Tryps-IN: A streamlined palaeoproteomics workflow enables ZooMS analysis of 10,000-year-old petrous bones from Jordan rift-valley

Theis Zetner Trolle Jensen a, b, Lisa Yeomans c, d, Louise Le Meillour a, Pia Wistoft Nielsen a, Max Ramsoe a, Meaghan Mackie a, b, d, Pernille Bangsgaard b, Moritz Kinzel c, Ingolf Thuesen e, Matthew J. Collins a, f, Alberto J. Taurozzi a, *.

a Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
b Department of Archaeology, BioArCh, University of York, Environment Building, Wentworth Way, York YO10 5NG, United Kingdom.
c Department of Cross-Cultural and Regional Studies, University of Copenhagen, Karen Blixens Plads 8, 2300 Copenhagen, Denmark.
d Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Blegdamsvej 3b, 2200 Copenhagen N, Denmark.
e German Archaeological Institute - Istanbul, Inonu Cad. 10, 34437 Istanbul, Turkey.
f McDonald Institute for Archaeological Research, University of Cambridge, West Tower, Downing St., CB2 3ER Cambridge, United Kingdom.

ARTICLE INFO

Keywords:
Zoogeography, Zooarchaeology, Palaeoproteomics, MPPNB, LC-MS/MS, ZooMS, Shkârat Msaied.

ABSTRACT

Poor preservation of collagen in dry and/or arid environments has hindered the application of Zooarchaeology by mass spectrometry (ZooMS) analysis in many regions of the world. As a result, many zooarchaeological investigations have relied exclusively on morphological assessment of fragmentary remains due to the inadequate preservation of biomolecules. The climatic conditions of Southwest Asia include extreme temperature fluctuations unconducive to preservation of proteins and DNA. We performed zooarchaeological analysis of remains from the 10,000-year-old site of Shkârat Msaied in Jordan and sub-sampled twenty-eight petrous bones, the hardest bone in the mammalian skeleton, for species identification by ZooMS. Using an unconventional and simplified extraction protocol we call Tryps-IN, in which digestion was performed without removal of the demineralising EDTA, we taxonomically identified several fragments, outperforming an established ZooMS workflow. A subset of identifications was subsequently confirmed using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) protein sequencing. The new methodology presented here opens the possibility of further bioarchaeological investigation of other fragmentary faunal assemblages within this region of archaeological significance.

1. Introduction

Some of the most studied archaeological questions pertain to the relationships our ancestors maintained with their environment. Studying past human-animal interactions is especially crucial for understanding the subsistence strategies that populations developed. It is also important when concerning the advent of domestication, one of the most important events in the (pre)history of human populations. Southwest Asia is important to our understanding of the shift from hunting and foraging to agriculture, as the region has yielded the earliest evidence for both plant and animal domestication (Bellwood, 2005; Martin and Edwards, 2013), early exploitation of secondary products (Marciniak, 2011) and the rise of urban economies (Zeder, 1988; Gaastra et al., 2020). The understanding of the changing nature of human-animal interactions could be improved if the identification of fragmentary bones was expanded beyond traditional zooarchaeological techniques (Zeder and Lapham, 2010). However, the morphological identification of faunal remains, which has been the backbone of zooarchaeological research for well over a hundred years, has encountered inherent limitations for identifying fragmentary bones which do not retain morphological criteria allowing taxonomic assignment (Zeder and Lapham, 2010). A comprehensive interpretation of past human-animal interactions, particularly when traditional approaches are insufficient, can be achieved through a combined methodology that leverages the strengths of both zooarchaeological (including osteometric, ageing and taphonomic data) and molecular (species identification) methods, as together they provide a more holistic and robust interpretation.

Here we present the zooarchaeological analyses of faunal remains...
from Shkārat Msaied, along with the development of a novel method for extracting preserved proteins from bones recovered in arid environments. To avoid losing the few molecules present, we devised a ‘one-pot’ extraction and digestion method where we demineralise bone in 0.5 M Ethylenediaminetetraacetic acid (EDTA) and digest proteins using trypsin, in the same tube without the removal of EDTA. We term this method of trypsin digestion in the EDTA ‘Tryps-IN’. To verify the compatibility of the trypsin enzyme with 0.5 M EDTA, we performed a trypsin activity assay, SDS-PAGE of undigested and trypsinised bone extracts, and LC-MS/MS analysis of cleavage efficiency using HeLa cell extract. We show that the activity of trypsin in 0.5 M EDTA was comparable to its activity in 50 mM Ammonium Bicarbonate (ABC). The Tryps-IN method was compared to a standard ABC extraction of the same Shkārat Msaied bone samples. We show that trypsin successfully cleaves proteins in 0.5 M EDTA, and that Tryps-IN yielded significantly more species-specific peptides than the classic ABC method. To verify our identifications, we performed peptide sequencing by liquid chromatography tandem mass spectrometry (LC-MS/MS) on eight extracts. Our results demonstrate that this method improves proteome species identification of samples from a region that is notorious for poor protein preservation, opening the door to future proteome studies of this and similar regions which until now have been out of reach.

2. Background

The advent of biomolecular archaeology methods in the past few decades has provided new insights into many archaeological questions. The well-established method of ancient DNA (aDNA) analysis is regrettably of very limited use in the semi-arid and fluctuating environments of the Levant due to the extensive degradation of endogenous DNA. An alternative to aDNA is palaeoproteomics, the study of the longer persisting biomolecules, namely proteins, which appears to be the most suitable strategy to build upon. It relies on the extraction of persisting proteins/peptides and their subsequent characterisation using mass spectrometry based techniques. So far, two different palaeoproteomics methods are commonly used for the recovery of proteins from archaeological faunal remains: Zooarchaeology by mass spectrometry (ZooMS) (Buckley et al., 2009; Buckley, 2018) and shotgun-based approaches (Hendy, 2021). Recent reviews of the discipline focusing on the fundamental principles, strengths, and limitations of the techniques were recently published and illustrate the range of research questions that ZooMS has been successfully employed to address (Hendy et al., 2018; Hendy, 2021; Warinner et al., 2022). However, most ZooMS studies have focused on temperate and/or more recent contexts, and questions that ZooMS has been successfully employed to address (Hendy et al., 2018; Hendy, 2021; Warinner et al., 2022). However, most ZooMS studies have focused on temperate and/or more recent contexts, and questions that ZooMS has been successfully employed to address (Hendy et al., 2018; Hendy, 2021; Warinner et al., 2022).

Our results demonstrate that this method improves proteome species identification of samples from a region that is notorious for poor protein preservation, opening the door to future proteome studies of this and similar regions which until now have been out of reach.

2. Background

The advent of biomolecular archaeology methods in the past few decades has provided new insights into many archaeological questions. The well-established method of ancient DNA (aDNA) analysis is regrettably of very limited use in the semi-arid and fluctuating environments of the Levant due to the extensive degradation of endogenous DNA. An alternative to aDNA is palaeoproteomics, the study of the longer persisting biomolecules, namely proteins, which appears to be the most suitable strategy to build upon. It relies on the extraction of persisting proteins/peptides and their subsequent characterisation using mass spectrometry based techniques. So far, two different palaeoproteomics methods are commonly used for the recovery of proteins from archaeological faunal remains: Zooarchaeology by mass spectrometry (ZooMS) (Buckley et al., 2009; Buckley, 2018) and shotgun-based approaches (Hendy, 2021). Recent reviews of the discipline focusing on the fundamental principles, strengths, and limitations of the techniques were recently published and illustrate the range of research questions that ZooMS has been successfully employed to address (Hendy et al., 2018; Hendy, 2021; Warinner et al., 2022). However, most ZooMS studies have focused on temperate and/or more recent contexts, and only a few papers so far have focused on archaeological material from either arid or tropical environments. To date, ZooMS analyses in arid environments have been restricted either to a few countries in Southwest Asia (Buckley and Kansa, 2011; Price et al., 2013; Pillaar Birch et al., 2019); and Africa (Coutu et al., 2021, 2016; Culley et al., 2021a, 2021b; Desmond et al., 2018; Prendergast et al., 2019, 2017), but no study has been conducted in Jordan. Even radiocarbon dates directly derived from oseous material are limited in these areas, instead, dating has relied almost entirely on indirect dates from charred botanical remains. Given the preservational limitations of collagen in most archaeological contexts of the region (Weiner and Bar-Yosef, 1990), there is significant potential for further research to determine if methodological improvements can be made that enable extraction and analysis of the limited and damaged persisting proteins. Making strides in this area is likely to increase our understanding of the early history of animal domestication and the post-depositional molecular degradation occurring in arid environments.

Here, we focused on the use of ZooMS as the cost per sample and relative ease of data analysis has become the dominant proteomic approach for large-scale studies. ZooMS provides additional information for the analysis of faunal assemblages by yielding peptide mass fingerprints of type I collagen (COL1) and in temperate regions has proven to give significant results that complement morphological identification. We chose to apply this methodology to a case study where characterising the preserved biomolecules could resolve the controversy that has surrounded the geographical distribution of the wild ancestors of domesticated sheep in the Early Holocene. Remains of wild sheep have been morphologically identified at several Late Epipalaeolithic/Late Pleistocene Southern Levantine sites - Wadi Mataha 2, Rosh Horesha, Abu Salem, Ramat Harif, Wadi Hammeh 27, Wadi Judayid, Shubayqa 1 and Hatoula (Butler et al., 1977; Davis, 1985; Davis et al., 1982; Henry et al., 1985; Martin and Edwards, 2013; Stock et al., 2005; Yeomans et al., 2017). These sites date between approximately 15,000 cal BP until the end of Younger Dryas at 11,700 cal BP. Evidence for the survival of a wild sheep population in the southern Levant into the Holocene is debated, as only three Pre-Pottery Neolithic A (PPNA) sites, Hatoula, Shubayqa 6 and Jericho (Clutton-Brock and Uerpmann, 1974; Davis, 1985; Yeomans et al., 2018), have faunal assemblages that include bones of wild sheep. These Holocene sites post-date the Younger Dryas but pre-date evidence of caprine domestication in the area. However, difficulties in species determination using morphological criteria on fragmentary remains has left no consensus as to the presence of wild sheep in the PPNA. It is generally argued that later Neolithic sheep were introduced as domesticated animals from the Northern Levant (Munro et al., 2018) and became well-represented in domesticated herds by the Late Pre-Pottery Neolithic B (LPPNB). Few sites dating to the intervening period between the PPNA and the LPPNB have been excavated in Southern Jordan and also yielded large, published faunal assemblages containing bones of sheep. Shkārat Msaied and Ayn Abu Nukhayla, dating to the Middle Pre-Pottery Neolithic B (MPPNB) (Henry et al., 2003), are notable exceptions. However, since the identification of sheep at Ayn Abu Nukhayla and Shkārat Msaied relies on morphological criteria, results might be disputed. Implications for the identification of sheep in the MPPNB are significant with three interpretations: 1) a small population of wild sheep survived in southern Jordan into the Early Holocene; 2) managed sheep were introduced into the region as early as the start of the MPPNB or 3) Trade with sites further north, such as ‘Ain Ghazal, included traded parts of animals or occasional live animals. An argument has recently been made that goats were being managed by the Early Pre-Pottery Neolithic B (EPPNB) in the southern Levant in a gradual process of incipient management of native wild goats with additional stock brought in from the north (Horwitz, 1993; Horwitz and Lernau, 2003; Munro et al., 2018; Sapir-Hen et al., 2016). No suggestion has been made for the management of sheep at such an early date. Sites with sheep dating to the MPPNB are ‘Ain Ghazal and Tel Ro’im West are located to the north of Ayn Abu Nukhayla and Shkārat Msaied (Agha et al., 2019; Von den Driesch and Wodtke, 1997; Wasse, 2002). These sheep are interpreted as imported founding stock. Shkārat Msaied, therefore, is an important site for the investigation of the shift from hunting to agriculture as it has provided more than 100 bones morphologically identified as sheep. Use of ZooMS to corroborate these results is highly relevant given the controversy.

3. Materials and methods

Site and sample description: Shkārat Msaied. The site is situated on the eastern slope of the Jordanian rift valley in present-day Jordan (Fig. 1A) c. 1000 m above sea-level (Fig. 1B), approximately 16 km north of Petra/Wadi Musa. Since 1999, thirteen excavation seasons at Shkārat Msaied have been conducted by the University of Copenhagen and an area of around 600 m² has been excavated in total (Hermansen et al., 2006; Jensen et al., 2005; Kinzel, 2018; Kinzel et al., 2016, 2011). Occupation at Shkārat Msaied is dated to 10,300–10,000 cal. BP based on 14 direct AMS dates obtained from charcoal (Hermansen et al., 2006) (SI1, Figs. S1, S2, Tables S1 & S2), with Phases I-III dated to the early MPPNB (Hermansen et al., 2006; Jensen et al., 2005; Kinzel, 2018, 2012). This period at Shkārat Msaied is characterised by circular stone buildings varying in size between 5–27 m². At least 25 circular stone
buildings have been exposed and reveal a long and complex history of the site, evident in the finds and the many modifications to the architecture (Jensen et al., 2005; Kinzel, 2019, 2013). Material from the site relates to daily life, and food preparation installations indicate the domestic nature of the buildings (Hermansen and Jensen, 2002; Kinzel, 2019).

Zooarchaeology. Zooarchaeological analysis was conducted by two authors (PWN and PB) using the reference collection at the Zoological Museum in Copenhagen and considering published criteria for the separation of sheep and goats (Boessneck, 1969; Boessneck et al., 1964; Payne, 2017; Rowley-Conwy, 1998; Zeder, 2006; Zeder and Pilaar, 2010). Bone fragments were assigned to taxonomic levels of identification and skeletal elements. Data was registered on a database, which also recorded element portion present, fusion data, tooth wear and anatomical measurements as well as any taphonomic alterations to the bones such as burning and cut-marks. Assessment of mortality profiles was also conducted on the identified caprines (SI2, Fig. S5). The current analysis of the faunal assemblage includes remains from recent excavations as well as data from the 1999–2000 seasons (SI2, Table S3) (Bangsgaard, 2004).

Sampling, protein extraction, enzymatic digestion, peptide purification. Based on zooarchaeological assessment we selected 28 petrous bones for analyses (SI2, Table S7). Soluble surface contamination was removed by incubating the samples with 100 μL of a 50 mM ammonium bicarbonate (ABC) (Sigma) solution pH = 7.8 for 16 h at RT. After which they were vortexed (15 s) before centrifugation for 1 min at 16,000 RCF. The supernatant was discarded and the samples were then dried, powdered and two 25 mg aliquots (Extraction 1 and Extraction 2) were made.

Extraction 1 - ABC buffered ZooMS extraction (van Doorn et al., 2011): Samples were incubated in 100 μL of 50 mM ABC at 65°C for 60 min, and the supernatants collected. 50 μL of the extracts were transferred to a new 1.5 mL Protein LoBind tube (Eppendorf). Digestion was performed by the addition of 0.4 μg of trypsin (Promega). The digestion was allowed to proceed for 16 h at 37°C. Subsequently, the digestates were pelleted (1 min, 16,000 RCF) and the pH lowered below 2 by the addition of 5 % (vol/vol) Trifluoroacetic acid (TFA, Sigma Aldrich). C18 ZipTips (PierceTM) were used to desalt and concentrate peptides which were subsequently eluted using 50 μL of 50 % acetonitrile (ACN) (Sigma Aldrich) / 0.1 % TFA (v/v).

Extraction 2 - Tryps-IN: 500 μL of 0.5 M disodium Ethylendiaminetetraacetic acid (EDTA) (Invitrogen) pH = 8 was added to each sample and demineralisation was performed for 24 h at 4°C with end-over-end rotation. Digestion was performed by the addition of 0.8 μg/μl of trypsin (Promega). The digestion was allowed to proceed for 16 h at 37°C. The ability of trypsin to enzymatically cleave proteins in 0.5 M EDTA was confirmed beforehand and is shown in Supplementary Information (SI4, Fig. S9). After digestion, the extractions were centrifuged at 16,000 RCF for 5 min without acidification with TFA as this causes precipitation of the EDTA. Purification and elution was performed as in Extraction Method 1.

MALDI-TOF-MS. Peptide eluates were spotted in triplicate onto a steel MALDI plate (Bruker Daltonics). 1 μL of peptide-containing solution was mixed with 1 μL of matrix solution consisting of α-cyano-4-hydroxy-cinnamic acid (Sigma Aldrich) dissolved in 50 % ACN / 0.1 % TFA (v/v). A Bruker Ultraflex III (Bruker Daltonics) in positive mode was used for MALDI-TOF data acquisition with mass-to-charge range = 800–3200, laser acquisition = 1200 and mode = reflector. Data analysis was performed using mMass v.5.5.0 (Strohalm et al., 2010). Triplicate spectra were merged and averaged before peak picking (signal-to-noise threshold = 4).

LC-MS/MS. LC-MS/MS data acquisition was performed using an EASY-nLC 1200 (Proxeon, Odense, Denmark) connected to a Q Exactive HF-X for the ABC ZooMS extractions or to an Exploris 480 (both mass
spectrometers: Thermo Scientific, Bremen, Germany) for those extracted by the Tryps-IN method. Ten μL of each sample were transferred to a 96-well MS plate and evaporated using a centrifugal evaporator until the volume was ~3 μL. Samples were then topped up with 5 (ABC) or 5 μL (Tryps-IN) of 0.1 % TFA, 5 % ACN. Five μL of the ABC ZooMS extractions were injected, whereas 3 μL were used for those extracted by Tryps-IN. The parameters for LC-MS/MS were the same as previously used for palaeoproteomics samples (Brandt et al., 2022; Mackie et al., 2018).

Bioinformatics. The raw files generated were then searched using the software MaxQuant (v.1.6.3.4 and v.2.1.3.0) (Cox and Mann, 2008). MaxQuant settings were as described in (Jensen et al., 2020). Confident protein identification requires at least 2 razor + unique peptides covering separate areas. Manual assessment of MS/MS spectra was performed to confirm identifications. Deamidation was measured as described in (Mackie et al., 2018). Raw files were then searched against a FASTA file database containing sequence data of alpha 1 and alpha 2 chains of type I collagen (COL1A1 and COL1A2) for those extracted by Tryps-IN. The parameters for LC-MS/MS were the same as previously used for palaeoproteomics samples (Brandt et al., 2022; Mackie et al., 2018).

The morphological analysis conducted here led to the identification of 7,464 remains (NISP, SI2, Table S5). A total of 6,499 bones were identified as subfamily Caprinae, of which 101 were identified as wild sheep or mouflon (Ovis orientalis), including two horncores. Due to the extensive fragmentation of the remains, the identification of the caprine remains was restricted, resulting in many elements being identified as sheep/goat (SI2, Fig. S3). Evidence that the subsistence economy at Shkārat Msaieid included a significant portion of hunted caprines is suggested by the size range of the Capra bones (SI2, Fig. S4 and Table S4), indicating that both Nubian ibex (Capra nubiana) and bezoar (Capra aegagrus) are present, with the former well represented. Both bezoar and ibex, based on horncores, have been also found at the nearby MPPNB site of Beidha (Hecker, 1975). Only two bezoar horncores (and no ibex horncores) were identified at Shkārat Msaieid. Since there is no evidence that ibex were ever managed or domesticated, hunting was probably still a major facet of the animal economy at Shkārat Msaieid, given the number of bones from the smaller species of Capra (likely to be ibex).

Mortality profiles were assessed to identify animal management, and indicate the preferential killing of males as juveniles. However, since both ibex and bezoar are present in the assemblage from Shkārat Msaieid, it is difficult to interpret these results. The overlap between the size of the two Capra species makes it impossible to determine if an unfused, juvenile bone is a large male ibex or a small female bezoar. Comparison of mortality profiles of the caprines from Shkārat Msaieid to sites where hunting and herding was the main subsistence strategy is inconclusive (SI2, Figure S5 and Table S5). Limited osteometric evidence (SI2, Figs. S6 & S7) suggests that the sheep from Shkārat Msaieid identified morphologically, were comparable in size to that of wild populations in the region. Given the time lag between livestock under early management practices and morphological change (Daly et al., 2021; Zeder, 2006)) the sheep could have been either early managed animals or wild. The presence of perinatal remains of caprines, which form 0.9 % of the identified caprine assemblage at Shkārat Msaieid, is tentative evidence that some early management of caprines took place at the site (Péllat et al., 2021). The frequency is not as high as at sites, such as Ganj Dareh and Aşıklı Höyük, where the economy was clearly focused on herding (Hesse, 1978; Stiner et al., 2014; Yeomans et al., 2021a). The zooarchaeological analysis suggests that there is probably some animal management at Shkārat Msaieid but hunting of wild fauna was still important.

Tryps-IN method validation. To confirm the viability of the Tryps-IN approach we performed a series of validation experiments to compare the performance of trypsin in 0.5 M EDTA with commonly used buffers such as 50 mM ABC and Tris-buffered 0.6 M GuHCl. Firstly, SDS-PAGE was used to analyse the trypsin digestion of both bone extract and purified collagen (SI2, Fig. S8). The results of this analysis show that trypsin performed similarly in 0.5 M EDTA and 50 mM ABC. We then performed trypsin activity assay which also showed that the activity of trypsin in 0.5 M EDTA is comparable to activity in 50 mM ABC. Interestingly, the activity of trypsin as measured by this assay was significantly lower in Tris-buffered 0.6 M GuHCl (SI4, Fig. S9). Lastly, we performed LC-MS/MS analysis of HeLa cell digests to investigate cleavage efficiency (Fig. 2A). We found that trypsin has a similar cleavage efficiency in 0.5 M EDTA as 50 mM ABC and that a higher rate of missed cleavages was seen in Tris-buffered 0.6 M GuHCl. The reduced cleavage efficiency of trypsin in Tris-buffered 0.6 M GuHCl may explain the poor performance of Tris-buffered 0.6 M GuHCl in the trypsin activity assay.

ZooMS peptide mass fingerprinting. Hot and arid regions with wide temperature fluctuations greatly accelerate molecular diagenesis and often result in the advanced breakdown of organic molecules. Previous attempts using ZooMS on bones from this region have also been unsuccessful (Yeomans et al., 2017). We, therefore, selected twenty-eight petrous (pars petrosa) bones initially visually identified as deriving from caprines, most of which derive from Building F. The inner-ear petrous bone is one of the most compact bones in the mammalian skeleton and has been shown to preserve DNA remarkably well (Pinhasi et al., 2015). We used ZooMS to obtain collagen peptide ions for taxonomic assignment (Buckley et al., 2009). It is important to note that it is not possible to separate wild progenitors from domesticated caprines with this method (Buckley et al., 2010). In addition, we devised a streamlined method (Tryps-IN, Extraction 2) for fragmented-collagen purification and extraction and compared this (Fig. 3) to the standard methodology (Extraction 1). Whilst the ABC ZooMS protocol was unable to assign any sample down to the order level, Tryps-IN allowed to confidently assign 14 of the 28 samples to the Capra spp (Fig. 3, Table 1). One sample (SM8) was identified as belonging to the order Carnivora. Six were assigned to the order level (Bovidae or Cervidae) and six did not contain enough peptides for identification (Table 1). Interestingly, one sample (SM12) differed from the Capra sp. as it lacked the higher molecular peak at 3093 Da, unique to Capra and Rangifer (Buckley et al., 2017) but contained miniscule trace isotope peaks at the mass range of 3033 Da (Fig. 3A), which is present in sheep and gazelles (Buckley et al., 2010; Janzen et al., 2021). However, due to incomplete
spectra we can only place this sample at the family level of Bovidae and potentially to sub-family Antilopinae.

To test for the presence of biomolecules in other skeletal elements, we also tested four long bones visually identified as either from *Capra*/*Ovis*. Neither of the methods used produced useful results (Table S8), although the Tryps-IN method did show more strong peaks, albeit in the low molecular range (Fig. S11). These results indicate that the preservation of collagen in petrous bones is superior to that of long bones.

**LC-MS/MS sequencing.** Based on our ZooMS results, we selected eight samples (SM4, SM5, SM8, SM10, SM12, SM18, SM23 and SM26) which contained ZooMS spectra of high intensity and were tentatively identified as either *Capra*, unresolved bovids or deriving from the order Carnivora. We used the remaining eluates from ZooMS for protein sequencing by liquid chromatography by mass spectrometry (LC-MS/MS). When comparing the deamidation levels of Shkarat Msaeid samples against publicly available modern reference bone proteomes (LC-MS/MS), we used the remaining eluates from ZooMS for protein sequencing by liquid chromatography by mass spectrometry (LC-MS/MS) and compared the deamidation levels of Shkarat Msaeid samples against publicly available modern reference bone proteomes with different methods used. The Tryps-IN method showed more strong peaks, albeit in the low molecular range (Fig. S11). These results indicate that the preservation of collagen in petrous bones is superior to that of long bones.

**Discussion.**

Demineralisation, typically performed using acid or EDTA, is a common step in palaeoproteomics sample preparation for mineralised tissues, although it is also common to omit this step and extract the already soluble protein fraction using ABC (van Doorn et al., 2011). We utilised EDTA instead of acid because it does not expose the likely degraded/fragmented proteins to low pH hydrolysing conditions (this was also the rationale behind choosing the ABC method which forgoes demineralisation). It has been known for some time that proteins are
liberated during demineralisation and that the demineralisation supernatant is an abundant source of protein (Cleland et al., 2012). Whilst the low pH of acid-derived demineralisation is definitely incompatible with trypsin (optimal pH = 7–9), there is also a strong belief in the field that EDTA used at demineralising concentrations (0.2–0.5 M) is also incompatible with trypsinisation and downstream sample preparation (Cleland, 2018; Cleland et al., 2012; Fagernás et al., 2020; Procopio and Buckley, 2017). To circumvent the apparent inhibitory effect, EDTA containing demineralisation supernatants are further processed by performing a buffer exchange before downstream digestion with trypsin. This is often done by dialysis, molecular weight cut-off filters, or more recently by single-pot, solid-phase-enhanced sample-preparation (SP3) (Cleland, 2018) and protein aggregation capture (PAC) (Battth et al., 2019). However, this practice is likely the result of some field specific dogma as trypsin does not require divalent cation (e.g. Ca\(^{2+}\), Mg\(^{2+}\) or Zn\(^{2+}\) etc.) co-factors for activity, although calcium binding by trypsin has been demonstrated to increase the stability of the enzyme (Green and Neurath 1953; Nickerson and Doucette 2022). Moreover, a combination of trypsin and EDTA has been used routinely in cell culture to subculture/passage cells for decades (de Oca et al., 1971). However, this concentration of EDTA is sufficient to strip Ca\(^{2+}\) from calcium-dependent adhesion molecules and even deplete intracellular calcium stores, i.e. Ca\(^{2+}\) is effectively sequestered by 0.5 mM EDTA in this setting (Hong et al., 2009).

All strategies to remove EDTA from the sample will inevitably incur some loss of protein/peptides. MWCO filters will lead to the loss of any peptide smaller than the cut-off. For SP3 and PAC methodologies, it has not been investigated whether precipitation of small peptides onto the bead surface is as effective as it is for intact or largely intact proteins. It is not inconceivable that efficiency of precipitation is proportional to solubility and smaller peptides are inherently more soluble. Therefore, the loss of small peptides that do not aggregate, is also likely to occur in washing steps. Even dialysis will incur protein/peptide loss, not through the semipermeable membrane but rather on it instead. Peptides/proteins are inherently ‘sticky’ and are always lost to some extent through non-specific binding (NSB) to surfaces. Reduction of NSB by a single-pot strategy is partially attributed to the success of PAC and SP3.

We have proven beyond doubt that trypsin is functional in EDTA concentrations sufficient for demineralisation (0.5 M). Surprisingly, we found that trypsin has a similar activity (SH, Fig. S9) and cleavage success rate (Fig. 2A) in 0.5 M EDTA as 50 mM ABC. By harnessing this activity we were able to omit predigestion clean-up/buffer exchange.

Fig. 3. Species identification by mass spectrometry. A) Peptide mass fingerprint of SM12 sample processed using Tryps-IN (upper) and ABC ZooMS (lower) workflows. Whilst the ABC ZooMS protocol failed to provide species level ID, the Tryps-IN method gave tentative species ID of sheep. B) Peptide mass fingerprint of SM23 sample processed using Tryps-IN (upper) and ABC ZooMS (lower) workflows. Whilst the ABC ZooMS protocol failed to provide species level ID, the Tryps-IN method gave tentative species ID of goat.
steps and directly digest the demineralisation solution and potentially additional proteins/peptides bound to the tube through NSB. In addition, C18 reverse phase clean-up/desalting appears to be sufficient to remove EDTA for LC-MS/MS applications, thereby reducing sample loss and reducing workflow complexity. Tryps-IN was shown to outperform the widely used ZooMS method of protein extraction by gelatinisation in this study. One explanation is that Tryps-IN is ideally-suited to the analysis of highly degraded protein-poor samples (such as those analysed here) where only low amounts of fragmented proteins are present, as these samples are the most at risk of losses during clean-up procedures. The ABC method also has minimal clean-up processing but it does not include any demineralisation which likely liberates additional protein/peptides. Whilst Tryps-IN has been shown to aid the application of palaeoproteomics to highly degraded samples from arid environments, it does not come without a limitation. In order to most efficiently digest a sample, it is valuable to have an accurate protein-to-enzyme ratio. However, 0.5 M EDTA is incompatible with BCA protein concentration (Bar-Oz and Dayan, 2007). Before further analysis, it is therefore recommended that petrous bones are fully documented by microCT (Pálsson et al., 2019).

Given that Shkārat Msaied is located in a region known for poor preservation of collagen, but also at the centre of many debates in zooarchaeology, the range of research questions that could be further explored by the application of this methodology is wide. However, despite the improvements, limitations still remain in that the petrous bones are usually separate from the rest of the crania eliminating any information about the age-at-death of the animal, and comparison to osteometric data from specimens is currently limited. An exciting future development of these findings is the potential to aid the development of integrated methodologies for simultaneous sample preparation for a combination of proteins/peptides, DNA, and isotopic analyses.

6. Conclusion

Trypsin activity in 0.5 M EDTA is comparable to that in 50 mM ABC, allowing digestion of proteins to be performed directly in EDTA demineralisation solution i.e. the Tryps-IN approach. Tryps-IN outperformed the established ABC ZooMS extraction in our sample set from Shkārat Msaied on remains dated from ca. 10,000 years ago, an arid environment with limited biomolecular preservation. Tryps-IN yielded 14 identifications to the sub-family level and 7 identifications to the order level. Skeletal element was found to be a primary driver of success

### Table 1

<table>
<thead>
<tr>
<th>Key</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
<th>Value 5</th>
<th>Value 6</th>
<th>Value 7</th>
<th>Value 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM1</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2792</td>
<td>2883.3</td>
<td>t 2883</td>
</tr>
<tr>
<td>SM2</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2883</td>
<td>2883.3</td>
<td>2899.3</td>
</tr>
<tr>
<td>SM3</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>3094.4</td>
<td>NoID</td>
</tr>
<tr>
<td>SM4</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2883</td>
<td>2883.3</td>
<td>2899.3</td>
</tr>
<tr>
<td>SM5</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2883</td>
<td>2899.3</td>
<td>NoID</td>
</tr>
<tr>
<td>SM6</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>3094.4</td>
<td>NoID</td>
</tr>
<tr>
<td>SM7</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>2999.4</td>
<td>NoID</td>
</tr>
<tr>
<td>SM8</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>3094.4</td>
<td>NoID</td>
</tr>
<tr>
<td>SM9</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>3094.4</td>
<td>NoID</td>
</tr>
<tr>
<td>SM10</td>
<td>1105.5</td>
<td>1427.7</td>
<td>1580.7</td>
<td>1648.8</td>
<td>2131.1</td>
<td>t 2899.3</td>
<td>3094.4</td>
<td>NoID</td>
</tr>
</tbody>
</table>

* t = trace isotope.

a = 1 Da shift due to deamidation.
Table 2: Genus/species of subfamily Antilopinae. Grey squares for some of the samples indicate that the corresponding peptide was not preserved in the sample (**As identified with available sequences, but most probably a local sp.).

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Assigned species identity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM_4</td>
<td>Capra sp.</td>
</tr>
<tr>
<td>SM_5</td>
<td>Capra sp.</td>
</tr>
<tr>
<td>SM_10</td>
<td>Nanger sp. / Antilopinae</td>
</tr>
<tr>
<td>SM_12</td>
<td>Capra sp.</td>
</tr>
<tr>
<td>SM_18</td>
<td>Capra sp.</td>
</tr>
<tr>
<td>SM_23</td>
<td>Capra sp.</td>
</tr>
<tr>
<td>SM_26</td>
<td>Capra sp.</td>
</tr>
</tbody>
</table>

The mass spectrometry proteomics data (including raw data, MaxQuant searches resulting files and averaged MALDI-TOF-Ms access to LC-MS/MS instruments and resources respectively.

CRediT authorship contribution statement

Theis Zetner Trolle Jensen: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing – review & editing. Lisa Yeomans: Data curation, Formal analysis, Investigation, Resources, Validation, Writing – original draft, Writing – review & editing. Louise Le Meillour: Data curation, Formal analysis, Validation, Writing – original draft, Writing – review & editing. Pia Wistoft Nielsen: Formal analysis, Investigation, Resources, Validation, Writing – review & editing. Perline Bangsgaard: Data curation, Investigation, Resources, Validation, Writing – original draft, Writing – review & editing. Meaghan Mackie: Formal analysis, Validation, Writing – original draft, Writing – review & editing. Moritz Kinzel: Data curation, Resources, Validation, Writing – review & editing. Ingolf Thuesen: Data curation, Resources, Writing – review & editing. Matthew J. Collins: Funding acquisition, Resources, Supervision, Writing – review & editing. Alberto John Tauruzzi: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The mass spectrometry proteomics data (including raw data, MaxQuant searches resulting files and averaged MALDI-TOF TXT files) have been deposited on Zenodo and are accessible using identifiers https://doi.org/10.5281/zenodo.7858995 and https://doi.org/10.5281/zenodo.7861359. LC-MS/MS data, MaxQuant results files and fasta file containing the sequences used for analyses have been deposited on ProteomeXchange with dataset identifier PXD045412 using the PRIDE partner repository tool (Perez-Riverol et al., 2022).
Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jasrep.2023.104238.

References

Hesse, B.C., 1978. Evidence for husbandry from the early Neolithic site of Gajn Darej in Western Iran.