Assessment of different screening methods for selecting palaeontological bone samples for peptide sequencing

Samantha Presslee1,2\*, Kirsty Penkman2., Roman Fischer3, Eden Richards-Slidel1,4, John Southon5, Carolina Acosta Hospitaleche6, Matthew Collins4 ,7, Ross MacPhee8

Affiliations

1. BioArCh, Department of Archaeology, University of York, York, UK
2. BioArCh, Department of Chemistry, University of York, York, UK
3. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK
4. Section for Evolutionary Genomics, The Globe Institute, Faculty of Health, University of Copenhagen, Copenhagen, Denmark
5. Department of Earth System Science, University of California, Irvine, USA
6. CONICET. División Paleontología Vertebrados. Museo de La Plata, La Plata, Argentina
7. McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK
8. Department of Mammalogy, American Museum of Natural History, New York, NY, USA

\* Corresponding Author: s[am.presslee@york.ac.uk](mailto:Sam.presslee@york.ac.uk), BioArCh, University of York, York, YO10 5DD.

Keywords: Palaeoproteomics, MALDI-ToF MS, Chiral AA analysis, ATR-FTIR, Collagen

# Abstract

Ancient proteomics is being applied to samples dating further and further back in time, with many palaeontological specimens providing protein sequence data for phylogenetic analysis as well as protein degradation studies. However, fossils are a precious material and proteomic analysis is destructive and costly. In this paper we consider three different techniques (ATR-FTIR, MALDI-ToF MS and chiral AA analysis) to screen fossil material for potential protein preservation, aiming to maximise the proteomic information recovered and saving costly time consuming analyses which may produce low quality results. It was found that splitting factor and C/P indices from ATR-FTIR were not a reliable indicator of protein survival as they are confounded by secondary mineralisation of the fossil material. Both MALDI-ToF MS and chiral AA analysis results were able to successfully identify samples with surviving proteins, and it is therefore suggested that one or both of these analyses be used for screening palaeontological specimens.

## Introduction

The analysis of proteins from fossils (‘palaeoproteomics’) has been valuable in the phylogenetic analysis of extinct taxa, as well as for understanding protein degradation [1,2,3]. The leading analytical method for palaeoproteomics is soft-ionisation mass spectrometry, particularly liquid chromatography tandem mass spectrometry (LC-MS/MS), which can characterise peptides and ultimately whole proteomes [[4,5,6,7,8,9]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). However, this is a costly and destructive technique to use on valuable fossil material. Therefore, several techniques have been developed to act as screening methods to select promising fossil samples for proteomic analysis [10, 11, 12, 13, 14]. However, the reliability of these techniques to accurately predict protein survival has yet to be tested.

In this study, bone samples were tested using 3 commonly used screening methods; attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR), chiral amino acid (AA) analysis using reverse phase-high performance liquid chromatography (RP-HPLC), and matrix assisted laser desorption ionisation- time of flight mass spectrometry (MALDI ToF-MS). A selection of these samples were also analysed using LC-MS/MS and the level of proteomic recovery is plotted alongside the individual screening method datasets, helping to determine overall trends and the reliability of these methods.

The three screening methods chosen for this study were selected because they are relatively inexpensive, rapid in throughput, and provide complementary information on inorganic and organic preservation. Attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) has been shown to be a fast, minimally destructive technique for assessing archaeological and fossil bones [15]. The presence of organics (such as proteins) can be inferred through the presence of amide peaks [13, 16], and the technique also allows evaluation of the inorganic content via assessment of the infrared splitting factor (IRSF) and carbonate to phosphate ratio (C/P) [17]. Measuring the IRSF of the mineral component of bone provides information about how ordered the crystals are in the matrix. Large, regularly organised crystals and those with fewer impurities (such as carbonate) will have a higher IRSF, while modern bone tends to have a low IRSF [17]. An increase in IRSF in bone shows the sample has undergone alteration and suggests loss of the organic content which would allow space for these new crystals to form. The carbonate to phosphate ratio (C/P) is another criterion that is commonly used to show alteration in the bone mineral and therefore potential organic loss [18].

MALDI-ToF MS has become commonly used in archaeological studies to identify peptides that provide taxonomic identification (often referred to as ZOOarchaeology through Mass Spectrometry or ZooMS [19, 20, 21, 22]. This technique is fast, inexpensive, and minimally destructive, requiring only 15-30 mg of bone. It is also advantageous because the remaining sample extraction can be sent directly for LC-MS/MS analysis without further sample preparation. This method has already been used as a screening method for fossil samples [[1, 23, 24]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk) and can provide a direct indication of surviving proteins.

Chiral AA analysis has been used for decades to analyse a wide variety of fossil samples, mainly as a means of relative dating using amino acid racemisation [25, 26, 27, 28, 29, 30, 31], but also to prove authenticity of surviving peptides [[2, 32, 33]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) and investigate protein degradation [34, 35]. It can also be used to investigate possible modern contamination. The D-form of alanine and glutamic acid can be found in certain bacterial cell walls [[36]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). Therefore, an increase in the D/L values of these amino acids, over other faster racemising amino acids, may imply bacterial infiltration of the bone. If an ancient sample shows low levels of racemisation, this can be indicative of modern contamination [[37]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) as well as an amino acid profile that is not consistent with bone [33].

The overall aim of this paper is to verify the use of ATR-FTIR, AA analysis, and MALDI-ToF as a means to accurately predict the level of surviving proteomic information from fossil samples prior to further destructive sampling. Once a reliable non/minimally destructive screening method(s) is established, it is hoped that more palaeontological samples will be able to be proteomically examined, enabling a better understanding of protein survival into deep time.

## Materials and methods

### 2.1 Sample information

133 bone samples from various locations (Arctic, North America, Caribbean, South America), covering a wide time span (Holocene-Cretaceous), and 34 collagen extracts left over from radiocarbon dating bone from Rancho La Brea tar pits, were analysed using a combination of ATR-FTIR, AA analysis and MALDI-ToF MS. Cortical bone was selected for these analyses when possible. 18 bone samples and 4 radiocarbon extract samples were also analysed using LC-MS/MS and the level of proteomic recovery was compared to the screening results. The radiocarbon samples from Rancho La Brea were selected as it has been shown that proteomic information may be retrieved from collagen samples purified for radiocarbon analysis, which would otherwise be discarded [[37, 38, 39]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). Sample information and the results of each of the analyses can be found in the SI. Fewer samples were tested for FTIR than for the other two methods due to external circumstances and this is taken into account when evaluating the results.

### 2.2 Methods

#### **2.2.1 ATR-FTIR**

All samples were crushed to a fine powder using a cleaned pestle and mortar before being analysed. The whole bone was chosen to be analysed as this would be the least destructive to the fossil material. Approximately 2-5 mg of the bone powder was analysed using an Alpha Platinum FTIR-ATR spectrometer with a diamond crystal, following the methods of [Kontopoulos et al.](https://paperpile.com/c/inFlrI/Bszal) [[41]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). In order to ensure good contact between the diamond crystal (used to pass the IR beam into the sample) and the sample itself, the pressure of the applicator was controlled using a pressure control spot. Both the applicator and the plate were cleaned with ethanol between samples. The resulting spectra were analysed using OPUS 7.5 software and the IRSF, and C/P were calculated using the formulas outlined in [Kontopoulos et al.](https://paperpile.com/c/inFlrI/Bszal) [[41]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY).

#### **2.2.2 Chiral AA analysis**

Samples were prepared for amino acid analysis using a modified version of the protocol outlined in [[42]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). Approximately 1 mg of sample was hydrolysed in 7 M hydrochloric acid (100 µL per mg) in a sterile glass vial under N2 for 18 hours at 110OC. After hydrolysis, the samples were placed into a centrifugal evaporator and spun to dryness before being re-hydrated in 0.01 mM L-homo-arginine (an internal standard). The samples were analysed by RP-HPLC using an Agilent 1100 HPLC following a slightly modified protocol set out by [Kaufman and Manley](https://paperpile.com/c/inFlrI/WrVam) [[43]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). It must be noted that during hydrolysis both asparagine and glutamine can undergo deamidation to aspartic acid and glutamic acid respectively [[](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)44]. Therefore, it is not possible to differentiate between these amino acids, and as such, are referred to as Asx and Glx respectively, in amino acid studies. The chromatograms were analysed using Chemstation, and an in-house VB application.

#### **2.2.3 MALDI-ToF MS**

##### **Bone sample preparation**

The samples were prepared using a slightly modified protocol set out by Welker et al. [21]. Between 15-30 mg of bone was demineralised in 0.6 M hydrochloric acid (HCl), washed in 0.01 M sodium hydroxide to remove any potential humic contaminants and three times in 50 mM ammonium bicarbonate (Ambic). The samples were then heated at 65OC for 1 hour to allow any available proteins to solubilise into solution. After heating, the samples were split; 50 µL was pipetted into a clean plastic microcentrifuge vial and stored in the freezer for additional analysis if needed. The other 50 µL was digested overnight using 1 µL of 0.5 µg/µL porcine trypsin in trypsin resuspension buffer (Promega, UK) at 37OC and the digestion was stopped by the addition of trifluoroacetic acid (TFA) at a concentration of 0.5-1% of the total solution. The samples were desalted using C18 zip-tips [[44]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) and eluted using 100 µL of 50% acetonitrile (ACN)/0.1% TFA (v/v). The zip-tipped samples were spotted in triplicate onto a MTP384 Bruker ground steel MALDI target plate. 1 µL of sample was pipetted onto each sample spot before being mixed with 1 µL of α-cyano-4-hydroxycinnamic acid matrix solution (1% in 50% acetonitrile / 0.1% trifluoroacetic acid (v/v/v)).

##### **Radiocarbon extract sample preparation**

The collagen from these samples had been previously extracted for radiocarbon dating; see Fuller et al. [46] for the detailed protocol and the radiocarbon results. In brief, samples were crushed, sonicated in a mix of toluene and methanol to remove the tar, demineralised, gelatinised, and ultrafiltered twice with the 3-30 kDa fraction being retained for radiocarbon analysis. A subsample (approximately 1 mg each) was taken from each sample of purified freeze-dried collagen. The protocol follows the same steps as the bone preparation from digestion onwards. The final zip-tipped samples were diluted 1:8 to prevent oversaturation, before being spotted on the MALDI plate.

##### **MALDI-ToF MS**

The samples were analysed on a Bruker Ultraflex III MALDI-ToF mass spectrometer. The resulting MS spectra were analysed using mMass [[47]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY), an Open Source mass spectrometry interpretation tool. The three spectra for each sample were averaged and the averaged spectrum was cropped between 800-3000 *m/z* and peak picking was carried out using a signal to noise ratio of 6 [[48]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). As part of the peak-picking, the baseline was corrected, peaks smoothed using the Savitzky-Golay method, and deisotoping was performed all using the predefined parameters set by mMass [47]. Shoulder peaks were also removed.

#### **2.2.4 LC-MS/MS analysis**

The tryptic digest and the left over 50 µL aliquot of the samples that were selected for LC-MS/MS were dried down and the 50 µL aliquot was re-suspended in 50 µL 100 mM tris(hydroxymethyl)aminomethane solution and digested with elastase (Worthington; USA) at the same concentration in 10% tris solution. The elastase digests were zip-tipped as the tryptic digests described above and then dried down for LC-MS/MS. Two different enzymes were used for digestion in order to increase the sequence coverage [[2, 21]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY).

The samples were analysed at the Discovery Proteomic Facility (DPF) at Oxford, UK, with the exception of *Mylodon* (Sample ID 16222) and *Scelidodon* (Sample ID 17480) which were analysed at the Novo Nordisk Foundation Centre for Protein Research (NNFCPR), University of Copenhagen. In general, the MS methods followed those outlined in Presslee et al. [3]. At Oxford, peptides were separated on a Dionex Ultimate 3000 UPLC (Thermo Fisher) using a linear gradient of 2-35% acetonitrile in 0.1% formic acid / 5%DMSO over 60 minutes and a 50 cm Easyspray column (ES803, Thermo Fisher). Peptide detection was achieved with either a Q-Exactive or a Fusion Lumos mass spectrometer, optimized for maximum sensitivity and spectrum quality (both Thermo Fisher). MS resolution was set to 70,000 on the Q-Exactive (120,000 on Fusion Lumos) with an AGC target of 3e6 (4e5). MS/MS resolution was set to 17,500 (30,000) with an AGC target of 1e5 (3E5). The maximum injection time was 128 ms (120ms) using the 15 most abundant precursor per duty cycle (Top Speed, 2 seconds duty cycle). Collision energy was set to 28% (28% ± 5%) and precursors were excluded for 27 seconds (7 seconds). The Mylodon and Scelidodon samples were analysed in Copenhagen under the following conditions: MS1: 120 k resolution, maximum injection time (IT) 25/45 ms, scan target 3E6. MS2: 60/30 k resolution, top 10 mode, maximum IT 108 ms, minimum scan target 2E5, normalized collision energy of 28, dynamic exclusion 20 s, and isolation window of 1.2 m/z.

The RAW files were converted to Mascot generic format files (MGFs) using proteowizard [[49]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) and the MGF files were analysed using PEAKS v.7.5 (Bioinformatics Solutions [[50]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)). The samples were searched against a list of common contaminants (common Repository of Adventitious Proteins: http://www.thegpm.org/ crap), a mammal collagens database, and the proteome of the relevant taxa if available on NCBI (i.e. Equid proteome for the horse samples). The following tolerances were used for the searches; mass tolerance of 0.5 Da for the fragment ions and 10 ppm for the precursor ions, a maximum of 3 missed cleavages and 3 PTMs were allowed per peptide. The protein tolerances were set at 0.5% false discovery rate (FDR), >50% average local confidence (ALC; *de novo* only) and -10lgP score ≥ 20. Various PTMs were allowed for each search including oxidation (MHW) and hydroxylation of proline (both +15.99), deamidation (NQ; +0.98) and pyro-glu from E (-18.01) as well as a fixed PTM of carbamidomethylation (+57.02) which occurs as part of the sample preparation [[2, 51]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY).

#### 2.2.5 Evaluation of the screening methods

In order to establish criteria for each of the screening methods to predict levels of proteomic survival, the results of each of the datasets were plotted alongside the samples that had undergone LC-MS/MS (Table 1). Any patterns and relationships between good sequence recovery and the screening results were investigated.

## 3. Results

All analytical results are recorded in the SI.

### 3.1 LC-MS/MS results

The results of the LC-MS/MS analysis are summarised in Table 1. The number of collagen 1 peptides (both A1 and A2 chains) and % coverage of collagen 1 were recorded for each of the sequenced samples. The evaluation of proteins other than collagen (NCPs) was not possible. Non collagenous proteins are less abundant in bone and are also less stable than collagen. Therefore, even when comparison proteomes were available, very few NCPs were identified.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sample ID** | **Species** | **Age** | **Location** | **# Col1 Peptides** | **% Col1 Coverage** | **# NCPs** |
| 14715 | *Paramylodon* | Unknown | Roseburgh, Oregon | 642 | 87 | N/A |
| 14723 | *Nothrotheriops shastensis* | Unknown | Rampart Cave, Arizona | 528 | 79 | N/A |
| 15191 | *Doedicurus* | Late Pleistocene  (Lujanian 8.5-128Ka) | Camet Norte, Argentina | 867 | 90 | N/A |
| 15194 | *Glyptodon* | Late Pleistocene  (Lujanian 8.5-128Ka) | Buenos Aires, Argentina | 731 | 84 | N/A |
| 15202 | *Scelidotherium* | Early Pleistocene (Ensenadan 0.4-1.8Ma) | Unknown | 475 | 76 | N/A |
| 15216 | *Glossotherium* | Middle Pleistocene (Bonarean 128-400Ka) | Unknown | 837 | 88 | N/A |
| 15225 | *Megatherium americanum* | Late Pleistocene (18, 000 BP) | Bariloche, Argentina | 520 | 81 | N/A |
| 15248 | *Palaeosphenicus biloculata* | Early Miocene (Burdingalian15.97-20.44Ma) | Patagonia, Argentina | 3 | 2 | 0 |
| 15255 | *Palaeeudyptes gunnari* | Eocene (Bartonian 37.8-41.1Ma) | Marambio/Seymour I, Antarctica | 1 | 1 | 0 |
| 15548 | *Neocnus comes* | Unknown | Trouing de la Scierie, Haiti | 699 | 84 | N/A |
| 15556 | *Parocnus serus* | Unknown | Trouing Marassa, Haiti | 575 | 82 | N/A |
| 15559 | *Acratocnus ye* | Unknown | Trouing Marassa, Haiti | 696 | 86 | N/A |
| 15564 | *Megalocnus zile* | Unknown | Trou Gallery; Haiti | 6 | 6 | N/A |
| 15565 | *Acratocnus* | Unknown | Trouing Marassa, Haiti | 629 | 87 | N/A |
| 15780 | *Neocnus comes* | Holocene (4486 +/- 39 BP) | Trouing Deron, Haiti | 591 | 84 | N/A |
| 15781 | *Neocnus Dousman* | Holocene (9867 +/- 65 BP) | Trouing Marrasa, Haiti | 614 | 74 | N/A |
| 16222 | *Mylodon* | Unknown | Cueva del Medio, Chile | 1371 | 96 | N/A |
| 17480 | *Scelidodon* | Pleistocene | Cueva Rosello, Peru | 1324 | 92 | N/A |
| **15063** | ***Panthera atrox*** | **Late Pleistocene (34860+/- 710 BP)** | **Rancho La Brea, Project 23 Collection** | **788** | **89** | **1** |
| **15066** | ***Canis dirus*** | **Late Pleistocene (33380+/- 800 BP)** | **Rancho La Brea, Project 23 Collection** | **895** | **89** | **0** |
| **15078** | ***Equus occidentalis*** | **Late Pleistocene (22890+/- 120 BP)** | **Rancho La Brea, Hancock Collection** | **793** | **92** | **0** |
| **15092** | ***Smilodon fatalis*** | **Late Pleistocene (36050+/- 600 BP)** | **Rancho La Brea, Project 23 Collection** | **861** | **90** | **1** |

**Table 1.** The LC-MS/MS results. Col1 includes both alpha1 and alpha2 chains. The samples in bold are the radiocarbon extract samples [45]. NCPs: non-collagenous proteins, a protein had to have 2 or peptides in order to be considered present in the sample. Comparison proteomes were not available for the majority of the samples (N/A) limiting the evaluation of protein survival other than collagen. The radiocarbon dates for samples 15780 and 15781 are from Steadman et al [52].

The majority of the sequenced samples produced good protein (collagen) recovery and are classified as good when being used to evaluate the screening method results. The exceptions are *Palaeosphenicus biloculata* (Sample ID 15248)*, Palaeeudyptes gunnari* (Sample ID 15255) and *Megalocnus zile* (Sample ID 15564), which showed very low collagen coverage. These three samples are recorded as poor when used to evaluate the screening method data. Sample 15248 dates to the Early Miocene and sample 15255 is even older, dating to the Eocene. Due to the age of these samples it is not surprising that very little protein content is recoverable, and it is important to note that the peptides recovered may be exogenous contamination This is explored further in the discussion below. Sample 15564 is from Haiti and likely dates to the Late Pleistocene or even Holocene [52] and the other Haitian samples all produced good sequence data. The Rancho La Brea (RLB) radiocarbon extracts all date to the Late Pleistocene and resulted in the identification of very few non-collagenous proteins (NCPs) despite good collagen recovery. This is not surprising as the extraction protocol used two filtration steps in order to remove the tar and purify the collagen. Therefore, any surviving NCPs were likely also removed [[46]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). More sample information and all the results can be found in the SI.

### 3.2 ATR-FTIR results

ATR-FTIR is the most rapid out of all the techniques considered here- taking a matter of minutes to produce and analyse the spectrum. This method also has the advantage of not requiring any prior sample preparation or protein extraction negating any possible extraction method bias. The mineral content of the samples were analysed using the splitting factor (IRSF) and carbonate phosphate ratio (C/P) which has been used in previous archaeological bone studies to predict degradation [[53]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). All the samples are plotted against IRSF and C/P and the sequenced samples are highlighted (Figure 1a). In general, a higher IRSF is indicative of an increase in homogeneity and crystal size and a decrease in C/P shows the mineral has undergone alteration [54].



**Figure 1. The ATR-FTIR results**. IRSF: infrared splitting factor, C/P: carbonate phosphate ratio. The modern reference sample is taken from [Kontopoulos et al.](https://paperpile.com/c/inFlrI/pQ3k) [[53]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). The results are shown in terms of the LC-MS/MS results (a), the sample location (b) and the age of the sample (c).

When the mineral content of the samples is compared to the level of proteomic recovery, there is no trend between mineral preservation and surviving collagen, as assessed by LC-MS/MS (Figure 1a), with both good and poorly preserved samples plotting close to the modern reference. Therefore, it was not possible to define a criterion in which to select samples for successful LC-MS/MS analysis. To investigate this lack of correlation between mineral alteration and protein degradation further, the results were also plotted according to location (Figure 1b) and the age of the sample (Figure 1c). No clear trend can be discerned in relation to the FTIR values and the location of the samples (Figure 1b) and this may be due to the smaller dataset for this method. However, the Antarctic samples all show a tightly grouped IRSF which suggests that cold environments may prevent or slow mineral alteration. The younger samples have signs of diagenetic alterations as expected (i.e. a higher splitting factor and lower C/P), but interestingly, as the age of the fossil increases, the FTIR values become more uniform and fall back in line with modern bone. This is especially evident when looking at the IRSF. These results suggest that FTIR can detect early diagenetic changes to bone mineral, but the complex process of fossilization and secondary mineralisation can affect the bone crystallinity in a way that disrupts the formation of more homogenous crystals and in turn lowers the IRSF.

Other studies have also shown modern bone values of IRSF and C/P in fossil samples, which they attributed to possible trace element uptake [[55]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). Therefore, it is likely that IRSF and C/P cannot distinguish between endogenous and exogenous mineralisation in all cases [[56]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY), and the addition of other foreign ions such as trace elements [[57]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) which may explain why we are seeing results in line with modern bone. The organic content of the bone was assessed using amide/phosphate (Am/P), which has been successfully used in other studies to predict collagen survival [58, 59, 60]. The predicted % weight of collagen was calculated using the Am/P values and the formula presented in [58]. The results of the sequenced samples are shown in Table 2 (all Am/P values are in the SI).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Am/P | % weight collagen | Sequence coverage | Number of peptides |
| *Megatherium americanum* | 0.215 | 26.0 | 81 | 520 |
| *Paramylodon* | 0.184 | 22.5 | 87 | 642 |
| *Nothrotheriops shastensis* | 0.180 | 22.0 | 79 | 528 |
| *Glyptodon* | 0.093 | 12.2 | 84 | 731 |
| *Parocnus serus* | 0.086 | 11.4 | 82 | 575 |
| *Neocnus comes* | 0.068 | 9.4 | 84 | 591 |
| *Doedicurus* | 0.047 | 7.0 | 90 | 867 |
| *Mylodon* | 0.038 | 6.0 | 96 | 1371 |
| *Neocnus comes* | 0.029 | 5.0 | 84 | 699 |
| *Acratocnus* | 0.025 | 4.5 | 87 | 629 |
| *Acratocnus ye* | 0.023 | 4.3 | 86 | 696 |
| *Glossotherium* | 0.011 | 2.9 | 88 | 837 |
| ***Megalocnus zile*** | **0.080** | **10.7** | **6** | **6** |
| ***Palaeosphenicus biloculata*** | **0.041** | **6.3** | **2** | **3** |
| ***Palaeeudyptes gunnari*** | **0.021** | **4.1** | **1** | **1** |

**Table 2.** The Am/P results of the sequenced samples. The samples that produced poor results are shown in bold. The predicted % weight of collagen was calculated following the formula in [56].

The majority of the good samples show a high predicted % weight of collagen (>6%), with a result higher than or near the 0.04 Am/P cut off suggested by Pothier Bouchard et al. [59]. In contrast, two of the poor LC-MS/MS samples (shown in bold in Table 2) were above this threshold but failed to produce meaningful LC-MS/MS results. However, more critically four of the samples that produced good quality LC-MS/MS data fell below this threshold, but still produced high sequence coverage.

These results suggest that while ATR-FTIR can be used to distinguish between modern and relatively young archaeological bone [17] using IRSF and C/P, in older palaeontological samples the results can be complicated due to secondary mineralisation, which prevents these parameters being able to predict the level of protein preservation. Am/P is a more reliable indicator of protein preservation, and samples with a clear amide band are usually shown to have good protein recovery. However, samples that showed a more degraded amide peak can still produce good sequence data, so using only this parameter for screening will risk samples containing valuable proteomic information being overlooked

### 3.3 Chiral AA analysis results

AA analysis requires a longer sample preparation time than ATR-FTIR; overall it takes three days for sample preparation and each sample takes 2 hours to run on the RP-HPLC. However, the data analysis is uncomplicated and a wealth of data is generated: concentration data, racemisation data, and amino acid percentage composition profiles. All of these datasets can provide useful indications of protein preservation; total concentration will show how many amino acids are present in the sample, racemisation data will give an indication of age, as well as possible bacterial and modern contamination [[36]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY), and composition data can be used to see if the bone sample contains the expected collagen-like amino acid profile [[33]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). The concentration data on its own is perhaps the least informative, but samples with low AA concentrations are unlikely to produce meaningful proteomic data, and it is possible that any amino acids present are not endogenous [33, 61]. The percentage composition and D/L values were analysed using PCA (R Studio) (Figure 2). The R script used to perform this analysis is available on request.



**Figure 2. The PCA of the AA composition and D/L values**. Grey circles indicate unsequenced samples, green shows the successful MS/MS results and red shows poor MS/MS results. A modern reference is shown in black. Only two poor MS/MS values are shown for the D/L plot as the *Palaeeudyptes gunnari* sample produced a very low AA total concentration and no D amino acids were detected.

Most of the LC-MS/MS “good” samples group closely together for both % composition and D/L values. The good samples show in general a low racemisation value, suggesting limited protein breakdown as, for the majority of amino acids, racemisation can only occur once the peptide chain breaks down [2]. The poor samples all show higher D/L values, suggesting more substantial protein degradation and loss. The same grouping of the good samples can be seen in the composition PCA, with a high glycine content as a key parameter. This is not surprising, as glycine is every third amino acid in the collagen molecule and this close grouping would represent the samples sharing a collagen like amino acid profile with minimal variability between the samples. The samples that place outside of this grouping suggest a non-collagen like amino acid composition, which may be suggestive of endogenous protein loss and the possible uptake of other contaminant amino acids. The one poor sample that groups near to the good samples in the composition PCA was *Palaeeudyptes gunnari.* This sample has a very low total AA concentration of ~3500 pmol/mg (See SI for details), and although the overall profile did not match that of collagen, it was dominated by glycine. This may be another indication of microbial contamination [33]. Therefore, it is suggested that total concentration should be the initial check alongside comparing D/L and % composition values.

### 3.4 MALDI-ToF MS results

This screening method is the most time intensive and destructive, with the sample preparation taking a week on average, and the use of data analysis software to analyse the results. However, the major benefit to this method is that once the sample has been prepared, it can be dried down and sent for LC-MS/MS without needing any further extraction. In order to evaluate the MALDI spectra, the total number of peaks and the highest peak intensity were compared (Figure 3). The results of the LC-MS/MS sequenced samples are highlighted in green (good proteomic recovery) or red (poor proteomic preservation). It must be noted that the peak intensity is not a measure of concentration, but rather a mixture of both peptide abundance and how well the peptide can be detected using MS.

****

**Figure 3.** **The MALDI-ToF MS results**. The number of MALDI peaks and the highest observed intensity for each of the samples were compared. The sequenced samples are highlighted; red indicates poor sequence data and green for good sequence data. Note the log scale on both the x and the y axis.

Overall, the samples that produced poor sequence data also produced poor MALDI-ToF spectra and vice versa. However, one sample (*Megalocnus zile* 15564)is an exception. The MALDI spectrum of this sample is comparable to the MALDI spectrum of *Megatherium americanum* (15225), which produced good LC-MS/MS results. The peak lists of both spectra were compared to possible ZooMS collagen peptides [19]. This resulted in 6 potential collagen peaks for both samples (see figure 4). When the spectra were compared to possible contaminants, sample 15564 had matches to both trypsin (due to sample digestion) and keratin, while only tryptic peaks were identified in sample 15225. However, other sequenced samples also contained keratin peaks (albeit with smaller intensities) in their MALDI spectra, so identification of keratin peaks does not always mean low levels of endogenous protein retrieval.



**Figure 4.** **A comparison between two MALDI spectra**. This shows the spectrum of sample 15225 (good sequence coverage) and sample 15564 (poor sequence coverage). The circles show potential collagen peaks, with trypsin peaks marked by T and the keratin peaks marked by K.

## 4. Discussion

ATR-FTIR is a minimally destructive technique that can provide key information about the early stages of bone mineral diagenesis. However, secondary mineralisation of the bone appears to influence the results, with some palaeontological samples placing near to the modern reference, thereby confounding its practical use for screening all samples. The presence of noise in the IR spectra has also been shown in other studies analysing palaeontological specimens [33], and ATR-FTIR has also been shown to be influenced by other factors; where and how the bone has been sampled[[62]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY), the extent to which the sample was ground [[41]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY), and other macroscopic effects such as electrostaticity[[63]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY). Overall, these variables and the inability to differentiate between endogenous apatite and secondary mineralisation, makes this technique an unreliable proxy for determining proteomic retrieval in fossil remains.

The chiral AA analyses produces a wealth of data, including concentration data, extent of racemisation, and an amino acid profile of the sample. Samples with low overall concentration levels of amino acids did not produce meaningful proteomic results. Therefore, it is suggested that concentration data is used as a first check, and this study has shown that samples over 100,000 pmol/mg are likely to be useful for further analysis for proteomic information. Over this threshold, the composition and D/L information provided distinct groupings (Figure 2) where proteomic success was likely, and conversely under the threshold proteomic success was unlikely. The PCA shows samples with low overall racemisation values and a collagen like AA profile dominated by glycine should be selected for further analysis. Once this collagen like amino acid profile is lost, it seems protein degradation can be quite rapid, with increased racemisation. This supports the hypothesis that when the triple helix structure of collagen is lost, the protein becomes more susceptible to breakdown [35]. The strength of the chiral amino acid analysis approach is in the complementary information sources; concentration data should be used as an initial indicator of overall protein survival, while racemisation and amino acid composition data show a strong relationship with good sequence recovery, and therefore is likely to be useful for screening.

Overall, the MALDI-ToF MS results showed a strong correlation with the “good” LC-MS/MS results producing spectra with a high intensity and circa 100 *m/z* peaks. There was one false positive, sample 15564 (*Megalocnus zile*), showed a good MALDI spectrum but produced poor LC-MS/MS results. While these are both mass spectrometric techniques, they use different methods of ionisation and detection, which affects the level of proteomic information retrieved. When the MALDI-ToF and LC-MS/MS data are compared for this sample, only 1 *m/z* peak was matched to a collagen peptide (Table 2). The majority of the potential collagen peaks were either not detected in the MS/MS analysis, or were a match to contaminant proteins. The presence of one peptide is not enough to conclude any endogenous protein survival and it is likely that this sample has been contaminated. Therefore, in general MALDI-ToF MS results are in agreement with LC-MS/MS data, but caution is needed when analysing samples that are not currently in the ZooMS database (such as xenarthrans) and contain a high number of contaminant peaks.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **MALDI-ToF marker** | |  | **LC-MS/MS results** | | |
| ***m/z*** | **Protein** | **Match to LC-MS/MS?** | **Mass** | **Protein** | **Peptide** |
| 1453.7 | Collagen? | Y | 1452.7 | Collagen | GLPGEFGLPGPAGPR |
| 1566.8 | Collagen? | Y | 1565.8 | **Trypsin** | LGEHNIDVLEGNEQ |
| 1592.8 | Collagen? | Y | 1591.8 | **Keratin** | SMDNNRSLDLDSII |
| 2131.1 | Collagen? | N |  |  |  |
| 2853.4 | Collagen? | N |  |  |  |
| 2869.4 | Collagen? | N |  |  |  |
| **842.5** | **Trypsin** | **Y** | **841.5** | **Trypsin** | **VATVSLPR** |
| **1045.5** | **Trypsin** | **Y** | **1044.5** | **Trypsin** | **LSSPATLNSR** |
| **1469.7** | **Trypsin** | **N** |  |  |  |
| **1165.6** | **Keratin** | **Y** | **1164.6** | **Keratin** | **LENEIQTYR** |
| **1179.6** | **Keratin** | **Y** | **1178.6** | **Keratin** | **YEELQITAGR** |
| **1300.5** | **Keratin** | **N** |  |  |  |
| **1716.9** | **Keratin** | **N** |  |  |  |

**Table 2.** The MALDI-ToF MS possible collagen and contaminant *m/z* markers compared to the LC-MS/MS sequence data for sample 15564. When a P is underlined it represents a hydroxylated proline. Contamination (trypsin and keratin) is shown in bold.

As well as predicting overall protein survival, both these methods can also give an indication of possible contamination. Recent studies have shown the presence of bacterial microbiomes in palaeontological bone specimens [33, 61], and the ability to provide evidence on the endogeneity of any peptides recovered is crucial to validate claims of proteins surviving further and further back in time. MALDI-ToF MS can provide direct identification of possible keratin peaks and chiral AA analysis can be used to identify bacterial contamination using Glx and Ala racemisation. Low racemisation values in samples of known ancient origin could also be an indication of modern contamination (particularly when associated with high AA concentrations), as well as an altered non-collagen like AA profile.

The two sequenced samples over 1 million years old both produced very few collagen peptides. The MALDI results for these samples were poor with no possible collagen peaks identified, and the AAR results showed signs of contamination with an altered AA profile and sample 15248 showing very low levels of Asx racemisation indicating modern contamination. Therefore, it is likely that the few collagen peptides recovered are exogenous. The *Megalocnus zile* sample (15564) that produced a good MALDI spectrum but poor LC-MS/MS data, showed very high racemisation of Glx (0.179) which was almost comparable to the Asx result of 0.196, even though Glx racemises at a much slower rate. This provides another line of evidence that the sample was contaminated.

## 5. Conclusion

Over 150 bone samples were analysed using a combination of ATR-FTIR, MALDI-ToF MS and chiral amino acid analysis. Twenty-two of these samples were also analysed using LC-MS/MS and the results of these analyses were used to evaluate the different methods for predicting protein survival (namely collagen). The oldest sample that produced good LC-MS/MS results dated to the Early Pleistocene (location unknown). This sample produced a good MALDI spectrum, had generally low levels of racemisation and a high glycine content, which can be used to help provide further evidence of surviving endogenous proteins. Three of the samples produced poor LC-MS/MS sequence data, with 2 of these samples dated to millions of years old (*Palaeosphenicus biloculata* and *Palaeeudyptes gunnari*). The left-over radiocarbon extracts from RLB all produced good sequence data and it is recommended that radiocarbon extracts are utilised in proteomic analysis.

ATR-FTIR has the benefit of not destroying the sample during analysis, which allows it to be used in conjunction with other analytical methods. Unfortunately, in this study, it has been shown to be unreliable for detecting surviving collagen in palaeontological samples using IRSF and C/P, due to the complex nature of fossilisation altering the bone mineral, and the potential for FTIR results to be affected by noise. Am/P is a more reliable indicator, however samples with a degraded amide peak may still hold valuable proteomic information. AAR and MALDI-ToF MS requires the loss of sample (albeit in small amounts, 1 mg for AA analysis and 15-30 mg for MALDI-ToF MS), but they give a better indication of protein survival, while being cheap and quick to carry out.

Based on this dataset, when conducting chiral amino acid analysis, samples that contain the following should be considered for further proteomic analysis:

* A total amino acid concentration over ~100,000 pmol/mg
* An amino acid composition high in glycine (>40%)
* Low overall racemisation values
* A proportionate D/L value for Glx and Ala in line with the relative racemisation rates of other amino acids, showing minimal bacterial contamination

When conducting MALDI-ToF MS as a screening method, samples with the highest intensity over ~1000 for a peak that is not the matrix used for the analysis, trypsin, or a contaminant (if identification is possible), and the spectrum contains more than 50 peaks should be considered for further proteomic analysis. One “poor” LC-MS/MS sequenced result also grouped with the samples that produced good proteomic recovery within the MALDI-ToF dataset and caution may be needed when screening species that are not within the PMF database. However, there was a clear grouping of the good sequence results which shows that this method will successfully filter out the majority of good samples from those with poor preservation. Proteins other than collagen were not able to be investigated, due to the limited number of NCPs identified.

Overall, this study has shown that chiral amino acid analysis and MALDI-ToF MS are both good screening methods for predicting endogenous protein survival in fossils. It is therefore recommended that one or both of these methods are used prior to further, more costly, proteomic analysis of palaeontological specimens.

Acknowledgements

SP wishes to thank Ioannis Kontopoulous for ATR-FTIR training and support, Sheila Taylor for AAR training and support, Beatrice Demarchi for useful discussion and support. Meaghan Mackie is thanked for LC-MS/MS analysis and valuable comments. The authors gratefully acknowledge the use of the Ultraflex III MALDI-ToF/ToF instrument in the York Centre of Excellence in Mass Spectrometry. SP and JS are grateful to Aisling Farrell, John Harris, Gary

Takeuchi and all the staff at Rancho La Brea for providing samples for analysis. JS also thanks Ben Fuller for assistance with collagen sample preparation. CAH is grateful to Oceanwide Expeditions, Vlissingen (NL) for support.

The authors are grateful for the comments of two anonymous reviewers; Reviewer 1 made numerous insightful comments which aided data interpretation and Reviewer 2 who encouraged us to include the Amide I/ phosphate ratio dataset.

Funding sources

This works was supported by the National Science Foundation (grant No. DEB 1547414), the Leverhulme Trust (grant No. PLP-2012-116) and the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Marie Curie Actions Grant Agreement No 676154

# References

[[1]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Welker, F., Collins, M.J., Thomas, J.A., Wadsley, M., Brace, S., Cappellini, E., Turvey, S.T., Reguero, M., Gelfo, J.N., Kramarz, A., Burger, J., Thomas-Oates, J., Ashford, D.A., Ashton, P.D., Rowsell, K., Porter, D.M., Kessler, B., Fischer, R., Baessmann, C., Kaspar, S., Olsen, J.V., Kiley, P., Elliott, J.A., Kelstrup, C.D., Mullin, V., Hofreiter, M., Willerslev, E., Hublin, J.-J., Orlando, L., Barnes, I., and MacPhee, R.D.E. (2015a) ‘Ancient Proteins Resolve the Evolutionary History of Darwin’s South American Ungulates’. *Nature* 522 (7554), 81–84](http://paperpile.com/b/inFlrI/pIfj)

[[2]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C.L., Woolley, J., Crisp, M.K., Wilson, J., Fotakis, A., Fischer, R., Kessler, B.M., Rakownikow Jersie-Christensen, R., Olsen, J.V., Haile, J., Thomas, J., Marean, C.W., Parkington, J., Presslee, S., Lee-Thorp, J., Ditchfield, P., Hamilton, J.F., Ward, M.W., Wang, C.M., Shaw, M.D., Harrison, T., Domínguez-Rodrigo, M., MacPhee, R.D.E., Kwekason, A., Ecker, M., Kolska Horwitz, L., Chazan, M., Kröger, R., Thomas-Oates, J., Harding, J.H., Cappellini, E., Penkman, K., and Collins, M.J. (2016) ‘Protein Sequences Bound to Mineral Surfaces Persist into Deep Time’. *eLife* [online] 5.](http://paperpile.com/b/inFlrI/1jEt)

[[3]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Presslee, S., Slater, G.J., Pujos, F., Forasiepi, A.M., Fischer, R., Molloy, K., Mackie, M., Olsen, J.V., Kramarz, A., Taglioretti, M., Scaglia, F., Lezcano, M., Lanata, J.L., Southon, J., Feranec, R., Bloch, J., Hajduk, A., Martin, F.M., Gismondi, R.S., Reguero, M., de Muizon, C., Greenwood, A., Chait, B.T., Penkman, K., Collins, M., and MacPhee, R.D.E. (2019) ‘Palaeoproteomics Resolves Sloth Relationships’. *Nature Ecology & Evolution* 1](http://paperpile.com/b/inFlrI/jcs2)

[4][Cappellini, E., Jensen, L.J., Szklarczyk, D., Ginolhac, A., da Fonseca, R.A.R., Stafford, T.W., Holen, S.R., Collins, M.J., Orlando, L., Willerslev, E., Gilbert, M.T.P., and Olsen, J.V. (2012) ‘Proteomic Analysis of a Pleistocene Mammoth Femur Reveals More than One Hundred Ancient Bone Proteins’. *Journal of Proteome Research* 11 (2), 917–926](http://paperpile.com/b/inFlrI/HjDV)

[[5]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Rybczynski, N., Gosse, J.C., Harington, C.R., Wogelius, R.A., Hidy, A.J., and Buckley, M. (2013) ‘Mid-Pliocene Warm-Period Deposits in the High Arctic Yield Insight into Camel Evolution’. *Nature Communications* 4, 1550](http://paperpile.com/b/inFlrI/WQkw)

[6][Warinner, C., Rodrigues, J.F.M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-García, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A.S., Barkow-Oesterreicher, S., Teoh, K.Y., Kelstrup, C.D., Olsen, J.V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C.M., Jr, Collins, M.J., Gilbert, M.T.P., Rühli, F., and Cappellini, E. (2014) ‘Pathogens and Host Immunity in the Ancient Human Oral Cavity’. *Nature Genetics* 46 (4), 336–344](http://paperpile.com/b/inFlrI/Esqd)

[7][Cleland, T.P., Schroeter, E.R., Feranec, R.S., and Vashishth, D. (2016) ‘Peptide Sequences from the First Castoroides Ohioensis Skull and the Utility of Old Museum Collections for Palaeoproteomics’. *Proceedings. Biological Sciences / The Royal Society* [online] 283 (1832). available from <](http://paperpile.com/b/inFlrI/whVC)<http://dx.doi.org/10.1098/rspb.2016.0593>[>](http://paperpile.com/b/inFlrI/whVC)

[8][Sawafuji, R., Cappellini, E., Nagaoka, T., Fotakis, A.K., Jersie-Christensen, R.R., Olsen, J.V., Hirata, K., and Ueda, S. (2017) ‘Proteomic Profiling of Archaeological Human Bone’. *Royal Society Open Science* 4 (6), 161004](http://paperpile.com/b/inFlrI/0qYM)

[9][Cappellini, E., Welker, F., Pandolfi, L., and Madrigal, J.R. (2018) ‘Early Pleistocene Enamel Proteome Sequences from Dmanisi Resolve Stephanorhinus Phylogeny’. *bioRxiv* [online] available from <](http://paperpile.com/b/inFlrI/H1jF)<https://www.biorxiv.org/content/early/2018/09/10/407692.abstract>[>](http://paperpile.com/b/inFlrI/H1jF)

[10][Lingham-Soliar, T. and Wesley-Smith, J. (2008) ‘First Investigation of the Collagen D-Band Ultrastructure in Fossilized Vertebrate Integument’. *Proceedings. Biological Sciences / The Royal Society* 275 (1648), 2207–2212](http://paperpile.com/b/inFlrI/0AlyZ)

[11][Zylberberg, L. and Laurin, M. (2011/7) ‘Analysis of Fossil Bone Organic Matrix by Transmission Electron Microscopy’. *Comptes Rendus. Palevol* 10 (5–6), 357–366](http://paperpile.com/b/inFlrI/0iwoa)

[12][Piga, G., Santos-Cubedo, A., Brunetti, A., Piccinini, M., Malgosa, A., Napolitano, E., and Enzo, S. (2011) ‘A Multi-Technique Approach by XRD, XRF, FT-IR to Characterize the Diagenesis of Dinosaur Bones from Spain’. *Palaeogeography, Palaeoclimatology, Palaeoecology* 310 (1–2), 92–107](http://paperpile.com/b/inFlrI/V6hRc)

[13][Edwards, N.P., Barden, H.E., van Dongen, B.E., Manning, P.L., Larson, P.L., Bergmann, U., Sellers, W.I., and Wogelius, R.A. (2011) ‘Infrared Mapping Resolves Soft Tissue Preservation in 50 Million Year-Old Reptile Skin’. *Proceedings. Biological Sciences / The Royal Society* 278 (1722), 3209–3218](http://paperpile.com/b/inFlrI/MliZ4)

[14][Lee, Y.-C., Chiang, C.-C., Huang, P.-Y., Chung, C.-Y., Huang, T.D., Wang, C.-C., Chen, C.-I., Chang, R.-S., Liao, C.-H., and Reisz, R.R. (2017) ‘Evidence of Preserved Collagen in an Early Jurassic Sauropodomorph Dinosaur Revealed by Synchrotron FTIR Microspectroscopy’. *Nature Communications* 8, 14220](http://paperpile.com/b/inFlrI/EYpz2)

[15][Hollund, H.I., Ariese, F., Fernandes, R., Jans, M.M.E., and Kars, H. (2013) ‘Testing an Alternative High-Throughput Tool for Investigating Bone Diagenesis: FTIR in Attenuated Total Reflection (ATR) Mode’. *Archaeometry* 55 (3), 507–532](http://paperpile.com/b/inFlrI/ZF90x)

[16][Chadefaux, C., Le Hô, A.-S., Bellot-Gurlet, L., and Reiche, I. (2009) ‘Curvefitting Micro-ATR-FTIR Studies of the Amide I and II Bands of Type I Collagen in Archaeological Bone Materials’. *E-PRESERVATION Science* 6, 129–137](http://paperpile.com/b/inFlrI/PE0Qo)

[17][Beasley, M.M., Bartelink, E.J., Taylor, L., and Miller, R.M. (2014) ‘Comparison of Transmission FTIR, ATR, and DRIFT Spectra: Implications for Assessment of Bone Bioapatite Diagenesis’. *Journal of Archaeological Science* 46, 16–22](http://paperpile.com/b/inFlrI/YiZm7)

[18] Thompson, T.J.U., Islam, M., Piduru, K., and Marcel, A. (2011) ‘An Investigation into the Internal and External Variables Acting on Crystallinity Index Using Fourier Transform Infrared Spectroscopy on Unaltered and Burned Bone’. Palaeogeography, Palaeoclimatology, Palaeoecology 299 (1–2), 168–174

[19][Buckley, M., Collins, M., Thomas-Oates, J., and Wilson, J.C. (2009) ‘Species Identification by Analysis of Bone Collagen Using Matrix-Assisted Laser Desorption/ionisation Time-of-Flight Mass Spectrometry’. *Rapid Communications in Mass Spectrometry: RCM* 23 (23), 3843–3854](http://paperpile.com/b/inFlrI/hnHwh)

[20][Buckley, M., Fraser, S., Herman, J., Melton, N.D., Mulville, J., and Pálsdóttir, A.H. (2014) ‘Species Identification of Archaeological Marine Mammals Using Collagen Fingerprinting’. *Journal of Archaeological Science* 41 (Supplement C), 631–641](http://paperpile.com/b/inFlrI/mwbGs)

[21][Welker, F., Soressi, M., Rendu, W., Hublin, J.-J., and Collins, M. (2015b) ‘Using ZooMS to Identify Fragmentary Bone from the Late Middle/Early Upper Palaeolithic Sequence of Les Cottes, France’. *Journal of Archaeological Science* 54, 279–286](http://paperpile.com/b/inFlrI/4EW3I)

[22][Harvey, V.L., Daugnora, L., and Buckley, M. (2018)](http://paperpile.com/b/inFlrI/q2iM5) *‘*[*Species Identification of Ancient Lithuanian Fish Remains Using Collagen Fingerprinting*](http://paperpile.com/b/inFlrI/q2iM5)*’*[. *Journal of Archaeological Science* 98, 102–111](http://paperpile.com/b/inFlrI/q2iM5)

[23][Welker, F., Smith, G.M., Hutson, J.M., Kindler, L., Garcia-Moreno, A., Villaluenga, A., Turner, E., and Gaudzinski-Windheuser, S. (2017) ‘Middle Pleistocene Protein Sequences from the Rhinoceros Genus Stephanorhinus and the Phylogeny of Extant and Extinct Middle/Late Pleistocene Rhinocerotidae’. *PeerJ* 5, e3033](http://paperpile.com/b/inFlrI/xOkFo)

[[24]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Buckley, M., Lawless, C., and Rybczynski, N. (2019) ‘Collagen Sequence Analysis of Fossil Camels, Camelops and C.f. Paracamelus, from the Arctic and Sub-Arctic of Plio-Pleistocene North America’. *Journal of Proteomics* 194, 218–225](http://paperpile.com/b/inFlrI/inpGk)

[[25]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Abelson, P.H. (1955) ‘Paleobiochemistry: Organic Constituents of Fossils’. in *Carnegie Institution of Washington, Yearbook, No. 54*. Carnegie Institution of Washington, 107–109](http://paperpile.com/b/inFlrI/eyUXu)

[[26]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Abelson, P.H. (1954) ‘Amino Acids in Fossils’. *Science* 119 (3096), 576](http://paperpile.com/b/inFlrI/Fnysn)

[27][Jones, J.D. and Vallentyne, J.R. (1960) ‘Biogeochemistry of Organic matter—I: Polypeptides and Amino Acids in Fossils and Sediments in Relation to Geothermometry’. *Geochimica et Cosmochimica Acta* 21 (1-2), 1–34](http://paperpile.com/b/inFlrI/lllfJ)

[[28]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Penkman, K.E.H., Preece, R.C., Bridgland, D.R., Keen, D.H., Meijer, T., Parfitt, S.A., White, T.S., and Collins, M.J. (2011) ‘A Chronological Framework for the British Quaternary Based on Bithynia Opercula’. *Nature* 476 (7361), 446–449](http://paperpile.com/b/inFlrI/Az70s)

[[29]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Demarchi, B., Williams, M.G., Milner, N., Russell, N., Bailey, G., and Penkman, K. (2011) ‘Amino Acid Racemization Dating of Marine Shells: A Mound of Possibilities’. *Quaternary International: The Journal of the International Union for Quaternary Research* 239 (1-2), 114–124](http://paperpile.com/b/inFlrI/RKLvu)

[30][Crisp, M., Demarchi, B., Collins, M., Morgan-Williams, M., Pilgrim, E., and Penkman, K. (2013) ‘Isolation of the Intra-Crystalline Proteins and Kinetic Studies in Struthio Camelus (ostrich) Eggshell for Amino Acid Geochronology’. *Quaternary Geochronology* 16 , 110–128](http://paperpile.com/b/inFlrI/RLdA4)

[[31]](https://paperpile.com/c/inFlrI/pIfj+xOkFo+inpGk)[Dickinson, M.R., Lister, A.M., and Penkman, K.E.H. (2019) ‘A New Method for Enamel Amino Acid Racemization Dating: A Closed System Approach’. *Quaternary Geochronology* 50, 29–46](http://paperpile.com/b/inFlrI/a00xP)

[[32]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Manning, P.L., Morris, P.M., McMahon, A., Jones, E., Gize, A., Macquaker, J.H.S., Wolff, G., Thompson, A., Marshall, J., Taylor, K.G., and Others (2009) ‘Mineralized Soft-Tissue Structure and Chemistry in a Mummified Hadrosaur from the Hell Creek Formation, North Dakota (USA)’. *Proceedings of the Royal Society of London B: Biological Sciences*](http://paperpile.com/b/inFlrI/RESDQ) 276**,** 3429–3437

[[33]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Saitta, E.T., Liang, R., Lau, M.C.Y., Brown, C.M., Longrich, N.R., Kaye, T.G., Novak, B.J., Salzberg, S.L., Norell, M.A., Abbott, G.D., Dickinson, M.R., Vinther, J., Bull, I.D., Brooker, R.A., Martin, P., Donohoe, P., Knowles, T.D.J., Penkman, K.E.H., and Onstott, T.C. (2019) ‘Cretaceous Dinosaur Bone Contains Recent Organic Material and Provides an Environment Conducive to Microbial Communities’. *eLife* [online]](http://paperpile.com/b/inFlrI/IxhQY)

[34][Bada, J.L., Wang, X.S., Poinar, H.N., Paabo, S., and Poinar, G.O. (1994) ‘Amino Acid Racemization in Amber-Entombed Insects: Implications for DNA Preservation’. *Geochimica et Cosmochimica Acta* 58 (14), 3131–3135](http://paperpile.com/b/inFlrI/99jpQ)

[[35]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Dobberstein, R.C., Collins, M.J., Craig, O.E., Taylor, G., Penkman, K.E.H., and Ritz-Timme, S. (2009) ‘Archaeological Collagen: Why Worry about Collagen Diagenesis?’ *Archaeological and Anthropological Sciences* 1 (1), 31–42](http://paperpile.com/b/inFlrI/e9LnZ)

[[36]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Bhattacharyya, S.K. and Banerjee, A.B. (1974) ‘D-Amino Acids in the Cell Pool of Bacteria’. *Folia Microbiologica* 19 (1), 43–50](http://paperpile.com/b/inFlrI/rXEz)

[[37]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[McCoy, V.E., Gabbott, S.E., Penkman, K., Collins, M.J., Presslee, S., Holt, J., Grossman, H., Wang, B., Solórzano Kraemer, M.M., Delclòs, X., and Peñalver, E. (2019) ‘Ancient Amino Acids from Fossil Feathers in Amber’. *Scientific Reports* 9 (1), 6420](http://paperpile.com/b/inFlrI/VzsQ)

[[38]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Brock, F., Geoghegan, V., Thomas, B., Jurkschat, K., and Higham, T.F.G. (2013) ‘Analysis of Bone “Collagen” Extraction Products for Radiocarbon Dating’. *Radiocarbon* 55 (2), 445–463](http://paperpile.com/b/inFlrI/mdzJa)

[[39]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Wadsworth, C. and Buckley, M. (2017) ‘Characterization of Proteomes Extracted through Collagen-Based Stable Isotope and Radiocarbon Dating Methods’. *Journal of Proteome Research* 17, 429-439](http://paperpile.com/b/inFlrI/oYN8Q)

[[40]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Cersoy, S., Zirah, S., Marie, A., and Zazzo, A. (2018) *Toward a Versatile Protocol for Radiocarbon and Proteomics Analysis of Ancient Collagen*. 101, 1–10](http://paperpile.com/b/inFlrI/WhDrf)

[[41]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Kontopoulos, I., Presslee, S., Penkman, K., and Collins, M.J. (2018) ‘Preparation of Bone Powder for FTIR-ATR Analysis: The Particle Size Effect’. *Vibrational Spectroscopy* 99, 167–177](http://paperpile.com/b/inFlrI/Bszal)

[[42]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[High, K., Milner, N., Panter, I., and Penkman, K.E.H. (2015) ‘Apatite for Destruction: Investigating Bone Degradation due to High Acidity at Star Carr’. *Journal of Archaeological Science* 59, 159–168](http://paperpile.com/b/inFlrI/5Pjl)

[[43]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Kaufman, D.S. and Manley, W.F. (1998) ‘A New Procedure for Determining Dl Amino Acid Ratios in Fossils Using Reverse Phase Liquid Chromatography’. *Quaternary Science Reviews* 17 (11), 987–1000](http://paperpile.com/b/inFlrI/WrVam)

[[44]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) [Hill, R.L. (1965) ‘Hydrolysis of Proteins’. *Advances in Protein Chemistry* 20, 37–107](http://paperpile.com/b/inFlrI/oa12Z)

[[45]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[van Doorn, N.L., Hollund, H., and Collins, M.J. (2011) ‘A Novel and Non-Destructive Approach for ZooMS Analysis: Ammonium Bicarbonate Buffer Extraction’. *Archaeological and Anthropological Sciences* 3 (3), 281–289](http://paperpile.com/b/inFlrI/Bkw1w)

[[46]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Fuller, B.T., Harris, J.M., Farrell, A.B., Takeuchi, G., and Southon, J. (2015) ‘Sample Preparation for Radiocarbon Dating and Isotopic Analysis of Bone from Rancho La Brea’. In: Ed. JM Harris *La Brea and beyond: The Paleontology of Asphalt-Preserved Biotas,. Natural History Museum of Los Angeles County, Science Series* (42), USA: Natural History Museum of Los Angeles County 151–167](http://paperpile.com/b/inFlrI/ZyBI)

[[47]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Strohalm, M., Kavan, D., Novák, P., Volný, M., and Havlícek, V. (2010) ‘mMass 3: A Cross-Platform Software Environment for Precise Analysis of Mass Spectrometric Data’. *Analytical Chemistry* 82 (11), 4648–4651](http://paperpile.com/b/inFlrI/NNqGn)

[[48]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Jonuks, T., Oras, E., Best, J., Demarchi, B., Mänd, R., Presslee, S., and Vahur, S. (2018) ‘Multi-Method Analysis of Avian Eggs as Grave Goods: Revealing Symbolism in Conversion Period Burials at Kukruse, NE Estonia’. *Environmental Archaeology* 23 (2), 109–122](http://paperpile.com/b/inFlrI/hPTUT)

[[49]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L., and Mallick, P. (2012) ‘A Cross-Platform Toolkit for Mass Spectrometry and Proteomics’. *Nature Biotechnology* 30 (10), 918–920](http://paperpile.com/b/inFlrI/6NBId)

[[50]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., Zhang, W., Zhang, Z., Lajoie, G.A., and Ma, B. (2012) ‘PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification’. *Molecular & Cellular Proteomics: MCP* 11 (4), M111.010587](http://paperpile.com/b/inFlrI/isDVz)

[[51]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C., and Collins, M.J. (2018) ‘A Guide to Ancient Protein Studies’. *Nature Ecology & Evolution* 1](http://paperpile.com/b/inFlrI/5l7eg)

[52]Steadman, D.W., Martin, P.S., MacPhee, R.D.E., Jull, A.J.T., McDonald, H.G., Woods, C.A., Iturralde-Vinent, M., and Hodgins, G.W.L. (2005) ‘Asynchronous Extinction of Late Quaternary Sloths on Continents and Islands’. *Proceedings of the National Academy of Sciences of the United States of America* 102 (33), 11763–11768

[[53]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Adamiano, A., Fabbri, D., Falini, G., and Giovanna Belcastro, M. (2013) ‘A Complementary Approach Using Analytical Pyrolysis to Evaluate Collagen Degradation and Mineral Fossilisation in Archaeological Bones: The Case Study of Vicenne-Campochiaro Necropolis (Italy)’. *Journal of Analytical and Applied Pyrolysis* 100, 173–180](http://paperpile.com/b/inFlrI/8I67)

[[54]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[King, C.L., Tayles, N., and Gordon, K.C. (2011) ‘Re-Examining the Chemical Evaluation of Diagenesis in Human Bone Apatite’. *Journal of Archaeological Science* 38 (9), 2222–2230](http://paperpile.com/b/inFlrI/q1dxs)

[[55]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Kontopoulos, I., Penkman, K., McAllister, G.D., Lynnerup, N., and Collins, M.J. (2019) *Petrous Bone Diagenesis: A Multi-Analytical Approach*. [online] available from <](http://paperpile.com/b/inFlrI/pQ3k)<https://www.researchgate.net/publication/330234441_Petrous_bone_diagenesis_a_multi-analytical_approach>[> [18 January 2019]](http://paperpile.com/b/inFlrI/pQ3k)

[[56]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) [Lebon, M., Reiche, I., Bahain, J.-J., Chadefaux, C., Moigne, A.-M., Fröhlich, F., Sémah, F., Schwarcz, H.P., and Falguères, C. (2010) ‘New Parameters for the Characterization of Diagenetic Alterations and Heat-Induced Changes of Fossil Bone Mineral Using Fourier Transform Infrared Spectrometry’. *Journal of Archaeological Science* 37 (9), 2265–2276](http://paperpile.com/b/inFlrI/FotkO)

[[57]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY) [Trueman, C.N., Palmer, M.R., Field, J., Privat, K., Ludgate, N., Chavagnac, V., Eberth, D.A., Cifelli, R., and Rogers, R.R. (2008) ‘Comparing Rates of Recrystallisation and the Potential for Preservation of Biomolecules from the Distribution of Trace Elements in Fossil Bones’. *Comptes Rendus. Palevol* 7 (2–3), 145–158](http://paperpile.com/b/inFlrI/zP7p7)

[58] Lebon, M., Reiche, I., Gallet, X., Bellot-Gurlet, L., and Zazzo, A. (2016) ‘Rapid Quantification of Bone Collagen Content by ATR-FTIR Spectroscopy’. *Radiocarbon* 1–15

[59] Pothier Bouchard, G., Mentzer, S.M., Riel-Salvatore, J., Hodgkins, J., Miller, C.E., Negrino, F., Wogelius, R., and Buckley, M. (2019) ‘Portable FTIR for on-Site Screening of Archaeological Bone Intended for ZooMS Collagen Fingerprint Analysis’*. Journal of Archaeological Science: Reports* 26, 101862

[60] Naito, Y.I., Yamane, M., and Kitagawa, H. (2020) ‘A Protocol for Using Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy for Pre-Screening Ancient Bone Collagen prior to Radiocarbon Dating’. *Rapid Communications in Mass Spectrometry* 34 (10), e8720

[61] Liang, R., Lau, M.C.Y., Saitta, E.T., Garvin, Z.K., and Onstott, T.C. (2020) ‘Genome-Centric Resolution of Novel Microbial Lineages in an Excavated Centrosaurus Dinosaur Fossil Bone from the Late Cretaceous of North America’. *Environmental Microbiome* 15 (1), 8

[[62]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Lebon, M., Zazzo, A., and Reiche, I. (2014) ‘Screening in Situ Bone and Teeth Preservation by ATR-FTIR Mapping’. *Palaeogeography, Palaeoclimatology, Palaeoecology* 416, 110–119](http://paperpile.com/b/inFlrI/53Jq8)

[[63]](https://paperpile.com/c/inFlrI/RESDQ+1jEt+IxhQY)[Aufort, J., Lebon, M., Gallet, X., Ségalen, L., Gervais, C., Brouder, C., and Balan, E. (2018) ‘Macroscopic Electrostatic Effects in ATR-FTIR Spectra of Modern and Archeological Bones’. *The American Mineralogist* 103 (2), 326–329](http://paperpile.com/b/inFlrI/7uSHd)