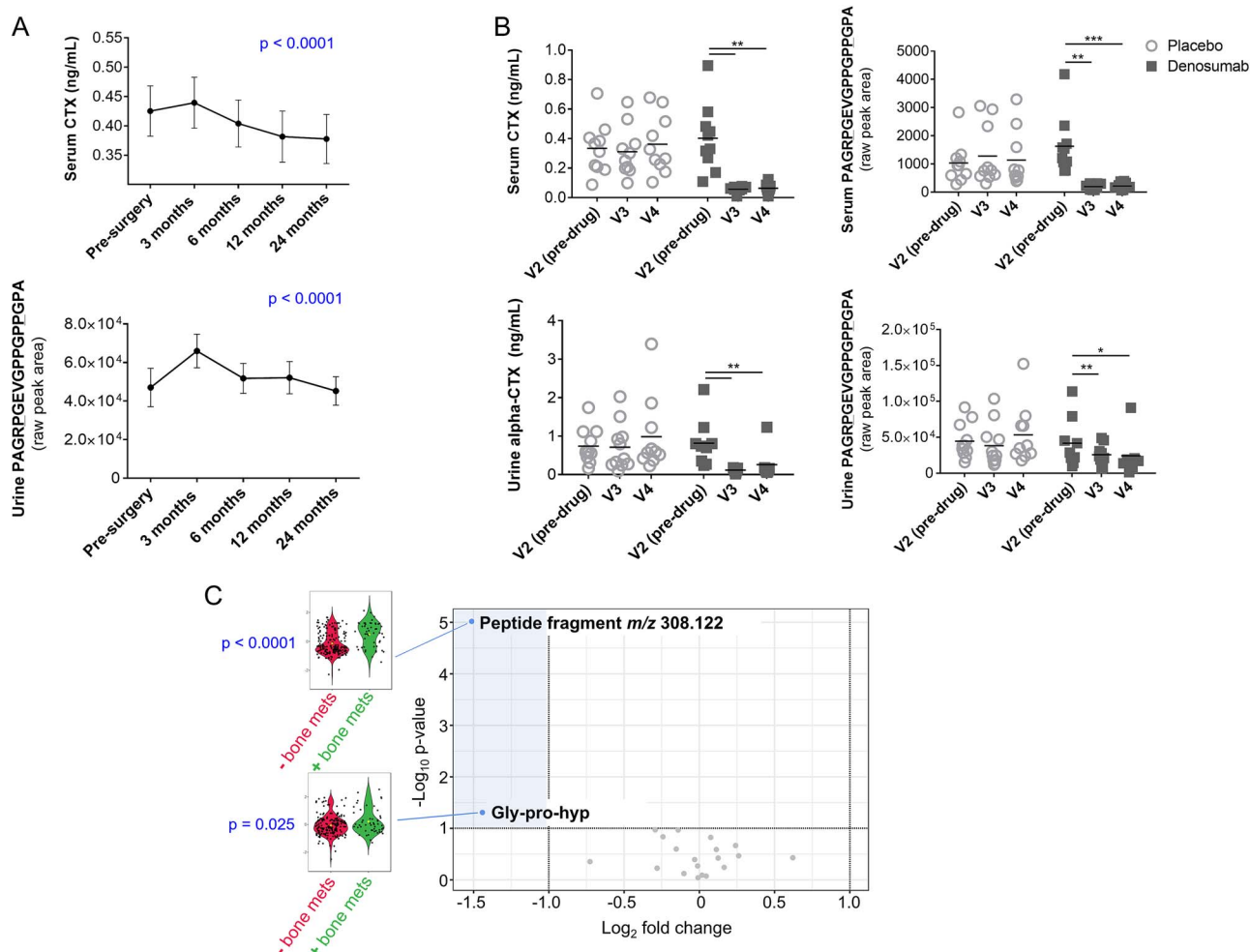


Figure 3. Oligopeptide products as potential markers of bone resorption and metastatic bone disease in humans. (A) Longitudinal abundance profiles of biochemical analytes from a total hip arthroplasty patient cohort.¹⁵ Serum and urine samples were taken at baseline (pre-surgery) then at 3, 6, 12, and 24 mo post-surgery. Analytes are serum CTX (data from Slullitell et al.¹⁵), and the largest COL1A1-specific peptide fragment identified in this study (compound 22, Table 1). Plots show data from all samples (serum, $n = 334$; urine, $n = 374$). p -values are from repeated measures one-way ANOVA on patients (serum, $n = 49$; urine, $n = 52$) with samples available for all time points, adjusted to false-discovery rate by Benjamini-Hochberg method. (B) Serum and urine biochemical abundance profiles in patients on denosumab/placebo at 4 (visit 3) and 8 wk (visit 4) on treatment versus baseline (visit 2). Compound 22 shown as an example peptide fragment with decreased abundance from baseline in patients on denosumab in serum and urine ($*p < .05$, $**p < .01$, $***p < .001$; Benjamini-Hochberg false discovery rate adjusted). This trend was in line with the published data on established bone markers in this cohort; decreases in serum and urine CTX on denosumab with no decrease in the placebo group.¹⁶ p -Values are from repeated measures ANOVA on patients (serum data, denosumab $n = 8$ patients, placebo $n = 11$ patients; urine data, denosumab $n = 7$ patients, placebo $n = 9$ patients) with samples available for all three time points. Dot plots show data from all samples analyzed (serum, $n = 58$; urine, $n = 58$). (C) Volcano plot highlighting the two oligopeptide compounds higher in urine from patients with bone metastasis among the lung cancer cohort (markers in shaded region indicating $p < .05$ and fold change > 2 between patients with bone metastasis versus controls). These compounds were gly-pro-hyp and an unidentified compound with m/z 308.122 (compound #13, Table 1). All remaining urinary oligopeptides detected showed no significant difference ($p > .05$ and fold change < 2) between patient groups (gray dots). $N = 24$ patients with lung cancer and bone metastasis (mets), $n = 88$ patients with lung cancer but without known bone metastasis.¹⁷

biochemical markers in serum and urine were available for this study. Pearson analysis showed positive correlations between 8 out of 11 serum peptide fragments and serum CTX ($p < .05$, correlation coefficient range = 0.21–0.71) (Figure 4). Out of the 21 urine peptide fragments detected, 12 compounds correlated ($p < .05$) with urine alpha-CTX (correlation coefficient range = 0.23–0.58) and 13 compounds correlated with beta-CTX ($p < .05$, correlation coefficient range = 0.27–0.61) (Figure 5). The largest, COL1A1-specific peptide fragments generally showed the most consistent correlations with established bone markers (Figures 4A and C and 5A and C); also including TRAP and P1NP in addition to serum and urine CTX. Table 2 summarizes the data from correlation analysis between all

peptide fragments and clinical bone biochemical markers investigated.

In our analysis of the lung cancer dataset, Gly-Pro-Hyp and an unidentified compound with m/z 308.122 (compound 13; Table 1) were more abundant ($p < .05$, fold change > 2) in urine from patients with bone metastasis; $n = 24$ patients with lung cancer and bone metastasis; $n = 88$ patients with lung cancer but without known bone metastasis (Figure 3E).

Discussion

Our data show culturing osteoclasts on dentin substrate provides a useful tool to model bone resorption as it occurs in vivo. The diversity in peptide fragments observed, including

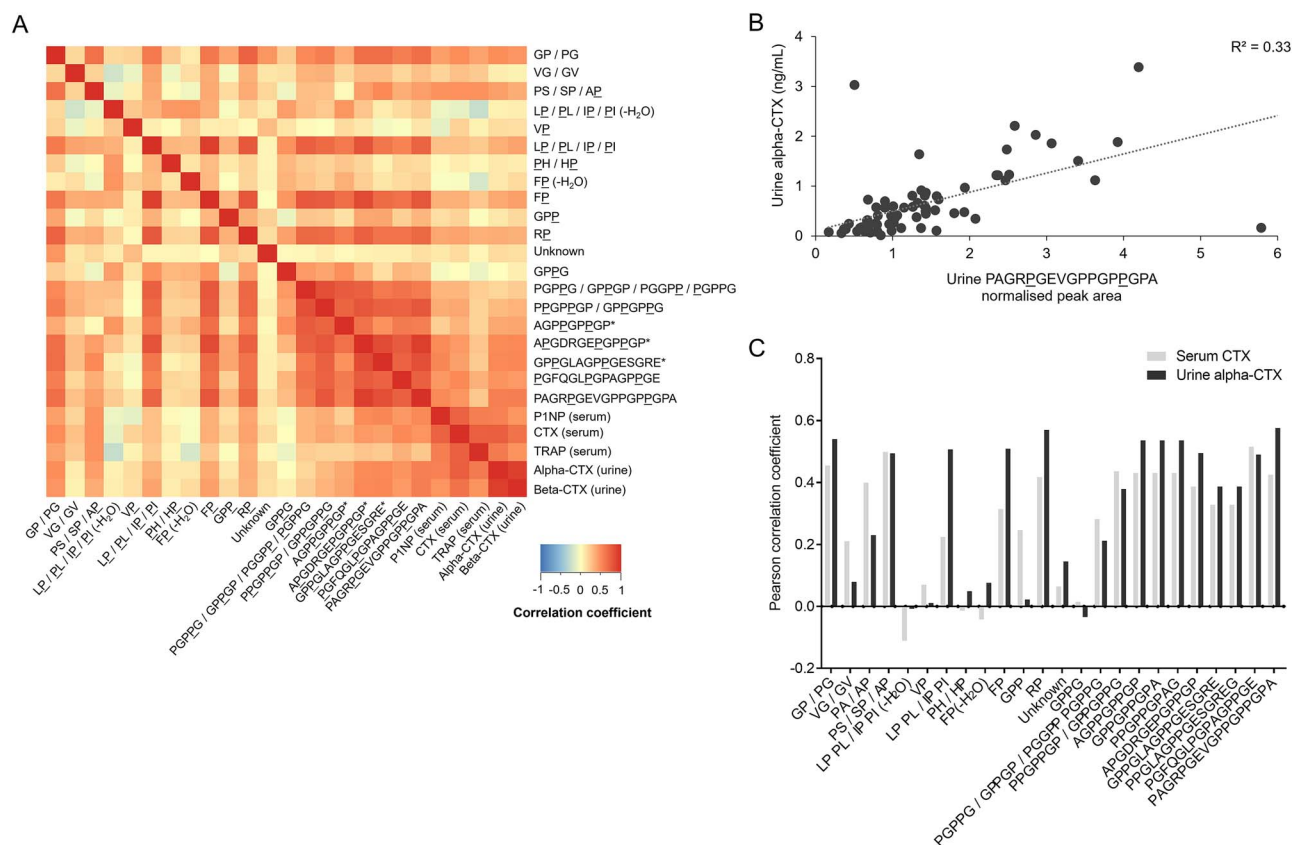


Figure 5. Correlation between urine collagen peptide fragments and established bone resorption markers. (A) Heatmap from Pearson correlation analysis between all collagen peptide fragments identified in urine and established biochemical bone markers. (B) Scatter plot showing correlation between urine alpha-CTX and the largest COL1A1-specific peptide identified in this analysis (compound 22, Table 1). (C) Summary of Pearson correlation coefficients between the collagen peptide fragments detected in urine and serum CTX and urine alpha-CTX. All data are from the denosumab clinical trial sample set ($n = 94$ samples).¹⁶

proteolysis, with early cleavage events after 1 h of digestion occurring preferentially at the telopeptide cross-linking domains at the extremities of the type I collagen molecule. With increasing digestion time, up to 16 h, cleavage occurred across the breadth of collagen, including the central helical domain, producing peptide fragments comprising fewer than 4 amino acid residues.¹⁰

Our data support biofluid collagen peptide fragments as biomarkers in disorders affecting bone. In the Canadian total hip arthroplasty cohort, DXA previously showed the expected post-surgical loss of femoral bone mineral density proximal to prosthesis insertion and associated with increased fracture risk.¹⁵ In these patients we observed concordance between the longitudinal profiles of oligopeptide fragments and serum CTX and P1NP, with marked increases in peptide fragment abundance at the time point immediately following surgery. In the denosumab study, we showed that positive correlations with established bone biochemical markers were generally the most consistent for the COL1A1-specific peptide fragments. The 5 largest fragments identified in this study in addition to the dipeptide Phe-Hyp, all COL1A1-specific, showed statistically significant correlations with serum P1NP, serum CTX, urine alpha-CTX, and urine beta-CTX (Table 2). These significant correlations held whether the peptide fragments were detected in serum or urine. The response of peptide fragments to denosumab is probably underestimated in this dataset due to the relatively small scale of the study, limiting

statistical power. Nevertheless, the data strongly support the largest peptide fragment identified in this analysis (compound 22, Table 1) as an excellent candidate for future exploration as a new bone biomarker. This molecule is COL1A1-specific, acutely increased immediately following total hip replacement (Figure 3A), decreased in serum and urine following denosumab treatment (Figure 3B) and correlates with all established bone biomarkers investigated in this study ($p < .05$; Figures 4B and 5B, Table 2), with exception of urine TRAP.

Our finding of elevated urine bone peptide fragments in bone metastasis is consistent with data on type I collagen markers in bone cancer.^{20–22} An osteolytic phenotype is frequently induced by presence of bone tumors, associated with stimulation of osteoclastogenesis through RANKL expression and cathepsin K activity.^{23–25} In our lung cancer cohort, we previously showed changes in bone peptide fragments in the last weeks of life.¹⁷ Two of the COL1A1-specific collagen peptides reported here progressively increased in urine leading up to death, from greater than 3 mo up to the last week before death; these were the dipeptide Phe-Hyp and the highest MW COL1A1 fragment observed here (compound #22; Table 1). Another unidentified bone marker (compound #13; Table 1) decreased leading up to death. These findings provide major insights into the fundamental biological changes that occur at the end of life, implicating divergent bone resorption pathways as part of a multifactorial process of dying.

Table 2. Summary of Pearson correlation analysis of collagen peptide fragments against established clinical biochemical bone markers.

Compound ID	Peptide fragment	Serum CTX		Serum TRAP		Serum PINP		Urine alpha-CTX		Urine beta-CTX	
		Serum	Urine	Serum	Urine	Serum	Urine	Serum	Urine	Serum	Urine
1	GPPG	0.413 ^b	0.455 ^b	0.262 ^a	0.390 ^b	0.310 ^b	0.514 ^b	0.222	0.540 ^b	0.256 ^a	0.493 ^b
2	VG/GV	0.364 ^b	0.210 ^a	0.241 ^a	0.083	0.161	0.143	0.155	0.079	0.106	0.070
3	PA/AP	0.099	0.399 ^b	-0.04	0.222 ^a	0.142	0.329 ^b	-0.013	0.230 ^a	-0.015	0.282 ^a
4	PS/SP/AP	-0.049	0.498 ^b	-0.125	0.483 ^b	-0.088	0.422 ^b	0.018	0.494 ^b	-0.022	0.408 ^b
5	LP/PL/IP/PI (-H ₂ O)		-0.110		-0.231 ^a		0.024		-0.008		-0.004
6	VP		0.070		0.079		0.052		0.011		0.025
7	LP/PL/IP/PI	0.685 ^b	0.224 ^a	0.597 ^b	0.145	0.459 ^b	0.543 ^b	0.630 ^b	0.507 ^b	0.526 ^b	0.605 ^b
8	PH/HP		-0.014		-0.024		0.005		0.049		0.037
9	FP (-H ₂ O)	0.210 ^a	-0.042	0.312 ^b	-0.164	-0.124	-0.044	0.267 ^a	0.076	0.257 ^a	0.125
10	FP	0.662 ^b	0.314 ^b	0.554 ^b	0.159	0.534 ^b	0.587 ^b	0.559 ^b	0.509 ^b	0.445 ^b	0.571 ^b
11	GPP	0.299 ^b	0.246 ^a	0.032	0.002	0.538 ^b	0.350 ^b	0.034	0.022	0.097	0.069
12	RP		0.417 ^b		0.365 ^b		0.539 ^b		0.570 ^b		0.506 ^b
13	Unknown		0.064		0.051		0.040		0.145		0.139
14	GPPG	0.136	0.015	-0.077	-0.028	0.342 ^b	-0.059	0.003	-0.034	0.124	0.052
15	PGPPG/GPPGP/PGGPP/PGPPG		0.282 ^b		0.061		0.457 ^b		0.212		0.271 ^a
16	GPPGPT										
17	PPGPPGP/GPPGPPG		0.436 ^b		0.166		0.626 ^b		0.379 ^b		0.425 ^b
18	AGPPGPPGP/GPPGPPGPA/PPGPPGPAG		0.431 ^b		0.242 ^a		0.287 ^b		0.536 ^b		0.537 ^b
19	APGDRGEFPPGP/PGDRGEPPGP	0.542 ^b	0.387 ^b	0.374 ^b	0.164	0.731 ^b	0.716 ^b	0.368 ^b	0.495 ^b	0.278 ^a	0.514 ^b
20	GPPGLAGPPGESGRE/PPGLAGPPGESGREG		0.328 ^b		0.108		0.746 ^b		0.387 ^b		0.396 ^b
21	PGFQGLPGAGPPGE		0.515 ^b		0.310 ^b		0.549 ^b		0.490 ^b		0.449 ^b
22	PAGRPGEVGPPGP	0.714 ^b	0.425 ^b	0.463 ^b	0.196	0.783 ^b	0.675 ^b	0.508 ^b	0.576 ^b	0.379 ^b	0.578 ^b

Values show Pearson correlation coefficients (two-tailed analysis) from the denosumab trial dataset. "Serum" and "urine" sub-columns correspond to the peptide fragment biofluid. N = 89 serum samples, n = 94 urine samples. ^aCorrelation sig. at *p* < .05 level. ^bCorrelation sig. at *p* < .01 level.

Quantification of global biofluid type I collagen peptide fragments has potential to provide a more detailed assessment of bone metabolism than established bone turnover markers. Established bone resorption markers are products of cathepsin K. We show that many bone breakdown products result from collagenolytic activity of other enzymes including MMPs. Altered activity of specific MMPs is a major factor in the pathogenesis of various common diseases associated with bone dysregulation, including osteoarthritis, bone cancer, and a range of conditions associated with osteolysis.²⁶ It is worth noting the peptide fragments observed in these patients that showed longitudinal changes inconsistent with CTX and P1NP despite being demonstrated as bone breakdown products in osteoclast cultures. Serum Leu-Hyp (Figure S4), for example, decreased following total hip replacement. The reason for this finding is unknown but suggests a multiplicity of bone resorption pathways revealed by monitoring the range of peptide fragments presented here. It is also possible that some peptide fragments observed in biofluids in vivo do not exclusively arise from bone collagen or from osteoclastic enzyme activity. Dietary protein and breakdown of connective tissues other than bone may contribute to the biofluid pools of some peptide fragments. MMPs of non-osteoclast origin, for example, MMP13 produced by periosteoclast cells, can also degrade bone collagen.²⁷

Previous analyses of type I collagen fragment peptides in urine^{4–8,28} or from in vitro digestion of human type I collagen protein by cathepsin K¹⁰ show considerable variation in the fragments observed. Many factors are likely to account for this variation, including the experimental conditions used in studies of bone or collagen digestion in vitro, in addition to instrumentation and analytical conditions and workflows employed in detection of peptides. The LC/MS technique used here is different to platforms used in other studies reporting collagen peptide fragments, which have frequently used capillary electrophoresis coupled to mass spectrometry.^{4–8,28} Our LC/MS approach has been employed extensively in small molecule analyses including metabolomic applications and is possibly more optimal for detection of low, versus high, MW peptides.¹³ This contrasts the proteomic approaches used in previous studies and shows the potential to detect low-medium MW bone collagen fragments in serum and urine metabolomic datasets. The data mining approach in terms of initial molecular feature extraction on raw LC/MS data was untargeted, in contrast to other studies reporting collagen-derived peptides. Identification of molecules of interest (ie, bone resorption products) was performed post hoc. In this way, the extraction of signals/peptide features did not depend on existing knowledge including the sequences of known peptides and data from existing peptide libraries. In other words, the profiles of ion features that were detected and extracted from the raw data, independent of their subsequent biochemical identifications, were included in the analysis. An advantage of this untargeted approach is that some peptide features remain incompletely characterized and are therefore not detectable using common proteomic data mining approaches which rely upon existing peptide libraries. For type I collagen, incompletely characterized features include the extent of cross-linking, known to occur at multiple telopeptide and helical regions, and post-translational modifications.¹⁰

In conclusion, our data demonstrate the powerful approach of untargeted high-resolution MS analysis for characterization of the range of type I collagen degradation products. The

overlap between collagen fragments observed in osteoclast cultures and in human serum and urine shows culturing of osteoclasts on dentin is relevant to identify potential bone resorption markers. The novel collagen fragments presented, some of which are COL1A1-specific, provide new insights into the pathways of collagen degradation and support the clinical potential of monitoring collagen peptide fragments in bone disorders.

Acknowledgments

The authors thank Agilent Technologies UK Ltd for kindly providing access to instrumentation for analysis of in vitro experiment samples in this study, with special thanks to Dr Gordon Ross of Agilent Technologies for his assistance in these analyses.

Author contributions

Brendan P. Norman (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing), Jane P. Dillon (Conceptualization, Investigation, Methodology, Validation, Writing—review & editing), Sahem M. Alkharabsheh (Data curation, Formal analysis, Investigation, Writing—review & editing), Francoise Congues (Investigation, Validation, Writing—review & editing), Peter J.M. Wilson (Investigation, Writing—review & editing), Andrew S. Davison (Investigation, Writing—review & editing), Elinor Chapman (Data curation, Investigation, Writing—review & editing), James Baker (Data curation, Investigation, Writing—review & editing), Seamus Coyle (Resources, Writing—review & editing), Chris Probert (Resources, Writing—review & editing), Paul E. Beaulé (Resources, Validation, Writing—review & editing), Lakshminarayan R. Ranganath (Funding acquisition, Resources, Supervision, Writing—review & editing), Jeremy M. Wilkinson (Conceptualization, Validation, Writing—review & editing), and James A. Gallagher (Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing—review & editing)

Supplementary material

Supplementary material is available at *JBMR Plus* online.

Funding

Funding for this work was provided by a PhD Studentship awarded to BPN with funds from the University of Liverpool, Royal Liverpool University Hospital and Agilent Technologies UK Ltd in addition to a Sireau Career Development Fellowship awarded to BPN by the Alkaptonuria Society (Registered Charity #1101052). A University of Liverpool Early Career Researchers & Returners Fund award and an ECMage Impact Fund award obtained by BPN also funded this research. The authors confirm independence from the funders; the content of the article has not been influenced by the funders. Graphical abstract figure was created with use of BioRender.com.

Conflicts of interest

All authors have no conflicts of interest to disclose.

References

1. Blair HC, Teitelbaum SL, Ghiselli R, Gluck S. Osteoclastic bone resorption by a polarized vacuolar proton pump. *Science*. 1989;245(4920):855–857. <https://doi.org/10.1126/science.2528207>
2. van der Rest M, Garrone R. Collagen family of proteins. *FASEB J*. 1991;5(13):2814–2823.

3. Maahs DM, Siwy J, Argilés A, et al. Urinary collagen fragments are significantly altered in diabetes: a link to pathophysiology. *PLoS One*. 2010;5(9):e13051. <https://doi.org/10.1371/journal.pone.0013051>
4. Magalhães P, Pontillo C, Pejchinovski M, et al. Comparison of urine and plasma peptidome indicates selectivity in renal peptide handling. *Proteomics Clin Appl*. 2018;12(5):e1700163. <https://doi.org/10.1002/prca.201700163>
5. Marx D, Anglicheau D, Caillard S, et al. Urinary collagen peptides: source of markers for bone metabolic processes in kidney transplant recipients. *Proteomics Clin Appl*. 2023;17(4):e2200118. <https://doi.org/10.1002/prca.202200118>
6. Delles C, Schiffer E, von Zur MC, et al. Urinary proteomic diagnosis of coronary artery disease: identification and clinical validation in 623 individuals. *J Hypertens*. 2010;28(11):2316-2322. <https://doi.org/10.1097/HJH.0b013e32833d81b7>
7. Good DM, Züribig P, Argilés A, et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol Cell Proteomics*. 2010;9(11):2424-2437. <https://doi.org/10.1074/mcp.M110.001917>
8. Bannaga AS, Metzger J, Kyrou I, et al. Discovery, validation and sequencing of urinary peptides for diagnosis of liver fibrosis—a multicentre study. *eBioMedicine*. 2020;62:103083. <https://doi.org/10.1016/j.ebiom.2020.103083>
9. Delaissé JM, Andersen TL, Engsig MT, Henriksen K, Troen T, Blavier L. Matrix metalloproteinases (MMP) and cathepsin K contribute differently to osteoclastic activities. *Microsc Res Tech*. 2003;61(6):504-513. <https://doi.org/10.1002/jemt.10374>
10. Demeuse J, Massonnet P, Schoumacher M, et al. Innovative workflow for the identification of cathepsin K cleavage sites in type I collagen. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2013;1228:123864. <https://doi.org/10.1016/j.jchromb.2023.123864>
11. Linde A, Lussi A, Crenshaw MA. Mineral induction by immobilized polyanionic proteins. *Calcif Tissue Int*. 1989;44(4):286-295. <https://doi.org/10.1007/BF02553763>
12. Walsh CA, Beresford JN, Birch MA, Boothroyd B, Gallagher JA. Application of reflected light microscopy to identify and quantitate resorption by isolated osteoclasts. *J Bone Miner Res*. 1991;6(7):661-671. <https://doi.org/10.1002/jbmr.5650060703>
13. Norman BP, Davison AS, Ross GA, et al. A comprehensive LC-QTOF-MS metabolic phenotyping strategy: application to alkaptonuria. *Clin Chem*. 2019;65(4):530-539. <https://doi.org/10.1373/clinchem.2018.295345>
14. Vorkas PA, Isaac G, Anwar MA, et al. Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: application to cardiovascular disease. *Anal Chem*. 2015;87(8):4184-4193. <https://doi.org/10.1021/ac503775m>
15. Slullitel PA, Mahatma MM, Farzi M, Grammatopoulos G, Wilkinson JM, Beaulé PE. Influence of femoral component design on proximal femoral bone mass after Total hip replacement: a randomized controlled trial. *J Bone Joint Surg Am*. 2021;103(1):74-83. <https://doi.org/10.2106/JBJS.20.00351>
16. Mahatma MM, Jayasuriya RL, Hughes D, et al. Effect of denosumab on osteolytic lesion activity after total hip arthroplasty: a single-Centre, randomised, double-blind, placebo-controlled, proof of concept trial. *Lancet Rheumatol*. 2021;3(3):e195-e203. [https://doi.org/10.1016/S2665-9913\(20\)30394-5](https://doi.org/10.1016/S2665-9913(20)30394-5)
17. Coyle S, Chapman E, Hughes DM, et al. Urinary metabolite model to predict the dying process in lung cancer patients. *Commun Med*. 2025;5(1):49. <https://doi.org/10.1038/s43856-025-00764-3>
18. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388-W396. <https://doi.org/10.1093/NAR/GKAB382>
19. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res*. 2018;46(D1):D624-D632. <https://doi.org/10.1093/nar/gkx1134>
20. Fizazi K, Massard C, Smith M, et al. Bone-related parameters are the main prognostic factors for overall survival in men with bone metastases from castration-resistant prostate cancer. *Eur Urol*. 2015;68(1):42-50. <https://doi.org/10.1016/j.euro.2014.10.001>
21. Rajpar S, Massard C, Laplanche A, et al. Urinary N-telopeptide (uNTx) is an independent prognostic factor for overall survival in patients with bone metastases from castration-resistant prostate cancer. *Ann Oncol*. 2010;21(9):1864-1869. <https://doi.org/10.1093/annonc/mdq037>
22. Demers LM, Costa L, Lipton A. Biochemical markers and skeletal metastases. *Cancer*. 2000;88(S12):2919-2926. [https://doi.org/10.1002/1097-0142\(20000615\)88:12+<2919::AID-CNCR7>3.0.CO;2-Z](https://doi.org/10.1002/1097-0142(20000615)88:12+<2919::AID-CNCR7>3.0.CO;2-Z)
23. Boopathi E, Birbe R, Shoyele SA, Den RB, Thangavel C. Bone health Management in the Continuum of prostate cancer disease. *Cancers (Basel)*. 2022;14(17):4305. <https://doi.org/10.3390/cancers14174305>
24. Lipton A, Chapman JW, Demers L, et al. Use of elevated bone turnover to predict bone metastasis. *J Clin Oncol*. 2008;26(15_suppl):591-591. https://doi.org/10.1200/jco.2008.26.15_suppl.591
25. Zhao H, Han KL, Wang ZY, et al. Value of C-telopeptide-cross-linked type I collagen, osteocalcin, bone-specific alkaline phosphatase and procollagen type I N-terminal propeptide in the diagnosis and prognosis of bone metastasis in patients with malignant tumors. *Med Sci Monit*. 2011;17(11):CR626-CR633. <https://doi.org/10.12659/MSM.882047>
26. Simsa-Maziel S, Sela-Donenfeld D, Monsonego-Ornan E. Matrix metalloproteinases in bone health and disease. In: Chakraborti S, Dhalla NS, eds. *Proteases in Health and Disease*. Springer; 2013:289-312. https://doi.org/10.1007/978-1-4614-9233-7_17
27. Garnerio P, Ferreras M, Karsdal MA, et al. The type I collagen fragments ICTP and CTX reveal distinct enzymatic pathways of bone collagen degradation. *J Bone Miner Res*. 2003;18(5):859-867. <https://doi.org/10.1359/jbmr.2003.18.5.859>
28. Siwy J, Zoja C, Klein J, et al. Evaluation of the Zucker diabetic fatty (ZDF) rat as a model for human disease based on urinary peptidomic profiles. *PLoS One*. 2012;7(12):e51334. <https://doi.org/10.1371/journal.pone.0051334>