

In the format provided by the authors and unedited.

A guide to ancient protein studies

Jessica Hendy ^{1,9*}, Frido Welker ^{2,3,9*}, Beatrice Demarchi ^{4,5}, Camilla Speller ⁵, Christina Warinner ^{6,7,8}
and Matthew J. Collins ^{3,5}

¹Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany. ²Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. ³Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark. ⁴Department of Life Science and Systems Biology, University of Turin, Turin, Italy. ⁵BioArCh, Department of Archaeology, University of York, York, UK. ⁶Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany. ⁷Department of Anthropology, University of Oklahoma, Norman, OK, USA. ⁸Institute for Evolutionary Medicine, University of Zürich, Zürich, Switzerland. ⁹These authors contributed equally: Jessica Hendy and Frido Welker. *e-mail: hendy@shh.mpg.de; frido_welker@eva.mpg.de

Supplementary Information

Supplementary Table 1. Reporting of extraction blanks, injection blanks, evidence of protein degradation and MS data reporting in MS/MS-based ancient protein analysis publications. Extraction and injection blanks are marked as present when they are explicitly mentioned in the manuscript; if marked as absent, this does not necessarily suggest that these blanks were in fact not run or analyzed in the experiment, but they are not reported. MALDI-TOF-MS, MALDI-TOF/TOF-MS and antibody-based studies are not included. Accession numbers in the final column refer to datasets stored in ProteomeXchange, otherwise the name of other repositories is given; in one case this refers to a university-based ftp page that can be accessed using details provided in the relevant paper.

¹Degradation noted by the presence of smeared gels.

Supplementary Table 2. Demonstration of misleading species assignments in Mascot outputs. Of the top 20 eukaryotic proteins (ranked by score) identified from sheep tooth cementum, only 4 are assigned to sheep. Although the protein identifications themselves are expected for bone/dentine/cementum, misleading species assignments to *Bos taurus*, *Homo sapiens* and *Mus musculus* are made when the SwissProt database lacks the relevant sheep reference protein.

Supplementary Dataset 1. FASTA formatted file containing proteins (in)frequently identified as likely contaminants in standard palaeoproteomic research. This contains commonly encountered contaminants in proteomics laboratories, and consists of the common Repository of Adventitious Proteins, with the addition of several protein sequences we sometimes encounter as contaminants in our experiments.