Extinct species identification from late middle Pleistocene and earlier Upper Pleistocene bone fragments and tools not recognizable from their osteomorphological study by an enhanced proteomics protocol

Fabrice Bray, Stéphanie Flament, Grégory Abrams, Dominique Bonjean, Christian Rolando, Caroline Tokarski, Patrick Auguste

Extinct species identification from late middle Pleistocene and earlier Upper Pleistocene bone fragments and tools not recognizable from their osteomorphological study by an enhanced proteomics protocol

Fabrice Bray[1] | Stéphanie Flament[1] | Grégory Abrams[2,3] | Dominique Bonjean[2,4] | Christian Rolando[1,5] | Caroline Tokarski[6] | Patrick Auguste[7]

[1]CNRS, UAR 3290 - MSAP, University of Lille, Lille, France

[2]Espace muséal d'Andenne, Sclayn, Belgium

[3]Faculty of Archaeology, Leiden University, Van Steenis Building, Leiden, The Netherlands

[4]Department of Prehistory, University of Liège, Liège, Belgium

[5]Shrieking Sixties, Villeneuve-d'Ascq, France

[6]University of Bordeaux, Bordeaux, France

[7]Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paléo, Unité Evolution, Lille, France

Correspondence

Fabrice Bray, CNRS, UAR 3290 - MSAP, University of Lille, F-59000 Lille, France.
Email: fabrice.bray@univ-lille.fr

ABSTRACT

Ancient preserved molecules offer the opportunity of gaining a deeper knowledge on their biological past. However, the development of a proteomic workflow remains a challenge. The analysis of fossils must involve a low quantity of material to avoid damaging the samples. In this study an enhanced proteomic pro- tocol was applied to 5-milligram samples of about 130,000-year-old mammalian bones ranging from the end of the Middle Pleistocene up to the earlier Upper Pleistocene, excavated from Scladina Cave (Sclayn, Belgium). Using sequence homology with modern sequences, a biological classification was successfully achieved and the associated taxonomic ranks to each bone were identified consistently with the information gained from osteomorphological studies and pal- aeoenvironmental and palaeodietary data. Amino acid substitutions on collagens were identified, thus provid- ing new information on extinct species sequences and helping in taxonomy-based clustering. Considering samples with no osteomorphological information, such as two fragments of bone retouchers, proteomics suc- cessfully identified the families providing paleontolo- gists new information on these objects. Combining osteomorphology studies and amino acid variations identified by proteomics, one of the retouchers was potentially identified as belonging to the Ursus spelaeus species.

KEY WORDS
amino acids substitution, mammalian bones, mass spectrometry, palaeoproteomics, peptides modifications, pleistocene, scladina cave, taxonomy

INTRODUCTION

Mass spectrometry (MS) based studies on bones have become an increasingly popular method. Indeed, bones are real safes protecting their constitutive proteins. MS has been proven to be a robust and accurate method allowing for the identification of proteins, their biological origins, and their modifications (Cleland & Schroeter, 2018; Dallongeville et al., 2016; Vinciguerra et al., 2016; Welker, 2018), overpassing DNA analysis considering the longer temporal scales (Cappellini et al., 2018; Demarchi et al., 2016; Schroeter et al., 2022). For example, the identification of collagen I from *Brachylophosaurus canadensis* dinosaurs from an earlier seminal study (Schweitzer et al., 2009) was confirmed (Schroeter et al., 2017). Numerous Plio-Pleistocene mammalian collagen sequences were identified from Camelidae fossil *Camelops* and c.f. *Paracamelus*, (Buckley et al., 2019), *Gigantopithecus blacki* (Welker et al., 2019), and Eurasian Rhinocerotidae (Cappellini et al., 2019). Studies on older bones and especially on dinosaur fossils are very difficult as they are hampered by the presence of contaminants. These studies show the importance of sample preparation in order to avoid contamination, of control quality of mass spectrometry analysis, and of the validation of data treatment (Buckley et al., 2008; Buckley, Warwood, et al., 2017).

The first step of a classical palaeoproteomic analysis is demineralization. The extracted proteins are digested into peptides using an enzyme, commonly trypsin. Several digestion methods have been reported, such as liquid digestion (Horn et al., 2019; Sawafuji et al., 2017), filter assisted sample preparation (FASP) (Cappellini et al., 2014; Kostyukevich et al., 2018), solid digestion of demineralized bones (Cleland, 2018b), and microwave digestion (Colleary et al., 2020). Recently a single-pot, solid-phase-enhanced sample preparation (SP3), an effective method allowing for collagen and non-collagenous proteins (NCPs) analyses together with the removal of co-extracted humic compounds, has been proposed (Cleland, 2018a). Multiprotease approaches aimed at increasing the coverage of protein sequences have also been described (Lanigan et al., 2020).

Two techniques are used for identifying peptides extracted from bones: ZooMS (Zooarchaeology by Mass Spectrometry) and LC–MS/MS-based shotgun analysis. Developed in 2009, ZooMS (Buckley et al., 2009) is based on peptide mass fingerprint analysis using MALDI-TOF MS. In MALDI-TOF mass fingerprinting, the mass spectrum is matched against the theoretical digestion of proteins present in the database without sequencing the peptides. ZooMS allows for taxonomic discrimination between a wide range of mammalian and non-mammalian bones through the identification of peptides substitutions among species. ZooMS allowed for the identification of various mammals species, such as hominin (Brown et al., 2016), marine mammals (Hofman et al., 2018), and other mammals such as sheep, goats (Birch et al., 2019; Brandt et al., 2018; Buckley et al., 2010), and kangaroos (Buckley, Cosgrove, et al., 2017), from bones but also from derived materials such as bone tools (Desmond et al., 2018; Martisius et al., 2020; McGrath et al., 2019). ZooMS has also been applied to non-mammals such as fish (Richter et al., 2011), reptiles (Harvey et al., 2018), and amphibians (Harvey et al., 2019). On the contrary, in LC–MS/MS-based shotgun analysis amino acid

sequences of peptides are recognized, thus enabling an accurate identification of proteins, including species-related amino acids substitutions. In the early 2010s the first application of shotgun analysis by LC–MS on a 43,000-year-old woolly mammoth bone (Cappellini et al., 2011) resulted in the identification of a hundred proteins. Since then, the method has been extensively applied to the study of other extinct species such as *Castoroides ohioensis* (Cleland et al., 2016), *Bison latifrons* (Hill et al., 2015), ancient birds (Cleland, Schroeter, & Schweitzer, 2015; Horn et al., 2019), and dinosaurs (Cleland, Schroeter, Zamdborg, et al., 2015; Schroeter et al., 2017). Peptide sequencing used in LC–MS/MS-based proteomics gives access to the phylogeny and evolutionary history of ancient species. This approach has been applied to the study of endemic South American 'ungulates' (Buckley, 2015; Welker et al., 2015), ancient hominin specimens (Welker et al., 2016), and tree sloths (Presslee et al., 2019).

We present here an enhanced proteomics protocol developed on LC–MS/MS-based shotgun for the study of extinct species, which has been applied to the identification of bones from the Scladina Cave (Andenne, Belgium). The Scladina Cave is located 400 m southwest of the village of Sclayn, between Namur and Andenne, on the right bank of the Meuse from which it is approximately 400 m away, along a secondary tributary (Figure 1a,b). The site has been under scientific excavation since 1978 and has delivered numerous Neanderthal occupations, Neanderthal remains, upper Paleolithic artefacts, as well as a plural Neolithic burial. The bone artefacts presented in this research were recovered among the faunal remains associated to the main Middle Paleolithic occupation of the site (Unit 5), which delivered an important bone tool assemblage made from cave bear remains (Abrams et al., 2014). The production of bone tools involves fracturing them in order to adapt their size to a particular use (Abrams, 2015). This fracturing makes their taxonomic identification difficult and therefore requires innovative methods in order to better understand Neanderthals and their management of exploited animal resources. To assert the results provided by the bone tools, other faunal remains were also studied in this paper. They were recovered in the lower part of the stratigraphic sequence of Scladina Cave, above and below Unit 5 (Unit 4B to Unit 6C). This lower sequence could be related to the end of the Middle Pleistocene up to the Upper Pleistocene (Pirson et al., 2008, 2014). The interest of these remains lies in the fact that these bones are more complete than the bone artifacts and therefore allowed for their classical identification to be compared our proteomic identification method.

Paleoproteomic analysis is generally performed on a few milligrams of bone samples. The starting material amount is low in comparison to other techniques usually applied to bone analysis, such as ^{14}C dating (from several hundreds of mg up to a g). Nevertheless, the amount required for the proteomics experiment still needs to be reduced to be minimally destructive for the analysis of bone artifacts preserved in museums. Two major improvements have been introduced in this proteomics workflow for increasing the sensitivity. The first one concerns demineralization, which is performed under conditions introducing less artificial post-translational modifications, and the second one being digestion, performed directly on bone powder without protein solubilization. This enhanced proteomics workflow described here works on less than 5 mg of sample allowing to investigate museum artifacts while preserving the full identification potential of proteomics. Eight ancient bones identified from osteomorphological studies and two fragments of bone retouchers with no osteomorphological information were studied. Our analyses resulted in a successful characterization of biological classification and associated taxonomic ranks for each bone, consistently with the information gained from morphological studies. Taxa identifications were also provided for the retouchers, which are not identifiable from their morphology, in agreement with the fauna found at the site. It can be pointed out that during this study, amino acid substitutions on collagens were identified providing new information on extinct species sequences and also helping in taxonomy-based clustering.

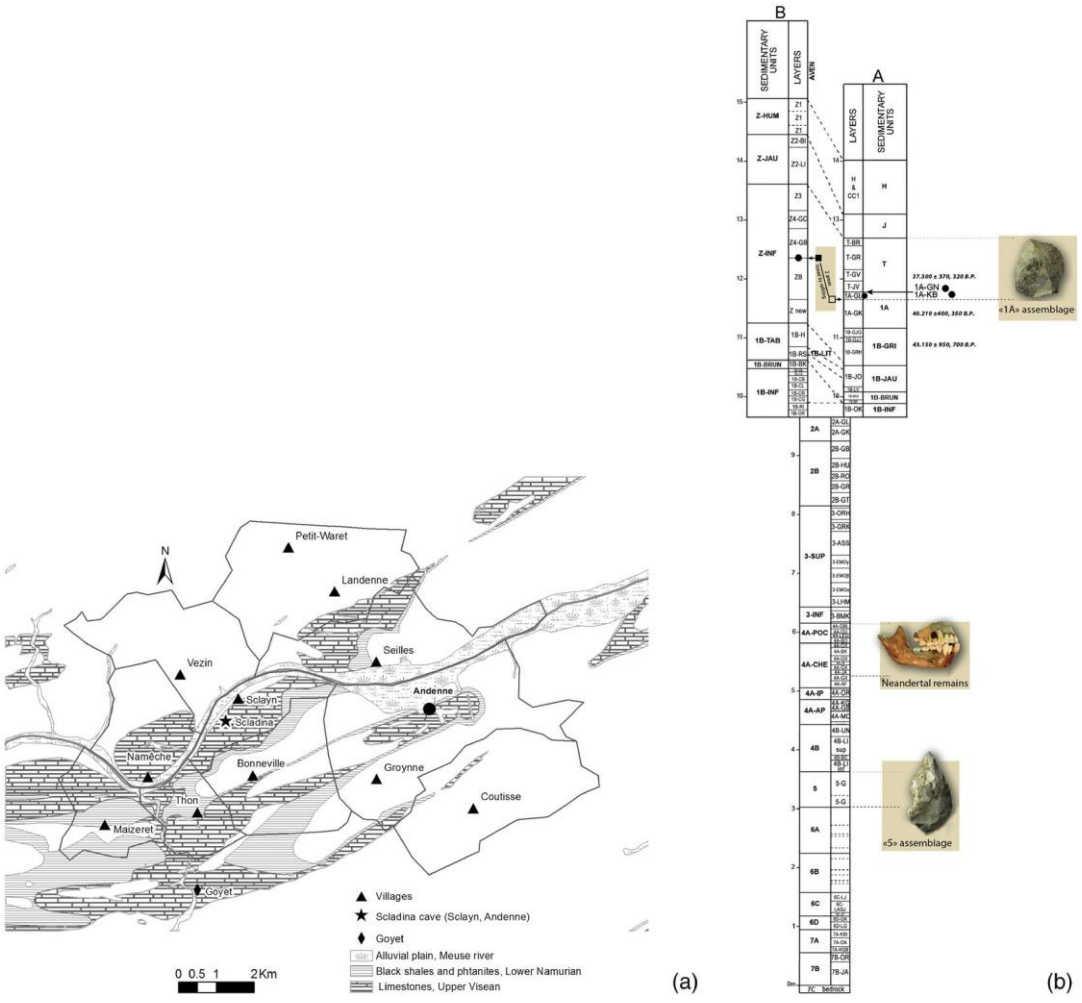

FIG URE 1 (a) Location of Scladina Cave. Grey areas: Distribution of the Palaeozoic carbonated rocks in Belgium. (b) Stratigraphic profile of Scladina cave composed of the entrance-sequence (lower and upper portions of a) and the aven sequence (upper portion of B; adapted from Bonjean and colleagues) (Bonjean et al., 2015)

MATERIALS AND METHODS

Samples

Ten bones excavated in Scladina Cave were selected from different stratigraphic units (Units 4B, 5, 6A, 6B, 6C, 4B). Several dating methods were used to estimate the age of this part of the stratigraphic sequence, which is older than 130,000 years old (Abrams et al., 2010; Pirson et al., 2014) (Bonjean et al., 2014). The faunal assemblage from Scladina Cave is presented in Supplementary Data 1. The species of 8 bones (Sc-1 to Sc-8) out of 10 were identified through their anatomy. The two remaining ones (Sc-9, Sc-10), having either unknown or hypothetical origins, are of particular interest, because they are directly associated with bone retouchers by refitting (Abrams et al., 2014). A retoucher is generally an object used for modifying a stone chip by striking it. Bone retouchers are used to reshape lithic artifacts. The repeated contact of

bone and lithic edges impresses the bone with microstriations, within which small lithic flakes can be found as evidence of their use (Pirson et al., 2014). Table 1 is referencing the 10 studied samples with archaeological information. Photographs of the studied bones are available in the Supplementary Data 1 Figure S1.1 to S1.10. All these bones are preserved in Espace muséal d'Andenne. The accession number of the bone in Espace muséal d'Andenne corresponds to the concatenation of excavation year and number (for example entry one for Sc-1 is referenced as 0.1983–236-52).

Preparation and analysis of archaeological bones

Bone fragments were mechanically ground into fine powder with a pestle in an agate mortar. Five milligrams of bone powder were demineralized with 5% trifluoroacetic acid (TFA) in water. The demineralization solution was evaporated for proteomics analysis. The bone powder was delipided (Folch et al., 1957). eFASP digestion was performed on both the bone powder and the demineralized fraction (Erde et al., 2014). The peptides obtained after digestion were analyzed by LC–MS/MS (Helle et al., 2019). Proteins and peptides were identified by bioinformatics analysis. More detailed experimental protocols are available in supporting information (Supplementary_data_S1).

RESULTS

Proteomics methodology and identified proteins

We applied an enhanced LC–MS/MS shotgun proteomics method to the selected bones (Supplementary_data_1). The new method starts from 5 mg of bones. A first delipidation step by the Folch method has been introduced to remove the lipids, which could interfere with the following extraction steps. We used an aqueous 5% TFA demineralizing solution. TFA is a commonly employed solvent in reverse-phase HPLC protein purification because of its effectiveness in solubilizing hydrophobic peptides, thus increasing coverage. Using TFA, proteins or peptides are less modified as the pKa for TFA is lower than the pKa for HCl, the acid commonly used in paleoproteomics for demineralization (0.3 and –6.3 respectively). After the TFA demineralization step two fractions are obtained: the bone powder fraction and the

TABLE 1 Bone samples from the Scladina cave

Excavation year	Excavation number	Geological unit	Location (square)	Description	Species	Sample name
1983	236–52	5	E13	Calcaneus	*Panthera spelaea*	Sc-1
2011	362–7	6C-LAGJ	G16	Metatarsal II	*Panthera spelaea*	Sc-2
1983	407–93	5	E12	Metatarsal III	*Panthera spelaea*	Sc-3
1981	127–3	6A	C4	Humerus	*Panthera pardus*	Sc-4
2013	46–2	6B	C37	Metatarsal III	*Ursus spelaeus*	Sc-5
2013	44–8	6A	D37	Metatarsal I	*Ursus thibetanus*	Sc-6
2003	590–2	6A	B36	Metatarsal III	*U. thibetanus*	Sc-7
1995	430–97	4B	C36	Rib	*Mammuthus* sp	Sc-8
1987	52	5	I22	Retoucher	Unknown	Sc-9
1985	F16–17	5	F16	Retoucher	Unknown	Sc-10

demineralization residual fraction. Both fractions were digested with the eFASP method to increase sensitivity, recovery, and coverage of the identified proteins, and to reduce chemical contamination (Erde et al., 2014). Contrary to most of the previous publications, protein extraction was not carried out before digestion. Trypsin digestion was performed directly on both the bone powder and the demineralized residual fractions. For all samples and fractions, the majority of the identified peptides matches mammalian collagen I alpha 1 (COL1A1) and collagen I alpha 2 (COL1A2) (Table S3.1). Considering COL1A1 and COL1A2 proteins, the peptides are distributed all along the sequences. Excluding the signal peptide and *N, C*-terminus propeptides, sequence coverage for the best protein hit, which is COL1A1, is roughly similar for the bone powder and the demineralized residual fractions, and it is equal to 62% and 63% respectively (Table S3.1). The mean of the Mascot protein COL1A1 protein best hit score is 7,863 for bone powder fractions and 7,014 for demineralized residual fractions. The difference between Mascot scores for both fractions is significantly different, showing that more proteins are present in the bone powder fraction than in the demineralized residual fraction. Indeed, the demineralization step both decalcifies the bone sample and removes contaminants, as keratins, which contribute to increasing the Mascot score in the bone powder fraction (Schroeter et al., 2016). The other identified proteins in both fractions with a low number of peptides are type II cytoskeletal 1 keratin, hornerin (less than 20 peptides for both proteins), which are common contaminants of archaeological excavations (Hendy et al., 2018). Others collagens are identified in both fractions for all samples as collagen II alpha 1 (COL2A1), collagen III alpha 1 (COL3A1), collagen V alpha 2 (COL5A2), and fibrillary collagen NC1 domain (Supplementary_data_2).

Taxa classification

All analyses were carried out blind, especially for taxa classification. The identification of taxa discriminant peptides of COL1A1 and COL1A2 was performed using BLASTp analysis against the nonredundant NCBI database limited to taxa identified by LC–MS/MS analyses. We used for (i) *Panthera* taxa, peptides identified from *Panthera tigris altaica, Panthera pardus, Panthera leo, Felis silvestris catus, Acinonyx jubatus, Puma concolor*, and *Lynx pardinus*; (ii) for *Ursus* taxa, peptides identified from *Ursus maritimus, Ursus arctos horribilis*, and *Ailuropoda melanoleuca*; (iii) for *Loxondota* taxa peptides from *Loxodonta africana; Mammut americanum* and (iv) for the *Bos* taxa peptides from *Bos primigenius taurus, Bos grunniens mutus, B. primigenius indicus, Bubalus Bubalis*, and *Bison bison*.

Supplementary Data 3, Table S3.2 shows the unique peptides identified in the studied samples. Their corresponding MS/MS spectra are shown in Supplementary Data 3 (Figure S3.1 – S3.19). All these unique peptides were identified in both fractions, that is, demineralization residual and bone powder. Most of the identified peptides carry post-translational modifications like methionine oxidation, proline hydroxylation, glutamine/asparagine deamidation, and carboxymethyl arginine. PEAKS X software identified other PTMs such as lysine methylation, lysine acetylation, and carboxylation of glutamic acid (MS/MS spectra are shown in Supplementary_data_3, Figure S3.20 – S3.22). These biological ageing-associated modifications of proteins were highlighted by previous studies on ancient bones (Cappellini et al., 2011; Cleland, Schroeter, & Schweitzer, 2015; Schroeter & Cleland, 2016; Wadsworth & Buckley, 2014).

Discriminant peptides matching the Ursidae family were identified in the samples Sc-5, Sc-6, and Sc-7. Considering sample Sc-8, the unique peptides match the Proboscidea order and two unique peptides (RGPNGEAGSAGPAGPPGLR and GPNGEAGSAGPAGPPGLR) show specificity to *Mammut americanum* species. The Proboscidea order taxa classification corresponds to the morphological identification realized by palaeontologists. Considering retouchers Sc-9 and Sc-10, which could not be identified by paleontologists on the basis of their appearance, discriminant peptides for the Bovinae subfamily as well as discriminant peptides

for Ursidae family were identified in samples Sc-9 and Sc-10 respectively. It can be noticed that two unique peptides from sample Sc-10 were also found in samples Sc-5, Sc-6, and Sc-7. Considering samples Sc-1, Sc-2, Sc-3, and Sc-4, discriminant peptides for the Felidae family were identified. Samples Sc-1, Sc-2, and Sc-3 were identified by paleontologists as *Panthera spelaea* bones and Sc-4 as *P. pardus* bone. Phylogenetic data show that their closest ancestors are *P. leo* (Barnett et al., 2016). Proteomics results may be refined allowing for the tracing of amino acid substitutions by introducing phylogenetic information. The species of closest ancestors *P. tigris altaica, P. pardus*, and *P. leo* were used to screen amino acid substitutions. In-house "All_Collagen" database gathers COL1A1 and COL1A2 sequences from all species present in National Center for Biotechnology Information (NCBI) database. The (NCBI) database contains four sets of sequences labeled *X1 X2, X3*, and *X4* of isoforms of COL1A1 from *P. pardus*. Sequence alignment shows that COL1A1 from *P. leo* is identical with COL1A1 from *P. pardus X1* and that COL1A1 from *P. tigris altaica, P. pardus X2, X3*, and *X4* are different (Figure S3.23 and Figure S3.24). The difference between COL1A1 from *P. leo and P. pardus X3* is that *C*-terminal part are absent on X3 isoform. COL1A2 from *P. tigris altaica, P. pardus*, and *P. leo* species shows no amino acid difference except for COL1A2 from *P. pardus* X4. So for samples Sc-1 to 3, collagen amino acids variations were studied from modern *P. leo* sequences and from modern *P. pardus* sequences X3 for sample Sc-4. Table S3.3 shows the closest species in the databases for the 10 studied bones and the phylogenetic data for each species. For retouchers, faunal data found on site, information from paleontologists, the results of proteomic analyses (number of peptides, sequence coverages), and phylogenetic data allowed targeting the closest species.

Asparagine and glutamine deamidation

The analysis of posttranslational modifications such as asparagine and glutamine deamidation or proline oxidation provides information on the state of preservation of bones. More ancient or more damaged samples tend to have a higher level of deamidation (Hill et al., 2015; Ramsøe et al., 2020; Schroeter & Cleland, 2016; Van Doorn et al., 2012; Welker et al., 2015). Formulas, sequences, mass, retention, and fragmentation spectra of both the native and modified peptides are presented in Supplementary Data 4 (Table S4.1, Figure S4.1 – S4.6).

For COL1A1, in the bone powder fraction the average percentage of deamidation is 76.4% for glutamine and 85.5% for asparagine in each of the samples. These percentages are slightly higher in the demineralized residual fraction, reaching 83.4% for glutamine and 91.4% for asparagine. The percentage of proline oxidation is stable at 99.1% for powder fractions and 99.4% for demineralized residual fractions (Figure S4.13). For COL1A2, in the bone powder fraction the average percentage of deamidation for all samples is 67.1% for glutamine and 80.7% for asparagine (Figure S4.14). These percentages are slightly higher in the demineralized residual fraction, reaching 83% for glutamine and 80.9% for asparagine. The percentage of hydroxyprolines is stable at 99.6% for powder fractions and 99.7% for demineralization residual fraction (Figure S4.14). All samples exhibit elevated deamidation frequencies in the range expected from Middle Pleistocene bones (Lanigan et al., 2020; Welker, Smith, et al., 2017). Pearson correlations show that the best correlation between bone powder fraction and the demineralization residual fraction is asparagine deamidation (r = 0.78 for COL1A1 and r = 0.62 for COL1A2), which proves that this latter is less impacted by the demineralization procedure.

Amino acid substitution

Nonreferenced amino acid substitutions in protein databases provide potential phylogenetic information. The excellent conservation state of proteins in bones from Scladina Cave

positively contributed to this study. Supplementary Data 5 (Figure S5.1 – S5.56) shows the MS/MS spectra of peptides with amino acid substitutions for all samples, and Table S5.1 indicates the peptides with amino acids substitutions. The analysis of amino acids variations was carried out with PEAKS X software and on the protein sequences of the closest species to the one identified by paleontologists (Table S3.3) and the protein sequences of extinct species identified as *Mammuthus sp* (Buckley et al., 2011; Welker et al., 2015) and *P. leo* (Buckley, Harvey, & Chamberlain, 2017; Welker et al., 2015).

Samples Sc-1, Sc-2, and Sc-3 show 12 substitutions on collagen sequences, referencing to the *P. leo* species. Samples Sc-1, Sc-2, and Sc-3 carry 10 identical amino acid variations in their protein sequences showing their belonging to the same species. These observations agree with the osteomorphological analyses concluding to their identification as *Panthera spelaea* (Table S5.1). Sample Sc-4 carries only four substitutions in the collagen I alpha 1 and alpha 2 sequences compared to the modern *P. pardus* sequences (Table S5.1).

Sample Sc-5 was identified as a bone from the *Ursus spelaeus* species by paleontologists. The closest ancestor sequence in the databases is *U. arctos horribilis*. Proteomic analysis shows one amino acid substitution on the sequence of COL1A1 and 11 amino acid substitutions on the sequence of COL1A2 (Table S5.1). The Sc-6 and Sc-7 samples were identified as *Ursus thibetanus* species bones by paleontologists. Its closest ancestor is again *U. arctos horribilis*. One amino acid substitution on the sequence of COL1A1 and seven amino acid substitutions on the sequence of COL1A2 were found in Sc-6 sample and four substitutions in Sc-7 COL1A2 sequence. There are no common substitutions for samples Sc-6 and Sc-7. Only two substitutions on the position 511 and 513 are common to both samples Sc-7 and Sc-5.

For sample Sc-8, Elephas and Mammuthus are considered to form a monophyletic clade (Shoshani & Tassy, 2005). The collagen sequence of *Mammuthus* sp. in Buckley and colleagues and Welker and colleagues are concatenated sequences of COL1A1 and COL1A2 (Buckley et al., 2011; Welker et al., 2015). From these sequences three substitutions are observed positioned on COL1A2 collagen sequence form *Mammuthus* sp. in Welker and colleagues, which are the best hit (Table S5.1).

Retoucher Sc-9 was identified as belonging to the *Bovidae* family. Considering the fauna at Scladina excavation site, the closest animals are *B. primigenius primigenius* (aurochs) and *Bison priscus* (bison). The collagen sequence of *Bos taurus* matched with a greater number of peptides and a slightly higher sequence coverage. Three amino acid substitutions were also identified on the sequence of COL1A1 and two substitutions on COL1A2 referring to *B. taurus* species. These results demonstrate that sample Sc-9 is a *B. primigenius taurus* (aurochs), agreeing with fauna at Scladina Cave.

Proteomics data indicate that retoucher Sc-10 belonged to a bone of an animal of the *Ursidae* family. The closest species in protein database referring to Scladina Cave fauna is *U. arctos horribilis*, as discussed above. Proteomics detected one substitution on the sequence of COL1A1, and seven substitutions on COL1A2. Seven of these substitutions are common to sample Sc-5. These results show a resemblance between samples Sc-10 and Sc-5, potentially concluding that sample Sc-10 may belong to the *Ursus spelaeus* species. The comparison of Sc-10 with either Sc-6 or 7 identified two common substitutions with Sc-6 (positions 833, 977) and one shared substitution with Sc-7 (position 511). These results also demonstrate that samples Sc-6 and Sc-7 do not belong to the same species as sample Sc-10.

DISCUSSION

Method

There is a strong impetus for developing methods that allow to obtain as much information as possible on the proteins present in bone samples from a minimum bone quantity (Schroeter

et al., 2022). For example, being able to identify the animal bone of the retouchers can help understanding the relationship between the human and the animal used. However, a crucial point for this kind of artifacts is that sampling should involve the smallest possible quantity in order to avoid damaging the object and allowing for subsequent analyses or displaying the artifact in a museum.

The first step in bone analysis is demineralization, which is carried out in either acidic or basic solutions. For the demineralization step we used aqueous trifluoroacetic acid (TFA) in our protocol rather than the commonly used hydrochloric acid (HCl). TFA has a lower pK_a than HCl. This means that TFA is a weaker acid and has a smaller impact on proteins. Furthermore, TFA is more hydrophobic, so it favors the solubility of the extracted proteins from their hydroxyapatite matrix. For digestion we use eFASP (Enhanced Filter Aided Sample Preparation), which outperforms in solution digestion or FASP (Erde et al., 2014). LC–MS/MS analyses of optimized-eFASP-digested paleontological bones afford a higher number of identified peptides, a higher percentage of coverage and the reliable identification of PTMs as very few spurious modifications are induced by eFASP digestion (Erde et al., 2014). Our method shows that both the demineralized fraction and the bone powder, allowed for the identification of type I collagen. By combining proteomic analysis, phylogenetic information, and modern species protein data bases, we unequivocally identified the species of bones from Scladina Cave in agreement with data provided by paleontologists. This methodology was applied for the specific identification of two bones used as retouchers (Sc-9 and Sc-10), which could not be identified precisely by their anatomy. Remarkably, on 5 mg of bone powder only, we were able to identifying proteins, genus, and PTMs from Pleistocene bones. Very recently Rüther and colleagues (Rüther et al., 2022) have developed a new method **"**Species by Proteome INvestigation**"** (SPIN), which allows high-throughput bone proteome analysis. In this paper the SPIN method is compared to different preparation methods including the FASP method but not the eFASP method, which is better than FASP. Our method works on 120,000 years old Pleistocene samples, whereas SPIN was used on specimens dating approximately between 30,000–60,000 BP.

Protein identification

Bone preservation and the presence of modern contaminants are two major difficulties in bone palaeoproteomics studies. In our study, keratins were poorly identified in comparison to endogenous bone proteins like collagens. Human honerin, and dermicin, very often described as contaminants in bones proteomic analysis, were detected in trace amounts (Sawafuji et al., 2017; Wadsworth & Buckley, 2014). The two chains of type I collagen, representing 90% of bone organic matter, were identified with up to 60% coverage on both chains. This result shows an exceptional bone preservation at Scladina Cave and the high sensitivity of eFASP protocol-based proteomic analysis (Erde et al., 2014) without protein extraction after demineralization from only 5 mg of starting material. It can be pointed out that starting material ranging from 30 mg to several hundred milligrams are commonly described in palaeoproteomics studies (Buckley et al., 2019; Buckley, Harvey, & Chamberlain, 2017; Cappellini et al., 2011; Welker, Smith, et al., 2017).

Other collagens have been identified as COL2A1, a fibrillar collagen found in cartilage, COL3A1, found as a major structural component in hollow organs, and COL5A2, a minor component of connective tissue (Shoulders & Raines, 2009). Welker and colleagues showed the presence of other types of collagen, as COL10A1 and COL27A1, in bone specimens from Grotte du Renne (Welker et al., 2016). Noncollagenous proteins (NCPs) were not detected despite being described in several studies on ancient bones (Cappellini et al., 2011; Sawafuji et al., 2017; Wadsworth & Buckley, 2014).

Post-translational modifications

In this study, the percentage of deamidation ranged between 71 and 95% on COL1A1 and 71–92% on COL1A2 peptides identified in all samples, which agrees with deamidation percentage described in literature (Welker, Smith, et al., 2017). In our study we see a slight difference between the percentage of deamidation between the demineralization fraction and bone powder. Many studies on ancient bones have studied deamidation rate. Glutamine and asparagine deamidation percentage of around 50% and 60% respectively were found for 40 and 50 ka bones from Kleine Feldhofer Grotte (Germany) (Lanigan et al., 2020). The analysis of deamidation of assemblage of 475 bones from Quinçay, France, dating from Châtelperronian (Upper Paleolithic, 44,500–36,000 BP) and Early Holocene showed an average deamidation rate of 60%. It was shown that no significant difference could be observed between layers (Welker, Soressi, et al., 2017). On Middle Pleistocene bone specimens of rhinoceros genus *Stephanorhinus* (337–300 ka), glutamine and asparagine deamidations reach nearly 100% (Welker, Smith, et al., 2017). Recent research show variable deamidation rate within a single bone assemblage over 45–50 ka until 284–305 ka (Brown et al., 2021). This study highlights the difficulty of using deamidation rate to relatively date samples. The difference in the rate of deamidation between publications for bones from the same period can be explained by the different protein extraction procedures for demineralization and/or once proteins induce variable deamidation (Hao et al., 2011; Ren et al., 2009; Simpson et al., 2016; Van Doorn et al., 2012; Wilson et al., 2012). The geological location and environmental conditions can also influence the percentage of deamidation (Schroeter & Cleland, 2016). The preservation of bone proteins may be deduced from both the percentage in weight of extracted collagen from bones and the rate of deamidation, which has been reported as a marker of degradation in ancient bones (Van Doorn et al., 2012).

Regarding hydroxyproline, the frequency percentage of hydroxylated prolines is close to 99% in the identified collagen sequence, having a 60% coverage. In modern samples, collagen hydroxyproline rate is around 20% (Zaitseva et al., 2015). This modification stabilizes the triple helix in collagen (Kotch et al., 2008). The presence of high frequency of hydroxyprolines could be explained by the age of the bones or the state of preservation. As with deamidation, environmental conditions can have an effect on this modification (Schroeter & Cleland, 2016).

Dihydroxyprolines were also identified in Scaldina Cave samples. This modification was reported in siliceous cell walls of diatoms (Nakajima & Volcani, 1969). These authors suggested that hydroxylated amino acids may play a role in the silicification of diatom cell walls. Bonjean and colleagues showed the presence of silicified materials in the area around Scladina Cave, which could explain the presence of dihydroxylation of prolines on the collagen sequences (Bonjean et al., 2015). It is also interesting to notice that hydroxylysines were identified in collagen sequences of Scladina Cave bones. Hydroxylysines are involved in crosslinks between the triple helix of collagen (Knott & Bailey, 1998). Hydroxylysines can be further modified by the sequential steps of *O*-linked glycosylation, producing G-Hyl (galactosylhydroxylysine) and GG-Hyl (glucosylgalactosylhydroxylysine) (Scott et al., 2012), playing multiple roles in normal mammalian physiology and pathology. Hill and colleagues, showed the presence of hydroxylysine glucosylgalactosylation in the extinct *Bison latifrons* (Hill et al., 2015).

Molecular sequence variation

Ancient proteins are commonly identified in reference to the closest (most often modern) species present in the database, as shown for Pleistocene samples of *Raphus cucullatus* from modern birds (Horn et al., 2019), *Stephanorhinus* from *rhinoceros* (Cappellini et al., 2019) or *kangaroos* from modern Australian vertebrate species (echidna, wombat, red kangaroo, etc.) (Buckley, Cosgrove, et al., 2017).

Samples Sc-1, Sc-2, and Sc-3 were identified as belonging to the Felidae family. Osteomorphological studies identified them as belonging to *Panthera spelaea* species. The *P. leo* sequences come from the PanLeo1.0 project present in NCBI (Armstrong et al., 2020). However, the alignment of the sequences of COL1A1 and COL1A2 does not show any difference compared to *P. pardus* isoform X1 sequences, which hampers the identification of the exact species. There are two other *P. leo* collagen sequences in the literature but protein alignments show that *P. leo* from Welker and colleagues has many missing portions in the sequence and the sequence from Buckley and colleagues are similar to *P. leo* sequences from NCBI except for three amino acids (Buckley et al., 2017; Welker et al., 2015). Nevertheless, the same amino acid substitutions on both COL1A1 and COL1A2 sequences were found for the three samples (numbered by reference to the *P. leo* modern sequence indicating they belong to the same species). A fourth sample, Sc-4, identified as belonging to the Felidae family showed different collagen sequence substitutions compared to Sc-1, Sc-2, and Sc-3 samples, ascribing it to another species. This Sc-4 bone was identified as *P. pardus* by paleontologists.

Samples Sc-5 to Sc-7 were identified as belonging to the Ursidae family by proteomics. Osteomorphological studies identified Sc-5 sample as belonging to the *Ursus spelaeus* and Sc-6 and Sc-7 samples to *U. thibetanus* species. The closest ancestor sequence in the databases is *U. arctos horribilis* for both species. Amino acid variations identified for the three samples did not provide conclusive information to be used in taxonomy-based clustering. For the two bones Sc-6 and Sc-7, there is no shared substitution, though osteomorphological studies conducted by paleontologists classified them as belonging to the same species. This lack of similarity may be explained either by variability between collagen protein sequences of the same species or by an inaccurate identification proposed by the osteomorphological analysis. Genome sequencing of the *U. thibetanus* species may enable precise identification of these samples.

Sample Sc-8 was ascribed to the *Mammut americanum* species using two discriminant peptides on the sequence of COL1A2, and other unique peptides correspond to the Proboscidea order. The identification of the Proboscidea order taxa classification is in full agreement with the osteomorphological identification. However, we found two peptides specific to the *Mammut americanum* species, in contradiction with osteomorphological and faunal studies. The *Mammut americanum* species and the *Mammuthus sp* species come from the same order, Proboscidae, despite the former belonging to the Mammutidae family and the latter being part of the Elephantidae family (Rohland et al., 2007). In protein databases (NCBI, Swissprot/Uniprot), only the sequence of the species *Mammut americanum* may be found. However, sequences for type I collagen of the species *Mammuthus sp* is described in previous papers (Buckley et al., 2011; Welker et al., 2015). These sequences allowed for the identification of type I collagen bearing three mutations. For the study of extinct species, it is extremely important to work with the right protein sequences. Mascot analysis makes it possible to make a definite genus identification. Moreover, this study also shows the importance of proteomics data with palaeontological studies and of working on protein sequences of the closest species.

The anthropogenically modified bones (Sc-9 and Sc-10) could not be identified by osteomorphologic studies. Retoucher Sc-9 was identified through proteomics as belonging to the *Bovidae* family with a number of identified peptides and higher sequence coverage to *B. primigenius primigenius* (aurochs). The type I collagen sequence of *B. primigenius* was reported by Welker and colleagues. (Welker et al., 2015), though this sequence shows multiple similarities with the one of *B. taurus* in Uniprot or NCBI databases. This shows the importance of determining the variations of amino acids for the validation of new collagen sequences of animal species. Sc-10 was ascribed to the genus *Ursus*. Whereas Sc-10 shows different substitutions than those observed for samples Sc-6 and Sc-7, similarities in amino acid substitutions to the Sc-5 sample allow its identification as *Ursus spelaeus*. Abrams and colleagues highlighted the presence of bear bone retouchers used by Neanderthals to reshape their lithic tools (Abrams et al., 2014). This new example reinforces the predominant role played by the exploitation of bear carcasses (brown and

cave bears) in Scladina Cave. The presence of anthropogenic modifications on urside remains remain a scarce event during the Middle Palaeolithic with only a few occurrences such as in Biache-saint-Vaast (Northern France) (Moigne et al., 2016). Therefore, this analytical protocol appears to be of great interest for Palaeolithic studies in regards of the limited quantity of bone employed.

Overall, the taxa identified by proteomics for the nine studied samples correlate with osteomorphological, palaeoenvironmental, and palaeodietary data (Bocherens et al., 1997, 1999), except for sample Sc-8. This may be due to the limited number of collagen sequence of the Proboscidea family available in the databases.

CONCLUSION

Our study demonstrated the possibility of identifying taxa of at least 130,000-year-old bones using an optimized proteomics method starting from 5-mg samples. A large number of peptides were identified along with high average sequence coverages up to of 60% for collagen I alpha 1. Deamidation frequency of about 80% and oxidation frequency of about 99% were measured. Despite the restricted databases for several species, several taxa were identified, including the Felidae family, the Elephantidae and Bovinae subfamily, and the *Ursus* genus. Proteomics allowed for unequivocally identifying the specific origin of Pleistocene-modified and strongly shrunken bones of bovid and ursid remains employed in tool manufacturing. It is a major contribution especially to zooarchaeological studies often struggling with species identification, as bone shape can be highly altered by anthropic modifications.

ACKNOWLEDGEMENTS

The authors thank the reviewers for their careful reading of the manuscript and for their valuable suggestions. The authors acknowledge the IBiSA network for the financial. The mass spectrometers were funded by the University of Lille, the CNRS, the Région Hauts-de-France, and the European Regional Development Fund. The authors thank the CNRS Mission pour l'Interdisciplinarité for the funding of the project ProtéoMamQuat (2017-2018) and the Région Hauts-de-France for the funding of the project ProtéOsHdF (2021-2022). The authors warmly thank the Scladina Cave excavation team for granting them the bone samples of this study and for sharing their paleontological data. The authors also gratefully acknowledge Caterina Bordin for her careful proofreading of the manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/arcm.12800.

DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data have been deposited on the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al., 2012) with the data set identifier PXD031386.

ORCID

Fabrice Bray https://orcid.org/0000-0002-4723-8206
Stéphanie Flament https://orcid.org/0000-0002-0135-675X
Grégory Abrams https://orcid.org/0000-0002-0616-4335
Dominique Bonjean https://orcid.org/0000-0002-0334-8307
Christian Rolando https://orcid.org/0000-0002-3266-8860
Caroline Tokarski https://orcid.org/0000-0001-7296-3330
Patrick Auguste https://orcid.org/0000-0002-8302-6307

REFERENCES

Abrams, G. (2015). Palaeolithic bone retouchers from Belgium: A preliminary overview of the recent research through historic and recently excavated bone collections: The origins of bone tool technologies." Retouching the Palaeolithic: Becoming Human and the Origins of Bone Tool Technology" Conference at Schloss Herrenhausen in Hannover, Germany.

Abrams, G., Bello, S. M., Di Modica, K., Pirson, S., & Bonjean, D. (2014). When Neanderthals used cave bear (Ursus spelaeus) remains: Bone retouchers from unit 5 of Scladina cave (Belgium). *Quaternary International*, *326-327*, 274–287. https://doi.org/10.1016/j.quaint.2013.10.022

Abrams, G., Bonjean, D., Di Modica, K., Pirson, S., Otte, M., & Patou-Mathis, M. (2010). Les os brûlés de l'ensemble sédimentaire 1A de Scladina (Andenne, Belgique): Apports naturels ou restes de foyer (s) néandertalien (s). *Notae Praehistoricae*, (30), 5–13.

Armstrong, E. E., Taylor, R. W., Miller, D. E., Kaelin, C. B., Barsh, G. S., Hadly, E. A., & Petrov, D. (2020). Long live the king: Chromosome-level assembly of the lion (Panthera leo) using linked-read, hi-C, and long-read data. *BMC Biology*, *18*(1), 1–14. https://doi.org/10.1186/s12915-019-0734-5

Barnett, R., Mendoza, M. L. Z., Soares, A. E. R., Ho, S. Y. W., Zazula, G., Yamaguchi, N., Shapiro, B., Kirillova, I. V., Larson, G., & Gilbert, M. T. P. (2016). Mitogenomics of the extinct cave lion, Panthera spelaea (Goldfuss, 1810), resolve its position within the Panthera cats. *Open Quaternary*, *2*, 1–11. https://doi.org/10.5334/oq.24

Birch, S. E. P., Scheu, A., Buckley, M., & Çakırlar, C. (2019). Combined osteomorphological, isotopic, aDNA, and ZooMS analyses of sheep and goat remains from Neolithic Ulucak, Turkey. *Archaeological and Anthropological Sciences*, *11*(5), 1669–1681. https://doi.org/10.1007/s12520-018-0624-8

Bocherens, H., Billiou, D., Mariotti, A., Patou-Mathis, M., Otte, M., Bonjean, D., & Toussaint, M. (1999). Palaeoenvironmental and palaeodietary implications of isotopic biogeochemistry of last interglacial Neanderthal and mammal bones in Scladina cave (Belgium). *Journal of Archaeological Science*, *26*(6), 599–607. https://doi.org/10.1006/jasc.1998.0377

Bocherens, H., Billiou, D., Patou-Mathis, M., Bonjean, D., Otte, M., & Mariotti, A. (1997). Paleobiological implications of the isotopic signatures (13C, 15N) of fossil mammal collagen in Scladina cave (Sclayn, Belgium). *Quaternary Research*, *48*(3), 370–380. https://doi.org/10.1006/qres.1997.1927

Bonjean, D., Abrams, G., Di Modica, K., Otte, M., Pirson, S., & Toussaint, M. (2014). Scladina cave: Archaeological context and history of the discoveries, *J the Scladina I-4A juvenile Neandertal. Palaeoanthropology, Context, Liege*, *143*, 31–48.

Bonjean, D., Vanbrabant, Y., Abrams, G., Pirson, S., Burlet, C., Di Modica, K., Otte, M., Auwera, J. V., Golitko, M., McMillan, R., & Goemaere, E. (2015). A new Cambrian black pigment used during the late middle Palaeolithic discovered at Scladina cave (Andenne, Belgium). *Journal of Archaeological Science*, *55*, 253–265. https://doi.org/10.1016/j.jas.2014.11.040

Brandt, L. Ø., Haase, K., & Collins, M. J. (2018). Species identification using ZooMS, with reference to the exploitation of animal resources in the medieval town of Odense. *Danish Journal of Archaeology*, *7*(2), 139–153. https://doi.org/10.1080/21662282.2018.1468154

Brown, S., Higham, T., Slon, V., Pääbo, S., Meyer, M., Douka, K., Brock, F., Comeskey, D., Procopio, N., Shunkov, M., Derevianko, A., & Buckley, M. (2016). Identification of a new hominin bone from Denisova cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. *Scientific Reports*, *6*(1), 1–8. https://doi.org/10.1038/srep23559

Brown, S., Kozlikin, M., Shunkov, M., Derevianko, A., Higham, T., Douka, K., & Richter, K. K. (2021). Examining collagen preservation through glutamine deamidation at Denisova cave. *Journal of Archaeological Science*, *133*, 105454. https://doi.org/10.1016/j.jas.2021.105454

Buckley, M. (2015). Ancient collagen reveals evolutionary history of the endemic south American 'ungulates'. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1806), 20142671. https://doi.org/10.1098/rspb.2014.2671

Buckley, M., Collins, M., Thomas-Oates, J., & Wilson, J. C. W. (2009). Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, *23*(23), 3843–3854. https://doi.org/10.1002/rcm.4316

Buckley, M., Cosgrove, R., Garvey, J., & Prideaux, G. J. (2017). Identifying remains of extinct kangaroos in Late Pleistocene deposits using collagen fingerprinting. *Journal of Quaternary Science*, *32*(5), 653–660. https://doi.org/10.1002/jqs.2964

Buckley, M., Harvey, V. L., & Chamberlain, A. T. (2017). Species identification and decay assessment of Late Pleistocene fragmentary vertebrate remains from Pin Hole Cave (Creswell Crags, UK) using collagen fingerprinting. *Boreas*, *46*(3), 402–411.

Buckley, M., Kansa, S. W., Howard, S., Campbell, S., Thomas-Oates, J., & Collins, M. (2010). Distinguishing between archaeological sheep and goat bones using a single collagen peptide. *Journal of Archaeological Science*, *37*(1), 13–20.

Buckley, M., Larkin, N., & Collins, M. (2011). Mammoth and mastodon collagen sequences; survival and utility. *Geochimica et Cosmochimica Acta*, *75*(7), 2007–2016. https://doi.org/10.1016/j.gca.2011.01.022

Buckley, M., Lawless, C., & 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. https://doi.org/10.1016/j.jprot.2018.11.014

Buckley, M., Walker, A., Ho, S. Y., Yang, Y., Smith, C., Ashton, P., Oates, J. T., Cappellini, E., Koon, H., & Penkman, K. (2008). Weighing the mass spectrometric evidence for authentic Tyrannosaurus rex collagen. *Science*, *319*(5859), 33–38. https://doi.org/10.1126/science.1147046

Buckley, M., Warwood, S., van Dongen, B., Kitchener, A. C., & Manning, P. L. (2017). A fossil protein chimera; difficulties in discriminating dinosaur peptide sequences from modern cross-contamination. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1855), 20170544. https://doi.org/10.1098/rspb.2017.0544

Cappellini, E., Gentry, A., Palkopoulou, E., Ishida, Y., Cram, D., Roos, A.-M., Watson, M., Johansson, U. S., Fernholm, B., & Agnelli, P. (2014). Resolution of the type material of the Asian elephant, Elephas maximus Linnaeus, 1758 (Proboscidea, Elephantidae). *Zoological Journal of the Linnean Society*, *170*(1), 222–232. https://doi.org/10.1111/zoj12084

Cappellini, E., Jensen, L. J., Szklarczyk, D., Ginolhac, A., da Fonseca, R. A., Stafford, T. W. Jr., Holen, S. R., Collins, M. J., Orlando, L., & Willerslev, E. (2011). Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *Journal of Proteome Research*, *11*(2), 917–926. https://doi.org/10.1021/pr200721u

Cappellini, E., Prohaska, A., Racimo, F., Welker, F., Pedersen, M. W., Allentoft, M. E., de Barros Damgaard, P., Gutenbrunner, P., Dunne, J., & Hammann, S. (2018). Ancient biomolecules and evolutionary inference. *Annual Review of Biochemistry*, *87*, 1029–1060.

Cappellini, E., Welker, F., Pandolfi, L., Ramos-Madrigal, J., Samodova, D., Rüther, P. L., Fotakis, A. K., Lyon, D., Moreno-Mayar, J. V., Bukhsianidze, M., Jersie-Christensen, R. R., Mackie, M., Ginolhac, A., Ferring, R., Tappen, M., Palkopoulou, E., Dickinso, M. R., Stafford, T. W. Jr., Chan, Y. L., … Willerslev, E. (2019). Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny. *Nature*, *574*(7776), 103–107. https://doi.org/10.1038/s41586-019-1555-y

Cleland, T. P. (2018a). Human bone Paleoproteomics utilizing the single-pot, solid-phase-enhanced sample preparation method to maximize detected proteins and reduce Humics. *Journal of Proteome Research*, *17*(11), 3976–3983. https://doi.org/10.1021/acs.jproteome.8b00637

Cleland, T. P. (2018b). Solid digestion of demineralized bone as a method to access potentially insoluble proteins and post-translational modifications. *Journal of Proteome Research*, *17*(1), 536–542.

Cleland, T. P., & Schroeter, E. R. (2018). A comparison of common mass spectrometry approaches for paleoproteomics. *Journal of Proteome Research*, *17*(3), 936–945. https://doi.org/10.1021/acs.jproteome.7b00703

Cleland, T. P., Schroeter, E. R., Feranec, R. S., & Vashishth, D. (2016). Peptide sequences from the first Castoroides ohioensis skull and the utility of old museum collections for palaeoproteomics. *Proceedings Biological Sciences*, *283*(1832), 20160593. https://doi.org/10.1098/rspb.2016.0593

Cleland, T. P., Schroeter, E. R., & Schweitzer, M. H. (2015). Biologically and diagenetically derived peptide modifications in moa collagens. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1808), 20150015.

Cleland, T. P., Schroeter, E. R., Zamdborg, L., Zheng, W., Lee, J. E., Tran, J. C., Bern, M., Duncan, M. B., Lebleu, V. S., & Ahlf, D. R. (2015). Mass spectrometry and antibody-based characterization of blood vessels from Brachylophosaurus canadensis. *Journal of Proteome Research*, *14*(12), 5252–5262.

Colleary, C., Little, N. C., & Cleland, T. P. (2020). Microwave-assisted acid hydrolysis for whole-bone proteomics and paleoproteomics. *Rapid Communications in Mass Spectrometry*, *34*(2), e8568.

Dallongeville, S., Garnier, N., Rolando, C., & Tokarski, C. (2016). Proteins in art, archaeology, and paleontology: From detection to identification. *Chemical Reviews*, *116*(1), 2–79.

Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C. L., Woolley, J., Crisp, M. K., Wilson, J., Fotakis, A., Fischer, R., & Kessler, B. M. (2016). Protein sequences bound to mineral surfaces persist into deep time. *eLife*, *5*, e17092. https://doi.org/10.7554/eLife.17092

Desmond, A., Barton, N., Bouzouggar, A., Douka, K., Fernandez, P., Humphrey, L., Morales, J., Turner, E., & Buckley, M. (2018). ZooMS identification of bone tools from the north African later stone age. *Journal of Archaeological Science*, *98*, 149–157.

Erde, J., Ogorzalek Loo, R. R., & Loo, J. A. (2014). Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. *Journal of Proteome Research*, *13*(4), 1885–1895.

Folch, J., Lees, M., & Stanley, G. S. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry*, *226*(1), 497–509.

Hao, P., Ren, Y., Alpert, A. J., & Sze, S. K. (2011). Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Molecular & Cellular Proteomics*, *10*(10), O111.009381. https://doi.org/10.1074/mcp.O111.009381

Harvey, V. L., Daugnora, L., & Buckley, M. (2018). Species identification of ancient Lithuanian fish remains using collagen fingerprinting. *Journal of Archaeological Science*, *98*, 102–111. https://doi.org/10.1016/j.jas.2018.07.006

Harvey, V. L., LeFebvre, M. J., deFrance, S. D., Toftgaard, C., Drosou, K., Kitchener, A. C., & Buckley, M. (2019). Preserved collagen reveals species identity in archaeological marine turtle bones from Caribbean and Florida sites. *Royal Society Open Science*, *6*(10), 191137.

Helle, S., Bray, F., Putaux, J.-L., Verbeke, J., Flament, S., Rolando, C., D'hulst, C., & Szydlowski, N. (2019). Intra-sample heterogeneity of potato starch reveals fluctuation of starch-binding proteins according to granule morphology. *Plants*, *8*(9), 324–341.

Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C., & Collins, M. J. (2018). A guide to ancient protein studies. *Nature Ecology & Evolution*, *2*(5), 791–799.

Hill, R. C., Wither, M. J., Nemkov, T., Barrett, A., D'Alessandro, A., Dzieciatkowska, M., & Hansen, K. C. (2015). Preserved proteins from extinct Bison latifrons identified by tandem mass spectrometry; Hydroxylysine glycosides are a common feature of ancient collagen. *Molecular & Cellular Proteomics*, *14*(7), 1946–1958.

Hofman, C. A., Rick, T. C., Erlandson, J. M., Reeder-Myers, L., Welch, A. J., & Buckley, M. (2018). Collagen fingerprinting and the earliest marine mammal hunting in North America. *Scientific Reports*, *8*(1), 10014–10019.

Horn, I. R., Kenens, Y., Palmblad, N. M., Van der Plas-Duivesteijn, S. J., Langeveld, B. W., Meijer, H. J. M., Dalebout, H., Marissen, R. J., Fischer, A., Florens, V. F. B., Niemann, J., Rijsdijk, K. F., Schulp, A. S., Laros, J. F. J., & Gravendeel, B. (2019). Palaeoproteomics of bird bones for taxonomic classification. *Zoological Journal of the Linnean Society*, *186*(3), 650–665.

Knott, L., & Bailey, A. J. (1998). Collagen cross-links in mineralizing tissues: A review of their chemistry, function, and clinical relevance. *Bone*, *22*(3), 181–187.

Kostyukevich, Y., Bugrova, A., Chagovets, V., Brzhozovskiy, A., Indeykina, M., Vanyushkina, A., Zherebker, A., Mitina, A., Kononikhin, A., Popov, I., Khaitovich, P., & Nikolaev, E. (2018). Proteomic and lipidomic analysis of mammoth bone by high-resolution tandem mass spectrometry coupled with liquid chromatography. *European Journal of Mass Spectrometry*, *24*(6), 411–419.

Kotch, F. W., Guzei, I. A., & Raines, R. T. (2008). Stabilization of the collagen triple helix by O-methylation of hydroxyproline residues. *Journal of the American Chemical Society*, *130*(10), 2952–2953. https://doi.org/10.1021/ja800225k

Lanigan, L. T., Mackie, M., Feine, S., Hublin, J.-J., Schmitz, R. W., Wilcke, A., Collins, M. J., Cappellini, E., Olsen, J. V., Taurozzi, A. J., & Welker, F. (2020). Multi-protease analysis of Pleistocene bone proteomes. *Journal of Proteomics*, *228*, 103889.

Martisius, N. L., Welker, F., Dogandžič, T., Grote, M. N., Rendu, W., Sinet-Mathiot, V., Wilcke, A., Mcpherron, S. J., Soressi, M., & Steele, T. E. (2020). Non-destructive ZooMS identification reveals strategic bone tool raw material selection by Neandertals. *Scientific Reports*, *10*(1), 1–12.

McGrath, K., Rowsell, K., St-Pierre, C. G., Tedder, A., Foody, G., Roberts, C., Speller, C., & Collins, M. J. (2019). Identifying archaeological bone via non-destructive ZooMS and the materiality of symbolic expression: Examples from Iroquoian bone points. *Scientific Reports*, *9*(1), 11027.

Moigne, A.-M., Valensi, P., Auguste, P., García-Solano, J., Tuffreau, A., Lamotte, A., Barroso, C., & Moncel, M.-H. (2016). Bone retouchers from lower Palaeolithic sites: Terra Amata, Orgnac 3, Cagny-l'Epinette and Cueva del Angel. *Quaternary International*, *409*, 195–212.

Nakajima, T., & Volcani, B. E. (1969). 3,4-dihydroxyproline: A new amino acid in diatom cell walls. *Science*, *164*(3886), 1400–1401. https://doi.org/10.1126/science.164.3886.1400

Pirson, S., Court-Picon, M., Haesaerts, P., Bonjean, D., & Damblon, F. (2008). New data on geology, anthracology and palynology from the scladina cave Pleistocene sequence preliminary results. *Memoirs of the Geological Survey of Belgium*, *55*, 71–93.

Pirson, S., Damblon, F., Balescu, S., Bonjean, D., & Haesaerts, P. (2014). The palaeoenvironmental context and chronostratigraphic framework of the Scladina Cave sedimentary sequence (units 5 to 3-SUP): The Scladina I-4A Juvenile Neandertal (Andenne, Belgium)., v. 4, Etudes et Recherches archéologiques de l'Université de Liège, 134, 23 p.

Presslee, S., Slater, G. J., Pujos, F., Forasiepi, A. M., Fischer, R., Molloy, K., Mackie, M., Olsen, J. V., Kramarz, A., & Taglioretti, M. (2019). Palaeoproteomics resolves sloth relationships. *Nature Ecology & Evolution*, *3*(7), 1121–1130.

Ramsøe, A., van Heekeren, V., Ponce, P., Fischer, R., Barnes, I., Speller, C., & Collins, M. J. (2020). DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *Journal of Archaeological Science*, *115*, 105080. https://doi.org/10.1016/j.jas.2020.105080

Ren, D., Pipes, G. D., Liu, D., Shih, L.-Y., Nichols, A. C., Treuheit, M. J., Brems, D. N., & Bondarenko, P. V. (2009). An improved trypsin digestion method minimizes digestion-induced modifications on proteins. *Analytical Biochemistry*, *392*(1), 12–21.

Richter, K. K., Wilson, J., Jones, A. K., Buckley, M., van Doorn, N., & Collins, M. J. (2011). Fish'n chips: ZooMS peptide mass fingerprinting in a 96 well plate format to identify fish bone fragments. *Journal of Archaeological Science*, *38*(7), 1502–1510.

Rohland, N., Malaspinas, A.-S., Pollack, J. L., Slatkin, M., Matheus, P., & Hofreiter, M. (2007). Proboscidean mitogenomics: Chronology and mode of elephant evolution using mastodon as outgroup. *PLoS Biology*, *5*(8), 1663–1671. https://doi.org/10.1371/journal.pbio.0050207

Rüther, P. L., Husic, I. M., Bangsgaard, P., Gregersen, K. M., Pantmann, P., Carvalho, M., Godinho, R. M., Friedl, L., Cascalheira, J., & Taurozzi, A. J. (2022). SPIN enables high throughput species identification of archaeological bone by proteomics. *Nature Communications*, *13*(1), 1–14. https://doi.org/10.1038/s41467-022-30097-x

Sawafuji, R., Cappellini, E., Nagaoka, T., Fotakis, A. K., Jersie-Christensen, R. R., Olsen, J. V., Hirata, K., & Ueda, S. (2017). Proteomic profiling of archaeological human bone. *Royal Society Open Science*, *4*(6), 161004.

Schroeter, E. R., & Cleland, T. P. (2016). Glutamine deamidation: An indicator of antiquity, or preservational quality? *Rapid Communications in Mass Spectrometry*, *30*(2), 251–255.

Schroeter, E. R., Cleland, T. P., & Schweitzer, M. H. (2022). Deep time Paleoproteomics: Looking forward. *Journal of Proteome Research*, *21*(1), 9–19.

Schroeter, E. R., DeHart, C. J., Cleland, T. P., Zheng, W., Thomas, P. M., Kelleher, N. L., Bern, M., & Schweitzer, M. H. (2017). Expansion for the Brachylophosaurus canadensis collagen I sequence and additional evidence of the preservation of cretaceous protein. *Journal of Proteome Research*, *16*(2), 920–932.

Schroeter, E. R., DeHart, C. J., Schweitzer, M. H., Thomas, P. M., & Kelleher, N. L. (2016). Bone protein "extractomics": Comparing the efficiency of bone protein extractions of Gallus gallus in tandem mass spectrometry, with an eye towards paleoproteomics. *PeerJ*, *4*, 2603. https://doi.org/10.7717/peerj.2603

Schweitzer, M. H., Zheng, W., Organ, C. L., Avci, R., Suo, Z., Freimark, L. M., Lebleu, V. S., Duncan, M. B., Heiden, M. G. V., & Neveu, J. M. (2009). Biomolecular characterization and protein sequences of the Campanian hadrosaur B. canadensis. *Science*, *324*(5927), 626–631. https://doi.org/10.1126/science.1165069

Scott, I., Yamauchi, M., & Sricholpech, M. (2012). Lysine post-translational modifications of collagen. *Essays in Biochemistry*, *52*, 113–133.

Shoshani, J., & Tassy, P. (2005). Advances in proboscidean taxonomy & classification, anatomy & physiology, and ecology & behavior. *Quaternary International*, *126-128*, 5–20.

Shoulders, M. D., & Raines, R. T. (2009). Collagen structure and stability. *Annual Review of Biochemistry*, *78*, 929–958.

Simpson, J., Penkman, K. E. H., Demarchi, B., Koon, H., Collins, M. J., Thomas-Oates, J. E., Shapiro, B., Stark, M., & Wilson, J. (2016). The effects of demineralisation and sampling point variability on the measurement of glutamine deamidation in type I collagen extracted from bone. *Journal of Archaeological Science*, *69*, 29–38.

Van Doorn, N. L., Wilson, J., Hollund, H., Soressi, M., & Collins, M. J. (2012). Site-specific deamidation of glutamine: A new marker of bone collagen deterioration. *Rapid Communications in Mass Spectrometry*, *26*(19), 2319–2327.

Vinciguerra, R., De Chiaro, A., Pucci, P., Marino, G., & Birolo, L. (2016). Proteomic strategies for cultural heritage: From bones to paintings. *Microchemical Journal*, *126*, 341–348. https://doi.org/10.1016/j.microc.2015.12.024

Wadsworth, C., & Buckley, M. (2014). Proteome degradation in fossils: Investigating the longevity of protein survival in ancient bone. *Rapid Communications in Mass Spectrometry*, *28*(6), 605–615. https://doi.org/10.1002/rcm.6821

Welker, F. (2018). Palaeoproteomics for human evolution studies. *Quaternary Science Reviews*, *190*, 137–147.

Welker, F., Collins, M. J., Thomas, J. A., Wadsley, M., Brace, S., Cappellini, E., Turvey, S. T., Reguero, M., Gelfo, J. N., & Kramarz, A. (2015). Ancient proteins resolve the evolutionary history of Darwin's south American ungulates. *Nature*, *522*(7554), 81–84.

Welker, F., Hajdinjak, M., Talamo, S., Jaouen, K., Dannemann, M., David, F., Julien, M., Meyer, M., Kelso, J., & Barnes, I. (2016). Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences*, *113*(40), 11162–11167.

Welker, F., Ramos-Madrigal, J., Kuhlwilm, M., Liao, W., Gutenbrunner, P., de Manuel, M., Samodova, D., Mackie, M., Allentoft, M. E., & Bacon, A.-M. (2019). Enamel proteome shows that Gigantopithecus was an early diverging pongine. *Nature*, *576*(7786), 262–265.

Welker, F., Smith, G. M., Hutson, J. M., Kindler, L., Garcia-Moreno, A., Villaluenga, A., Turner, E., & 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*, 3033. https://doi.org/10.7717/peerj.3033

Welker, F., Soressi, M. A., Roussel, M., van Riemsdijk, I., Hublin, J.-J., & Collins, M. J. (2017). Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen. *STAR: Science & Technology of Archaeological Research*, *3*(1), 15–27.

Wilson, J., van Doorn, N. L., & Collins, M. J. (2012). Assessing the extent of bone degradation using glutamine deamidation in collagen. *Analytical Chemistry*, *84*(21), 9041–9048.

Zaitseva, O. V., Shandrenko, S. G., & Veliky, M. M. (2015). Biochemical markers of bone collagen type I metabolism. *Ukrainian Biochemical Journal*, *87*(1), 21–32.