BACK TO THE PAST. DECYPHERING CULTURAL HERITAGE SECRETS BY PROTEIN IDENTIFICATION

Maria Gabriella Giuffrida[a], Roberto Mazzoli[b], and Enrica Pessione[b,1]

Affiliations:

[a]National Research Council, Institute of Sciences of Food Production (CNR, ISPA), Largo P. Braccini 2, 10095 Grugliasco, Torino, Italy.

[b]Università di Torino - Dipartimento di Scienze della Vita e Biologia dei Sistemi, Via Accademia Albertina 13, 10123 Torino, Italy.

[1]To whom correspondence should be addressed: Enrica Pessione,
Department of Life Sciences and Systems Biology. Università of Turin. Via Accademia Albertina 13.
10123 Torino. Italy.
Tel. +39 011 6704644
Fax +39 011 6704508
E-mail: enrica.pessione@unito.it

ABSTRACT

The present review article reports the most innovative methods to detect proteins in historical and archeological samples as well as to characterize proteins used as binders in artworks. Difficulties to ascribe proteins to a certain animal species are often due to post-translational modifications originated by chemical or microbial deterioration during aging. Combining different techniques such as peptide mass fingerprinting and tandem mass spectrometry can solve some of these problems and also allow discrimination between taxonomically related species like sheep and goat. The most studied proteins in bones and textile samples are osteocalcin, collagen and keratin, whereas egg yolk and white proteins, casein and collagen are the most relevant for binders used in old paintings. With the suitable approaches (immune-based methods, DOT-blot, etc…) it is also possible to obtain *in situ* characterization or analyze the samples directly in the museum laboratories, with the advantage of avoiding artwork damage and expensive external commitments. Recent cutting-edge strategies allowed detection of proteinaceous infection markers that, for instance, were used to establish the cause of death of old Inca mummies and also proved the presence of *Yersinia pestis* in old documents dating from the period in 17th century in which the plague ravaged Europe.

Key-words: mass spectrometry, textiles, wall-paintings, archeology, keratin, infections, egg yolk.

INTRODUCTION

The analysis of archeological samples by chemical and/or biological techniques is of valuable interest for adding knowledge to the historical context and for obtaining information about daily life, human-environment interactions, historical transition periods, dietary habits and so on. On the other hand, characterizing ancient artworks is of interest, not only for shedding light on the manufacturing techniques used, but also to detect previous restoration interventions and in view of conservation strategies.

Technical aspects of protein identification from paleontological, archeological and art-work samples have been discussed in details in the extensive review article by Dallongeville et al. (2016) were almost all available methods for detecting ancient proteins are described. Cutting-edge technologies such as proteomics and mass spectrometry have emerged in the last two decades and, although mainly applied in the human health sector, they may be very useful to obtain insights into cultural heritage items and for complementing more traditional biochemical approaches such as enzymology. The present mini-review will explore the most suitable analytical methods reported in the literature for protein detection and characterization.

IDENTIFICATION OF PROTEINS IN ARCHEOLOGICAL AND HISTORICAL SAMPLES

Characterizing "everyday life" objects and/or ascribing bone remains to a certain species can supply useful information about human evolution and habits. The scientific approach to history such as molecular paleontology exploits the combination of different techniques, as a winning strategy to obtain in-depth characterization of archeological items such as bones, textiles, shells, potsherds and parchment-made objects.

Among the different methods used, the identification of ancient proteins by mass spectrometry (MS) proved to be the best fitting strategy (Tab 1). Like DNA, proteins contain an enormous reservoir of information that allows phylogenetic reconstruction (Ostrom et al. 2000). A novel term, *i.e.* paleoproteomics, has been introduced to open the way to a very new sector of proteomic investigations. However, as compared to the best known field of human health, paleoproteomics has to face at least two challenges: i) the small amount of protein samples and their frequent dispersion in a heterogeneous environment ii) the modifications that occurred during aging and the deterioration due to physical and microbial agents (Vinciguerra et al. 2016).

Other protein-based approaches to characterize archeological samples include ELISA (to detect hemoglobin and albumin) (Smith and Wilson 1990; Tuross et al. 1989) and individual amino

acid racemization (for dating archeological bones) (Demarchi et al. 2011). However, before starting discussing the protein-based methods it is worth reminding that also DNA recovery after PCR amplification (Brown et al. 2001) and lipid analyses (Evershed et al. 1995) have been used to obtain information on archeological bones.

Unlike the DNA-based analyses whose limit is the extreme fragility of nucleic acids that frequently undergo degradation or contamination (Corthals et al. 2012), lipid-based investigations proved to be successful since hydrophobic molecules, in the absence of surfactants, resist to hydrolase-mediated microbial degradation. In a paper by Evershed et al. (1995), steroidal compounds such as cholesterol, bile acids and diagenetic-cholesterol products were chosen as markers to assess paleo-diet. Dietary habits of prehistorical humans have been often established by measuring ^{15}N and ^{13}C in collagen. However, also the carbon skeleton of steroids is unaffected by diagenesis, therefore it can constitute a valuable source of ^{13}C to evaluate the ratio of C_3 (temperate zone) versus C_4 (tropical zone) plants in the diet. Gas chromatography (GC)/isotope ratio-MS can be used for paleo-diet investigations. The advantage of analyzing steroids lies on the fact that bacteria, possibly contaminating the bones, do not synthesize cholesterol but only hopanoids therefore any microbial contamination is easily detectable by the chromatographic profiles. A further possible source of chemical contamination is the burial ground. In this case, the authors demonstrated that the total lipid composition of soils immediately surrounding the burial ground (mainly consisting of alkanes, wax esters, fatty acids and long-chain alcohols) is very different from the bone samples, whose major components are cholesterol and its diagenetic products. This investigation constitutes a valuable example of tracing human diet by analytical approach.

Bones

The possibility to identify archaeological or ancient bones at the species level through methods based on protein characterization is of great interest because it offers precious phylogenetic and diagenetic information. Until twenty years ago, it was not possible to exploit the potential information contained in ancient proteins due to the sensitivity limits of the techniques used for protein identification at that time, as Edman degradation or amino acid analysis. More recently, the application of MS to peptide and protein analysis has allowed a rapid and easier protein identification and characterization in this peculiar field. The most frequently used techniques are Matrix Assisted Laser Desorption Ionization-Time Of Flight Mass Spectrometry (MALDI-TOF MS) and Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS), both useful for protein characterization (Ostrom et al. 2000).

One of the first applications of MS in fossil organic material was on *osteocalcin*, a protein presents only in vertebrate mineralized tissue and not in common fossil contaminants such as microbes, invertebrates and plants. This protein has a highly variable N-terminal sequence, hence is very useful for taxonomic studies. Ostrom and colleagues (2000) developed a method based on SDS-PAGE, radioimmunoassay and MALDI-TOF analysis on both intact osteocalcin (obtained by matrix demineralization followed by reverse phase HPLC purification) and its tryptic peptides to establish taxonomy affinities and diagenetic changes. Information on intact protein was not so useful because of the variability of the molecular weight of the protein in different species. More informative were the data collected using trypsin-digested proteins. More in detail, chemical derivatization with tris-trimethoxyphenyl phosphonium acetyl N-hydroxy succinamide ester was used to form N-terminal tris-trimethoxyphenyl phosphonium acetates at the N-terminus of the tryptic peptides. This novel derivatization approach of peptides permits a complete sequence fragment series of predominantly "a-type ions" to be obtained when a single trypic peptide is analyzed by MALDI-Post Source Decay (PSD). It is so possible to collect direct information on peptide sequence of ancient species opening new possibilities for molecular phylogeny, comparative biochemistry, and an understanding of the diagenetic changes of ancient macromolecules.

A deep characterization of osteocalcin from a Neanderthal was achieved by the combination of MALDI-TOF MS, sequencing after high-energy Collision Induced Dissociation (CID) for fragmentation of peptides and N-terminal amino acid sequencing (Nielsen-Marsh et al. 2005). This investigation revealed that the amino acid sequence of the Neanderthal's osteocalcin is identical to that of modern humans. Moreover, the osteocalcin sequences of Neanderthal, modern human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), and orangutan (*Pongo pygmaeus*) are unusual among mammals in that the ninth amino acid is proline (Pro-9), whereas most species have hydroxyproline (Hyp-9). The hydroxylation of Pro-9 in osteocalcin requires adequate concentrations of C vitamin and it depends on enzyme recognition of the target proline substrate consensus sequence Leu-Gly-Ala-Pro-9-Ala-Pro-Tyr occurring in most mammals. The researchers suggested that the absence of hydroxylation of Pro-9 in the four investigated species may reflect a response to a selective pressure related to a decline in vitamin C in the diet during adaptation to omnivorous habit.

Despite promising features of osteocalcin for species identification from archeological bones, many studies have been focused on the detection of peptides from bone *collagen*. The advantage is that collagen (particularly type I, the dominant protein in mineralized tissues) is stable over longer time-periods than ancient DNA and it can also be sampled directly from bone. One of the first publication in this field was based on conventional trypsin digestion of collagen and peptide analysis by using solid-phase extraction followed by Peptide Mass Fingerprinting (PMF). This approach,

called Zooarchaeology by mass spectrometry (ZooMS), uses the persistence and slow evolution rate of collagen (Fig. 1) as a molecular barcode to read the identity of bones (Buckley et al. 2009). Bones are identified by differences in peptide mass, resulting from sequence differences between species. The method was set up analyzing collagen from 32 different mammalian species and selecting 92 peptide markers normally used in species identification in processed food and feed. The combination of the selected markers could determine 26 of the 32 mammal species and was successfully applied to archaeological bones (actually, discrimination between sheep and goat bones was not possible). The difference between sheep and goat was later solved identifying a single collagen peptide with two different amino acid positions in bones from the Neolitic site of Domuztepe (Turkey) giving a correct species identification (Buckley et al. 2010). The ZooMS method was also able to identify different animal species dating from the lower-Pleistocene present in Weybourne Crag (1.5 million years) and in Happisburgh (900,000 years) old sites in the UK, where the signs of the earliest humans in Britain have been found (Buckley 2011). The study of archaeological marine mammals using ZooMS is interesting to understand the prehistoric and early historic human interaction with these animals (Buckley et al. 2014). A new collection of collagen peptides from modern marine mammals was generated and used for distinguishing a wide range of marine mammal species. This method could separate cetaceans and pinnipeds at least to the subfamily level. The characterization of bone collagen of extinct animals such as a 160,000- to 600,000-year-old mastodon (*Mammuth americanum*) and a 68 million years dinosaur (*Tyrannosaurus rex*) was obtained by a two-step proteomic approach from ion-trap MS fragmentation patterns (Asara et al. 2007). The authors claimed that this approach is a valuable tool to study the evolution and adaptation of ancient taxa when genomes are unlikely to be obtained. Ten years later the Buckley's group argued that the predicted sequences could be derived from a laboratory contamination, underlining the difficulty to predict peptide sequences without genome knowledge (Buckley et al. 2017).

The advent of more advanced proteomic techniques also allowed to identify *non-collagenous proteins* (NCPs) in archaeological sites. NCPs present in the bone extracellular matrix were investigated using conventional shotgun proteomics methodology following digestion with trypsin after elimination of collagen by bacterial collagenase (Wadsorth and Buckley 2014). The method was applied to a set of 19 bovine sub-fossil specimens (ranging in age from approximately four thousand to one and a half million years old) to determine NCPs degradation rate (in comparison to collagen) occurring over time in a temperate climate. Of the 44 total NCPs identified, 16 could be classified as blood or serum proteins, 12 were non-collagenous extracellular matrix (ECM) proteins, 10 were intracellular and just two (osteonectin and osteomodulin) were only found in bone. The protein displaying the longest life was identified as type 1 collagen. Among the NCPs, alpha-2-HS-

glycoprotein (A2HSG), serum albumin and the glycoprotein biglycan appeared to be the longest lived proteins frequently identified in ancient bones. The authors concluded that although type 1 collagen is the longest surviving protein in paleoarcheological remains, the potential for greater taxonomic resolution in NCPs sequence render them a better source of information than current collagen-based methods for establishing human-animal interactions occurring in the past (Wadsorth and Buckley 2014).

Of extreme interest are the recent investigations at the site of Denisova Cave, Russia, a Paleolithic site that contains a large number of bones most of which lack the diagnostic features necessary for traditional morphological identification (Brown et al. 2016). In order to facilitate the discovery of human remains, the *"collagen fingerprinting" combined with mitochondrial DNA* and radiocarbon analysis was applied to 2000 fragmented bones. Only one bone fragment, probably a distal phalanx excavated from a Pleistocene level, led to the discovery of a previously unknown hominid population more than 50,000 years old, genetically distinct from both anatomically modern humans (AMH) and Neanderthals, named Denisovans. The genome studies revealed that Neanderthals had contributed DNA to Denisovans, just as Neanderthals and Denisovans had contributed DNA to AMHs. This suggests that Denisovans and Neanderthals may have inhabited the Altai region of Russian Siberia in close chronological proximity to one another, and even perhaps co-existed here periodically.

The proteomic approach was also used to predict the possibility to find ancient DNA (aDNA) in skeletal remains (Wadsworth et al. 2017). The techniques to extract aDNA are time-consuming and expensive and predicting its presence by alternative methods (i.e. proteomics) is extremely intriguing. The proteome of 69 archaeological cattle tooth and bone samples from multiple European sites were obtained by nanoflow liquid chromatography/electrospray ionization tandem mass spectrometry (nano-LC/ESI-MS/MS) using a LTQ-Orbitrap as detector. The comparison of these data with mitochondrial aDNA and amino acid racemization (AAR) data, including estimations of the relative abundances for seven selected non-collagenous proteins, indicated that the survival of aDNA in bone or dentine may correlate with the survival of some proteins, and that proteome complexity is a more useful predictor of aDNA survival than protein abundance or AAR. The lack of a strong correlation between the recovery of aDNA and the proteome abundance may indicate that the survival of aDNA is more closely linked to its ability to associate with bone hydroxyapatite crystals rather than to associate with proteins. This study did not provide potential biomarkers for aDNA, but suggested that the proteome complexity could be used to predict the presence of aDNA. Of course, more studies are needed to develop a validated method.

Other archeological specimens

MS was used to characterize protein remains in biological samples other than bones as well. By analyzing few milligrams in potsherds fragments through nano-LC/ESI-Fourier Transformed (FT) MS/MS, Solazzo and co-workers (2008) proved that Alaskan diet of the period 1200-1400 AD was based on seal meat (muscle tissue) because they found specific peptide markers of seal (*Phoca vitulina*) *myoglobin*, thus confirming the importance of seal (and not only whale) hunting in this area.

Five years later, the same group focused attention on species identification through selected peptides from *keratinous material* (Solazzo et al. 2013). In ancient and/or damaged artefacts, the species origin of the materials can be difficult to identify through visual examination. The possibility of distinguishing α-keratin-made materials (wool, hair, horn, hoof, nail, baleen, claws and quills) at the genus level was limited by the lack of keratin sequences in the public database. The authors overcame this problem by searching theoretical peptide sequences created by substitution of variable residues, thus producing a large range of possible new sequences. Combining PMF method and nanoLC/ESI/ Q-TOF MS/MS they were able to discriminate in all horn and hoof materials between important species used in the past. The best matches were manually confirmed and important species markers were characterized; one in particular exists in many variations in every genus (YSCQLSQVQSLIVNVESQLAEIR in *Ovis*). New sequences of this peptide were characterized for unknown species, for instance in a baleen sample of unknown origin. For example, it should be possible to distinguish the bowhead whale (*Balaena* genus, split time ~5.4 million years) from the grey whale (*Eschrichtius genus*, split time ~9 million years) based on this single peptide.

Corthals and co-workers (2012) combined proteomic strategies and DNA-based methods to establish the cause of death of three 500 years old Inca mummies by sampling lip tissue by a cotton swab. They found the presence of *Mycobacterium tuberculosis* (Fig. 2) by DNA analyses but they were not fully satisfied since infection not always results in disease. Therefore, they tried to detect the immune response of the host by shotgun proteomics thus having positive evidence of the pathological event probably causing death of the three young humans. Actually, the MS-identified proteins (*Cathepsin G, Serine-protease inhibitor, apolipoprotein A1 and A2, transthyretin, vitamin D-binding protein*) are all related to acute/chronic lung inflammation or even to mycobacterial pulmonary disease. This work had the merit to extend and enrich previous DNA-based researches for assessing the microbial agent causing death for infectious disease such as: i) the evidence of *Plasmodium falciparum* malaria infection as the most likely cause of death of Tuthankhamon (Hawass et al. 2010) and; ii) the presence of *Yersinia pestis* in the victims of the Black Death (Bos et al. 2011).

Very recently, a multidisciplinary international team identified proteins from *Y. pestis* in the register of death dating summer 1630, in the archives of Milano by the EVA film technology (ethyl vinyl acetate film studded with crushed strong anion and cation exchangers as well as C_8 resins) followed by LC−MS/MS analysis (D'Amato et al. 2018). Humidified EVA plastic films allow sample harvesting from whatever surface through simple contact for few (15-30) minutes. Proteins are eluted from the film, reduced, alkylated and finally digested with trypsin overnight prior to LC-ESI/TripleTOF analysis. This innovative methodology for biological sample harvesting, is particularly suitable in studying cultural heritage items since is based on non-destructive extraction of proteins and can be easily performed directly *in situ*, thus avoiding any damage to precious historical items (Manfredi et al. 2017). Additionally, this new method, that is fully validated for the quantification of proteins (e.g. BSA and ovalbumin from egg tempera as markers), can also be applied to small molecules such as dyes (e.g. carminic acid, alizarin, and indigotin) on several types of supports.

Cloths and Ornaments

The study of cloths and ornaments can reveal important features concerning the human history. A very interesting study on the Oetzi's clothing has been performed by MALDI-TOF MS by a German research group (Hollemeyer et al. 2008). Oetzi is a mummy belonging to the Neolithic period found in 1991 in the ice of the Tyrolean Alps. Because of the very low temperature, the body and the clothing were found to be exceptionally well-conserved since more than 5300 years.

Knowing how Neolithic-age humans created their cloths can reveal unsuspected abilities and in particular give information on animals used for textile/skin supplying. As an example, these data can be of interest in understanding whether wild animals (like deer, otter, wolf and bears) were used (thus indicating hunting-gathering habits), or alternatively if clothing was prepared from domesticated animals such as goat, sheep and cows (consistently with an agro-pastoral economy). Traditional methods to identify animal hair are based on hair morphology. However, physical, chemical and biological degradation can alter structural fiber rendering both macroscopic and microscopic evaluation questionable. Also a more recent approach, (*i.e.* the use of DNA amplification) has shown some limits because prehistoric humans used to chew leather, fur and other skin-derived items to soften their texture, hence creating the risk of human DNA contamination.

In the study by Hollemeyer et al. (2008), *keratin* was chosen as a biomarker, since it is the most abundant protein in hair (Fig. 3). Amino acid sequences of keratins, although similar, vary among different animal species, thus allowing a good degree of identification at the species level.

Analyses were performed on coat, leggings and moccasins belonging to Oetzi. The PMF from ancient samples and from reference species, mostly occurring in the Alpine surroundings, were collected and compared to each other. Multidimensional scaling (useful to discriminate the zoological order level), binary hierarchical cluster tree analysis (reaching the family or subfamily level) and screening of the spectra against protein database allowed the identification of mammals down to single species level. For example, canid species were discovered in Oetzi's leggings, but could not be differentiated to species level. On the contrary, red deer was found in his shoe vamp, goat in the leggings, cattle in his shoe sole and at his quiver's closing flap as well as sheep and chamois in his coat.

Skin clothing belonging to Danish museum collections were analyzed by a research team of the Copenhagen University together with the Center for Textile Research of Copenhagen (Schmidt et al. 2011). To maximize the probability for obtaining a correct species attribution three methods were used: microscopic observation, DNA analyses and MS-based protein sequencing. Even in this case microscopy gave poor results: actually, while hair from wild animals conserved its structure unchanged over times rendering comparison with present species easy, hair form domesticated species are very different form original prehistorical samples because of the breeding activities and this makes comparison with reference species very problematic. Similarly, DNA-based analyses after PRC amplification had some constraints since these textile materials were conserved in a very acidic environment that altered DNA stability. The innovative approach was therefore to analyze *collagen* as the reference protein for species determination. Collagen type 1 alpha 1 and alpha 2 (present in all animal tissues) as well as collagen type 3 alpha 1 (very abundant especially in skin) were found after MS analysis. Sequence analyses allowed species attribution, although differentiation between sheep and goat was questionable. However, the overall results brought important information such as that some clothing were from cattle while some others from sheep/goat but no wild animal skin was used, thus confirming the attribution of these textiles to the agro-pastoral period of human history.

Collagen PMF was also employed to characterize ancient combs found in the Northern Scotland and presumably belonging to the 8^{th}-9^{th} centuries AD (Von Holstein et al. 2014). These objects were used for grooming but also as ornaments or gifts. Two types of combs were considered, that is native-type and Norse-type. MS was applied to 20 combs and revealed that 11 were from red deer antler, 4 from reindeer antler and one from whale bones. It is important to underline that Norse-type combs were all from reindeer, whereas native-type combs were from red deer. The results, obtained in a non-destructive manner on a small sample and without previous demineralization of bone or antler samples, were in agreement with analyses on DNA whose amplification had been possible only for ten samples. The value of this study is consistent with the importance of establishing whether peaceful contacts between Atlantic Scotland inhabitants and Scandinavians occurred before

the Norse political domination of the Viking Age and the late medieval period. Clearly, these results demonstrate that Norse-type and native type combs were made using different local materials (animals), supporting the idea that no cultural contamination occurred before the 10th century.

PMF was applied together with bulk amino acid composition/racemization and Raman spectroscopy to assess the species of mollusk shells used for necklace fabrication during the Bronze Age (De Marchi et al. 2014). The mineral skeleton preserved the proteins responsible for the biomineralization process during time and the authors were able to complete species attribution with minimal destruction and minimal sample requirement (less than 2 mg of shell powder).

Parchments

Parchments have been used as a writing medium for over two millennia. They are a writing material made from specially prepared untanned skins of animals, primarily sheep, calves, and goats. Traditionally, the species identification of these materials relies on visual (macroscopic and microscopic) and tactile examination. The application of the proteomic techniques proved to be very useful to clarify the nature of the parchment of a pocket Bible (Fig. 4) delivered by a Franciscan friar to the Mogul Emperor at the end of XIII century (Toniolo et al. 2012). Because of the thinness of the parchment, the experts opined the pages were produced from foetal lambskins. The nLC/ESI/LTQ MS/MS analysis on the tryptic peptide mixture obtained from a tiny fragment of the margins of a foil yielded the identity of 8 unique proteins from *Bos taurus* which indicated the origin of the parchment from calfskins rather than from foetal lambskins at least for that fragment.

More recently, identification of the taxonomical origin of parchment specimens named "uterine vellum" (used for XIII A.D. century Bibles) was obtained by a particularly sensitive and non-invasive strategy called "eZooMS" (electrostatic ZooMS). The method consists of a tribo-electric extraction of proteins, that is by using an electrostatic charge generated by gentle rubbing of a PVC eraser on sample surfaces, and the subsequent MS analysis by conventional MALDI-TOF based PMF (Fiddyment et al. 2015). Advantages of this approach include that no special equipment is required, therefore protein sampling can be obtained without the necessity to transport the artefact. Proteins harvested by this method, eluted by ammonium bicarbonate and tryptic digested, were analyzed by MALDI-TOF MS. Protein identification confirmed that ultrafine uterine vellum was not necessarily produced from abortive/new-born or thin-skinned animals (e.g. rabbits and squirrels) as previously supposed. They could be obtained from the skins of maturing animals of several species (more than one mammal species in a single manuscript, consistent with the local availability) through processes able to generate thin and soft high-quality parchment. Following studies have demonstrated that eZooMS can be extended to harvest other molecular species (e.g. DNA) relevant for studying ancient

items, such as York Gospels dating from the XI century A.D. (Teasdale et al., 2017). In the near future, tribo-electric extraction of biological samples may become one of the election methods in exploring ancient artefacts.

IDENTIFICATION OF PROTEINS IN ARTWORKS

Artworks can contain proteins telling the story of their birth. This has a huge importance both for their authentication and for getting insights into the painting technique as well as into the attitudes of the author who created the artwork.

As an example, a recent study by Zilbestein and co-workers (2016) demonstrated the presence of a morphine-derivative (6-O-acetyl morphine) in the last draft of Michail Bulgakov's "Master and Margarita" that revealed the opioid addiction of the writer. The scientific approach to discovering these interesting aspects was based on a cation exchange separation followed by GC-MS. In a following study, three *proteinaceous biomarkers of a nephrotic syndrome* were also found among the detected proteins, suggesting that probably the kidney intense pain related to this syndrome rendered necessary this drug treatment in the last years of Bulgakov's life (Zilberstein et al. 2017). However, the majority of literature articles concern the detection of *binders* used both in wall and paper paintings.

Binders in art samples

Knowing the binder composition of wall and easel paintings is a precious requisite for both revealing possible previous restoration procedures on the original painting and approaching the correct conservation method when necessary. Several proteins/proteinaceous materials were of common use as binders in ancient times: *caseins, whey globulins* and *albumins* from milk, *gelatin, collagen* from different animal species, as well as *egg yolk* and *white*. Traditional strategies for protein identification include the use of acid hydrolysis, amino acid derivatization and quantitative determination by chromatographic methods (Colombini and Modugno 2004).

A procedure for extracting proteins from small samples without protein hydrolysis was set up by testing different protocols on egg-based Renaissance paintings (15^{th} and 16^{th} centuries). As shown in Fig 5, *egg white* and *yolk* proteins were identified by both MALDI-TOF and tandem MS (nano-LC/nanoESI/Q-q-TOF MS/MS) analyses (Tokarski et al. 2006). More recently, Gambino and co-workers (2013) compared the MS-based proteomic strategy with a DOT-blot immunoassay. The latter approach showed very good discrimination ability between egg yolk and egg white proteins because of the high antibody specificity. They applied the method also to more complex matrices that included

pigments and artificially aged samples. The results were satisfactory, hence the technique was applied to gilding samples dating from the 13th century, which allowed identification of egg white and egg yolk proteins. The authors highlighted the importance of this fast, non-destructive method to find proteins in complex (mixed layers) and very old samples. Furthermore, they underlined the fact that this procedure is low-cost and easily adoptable by laboratories belonging to museums and conservation centers. Tripkovic and co-workers (2015) obtained good results in the characterization of ancient proteins by complementing MALDI TOF and nano-LC/ESI-MS/MS using a LTQ-Orbitrap as detector. They studied both model samples and 19th century icons from the orthodox churches of Holy Virgin in Baric (Serbia) and of Annunciation of the Holy Mother of God in Idvor (Serbia). They found that the combination of the two techniques allowed detection of peptides belonging to the same protein but having different physico-chemical properties. In particular, egg yolk proteins Vitellogenin-2 and Apovitellenin -1 were better identified by LC-MS/MS whereas bovine collagen (alpha1 and alpha 2 chains) was detected by both methods. The authors also compared the interference of two different pigments (French ochre and Zinc white) on the protein "detectability" and they found that the former has a negative effect on the egg yolk and egg white proteins.

Leo and co-workers (2009) investigated binders from samples of a Giotto's painting decorating St. Francis Church in Assisi (Italy) by CHIP/Ion Trap MS/MS and/or nLC/ESI-Q Trap MS/MS, by using direct tryptic cleavage on the artwork instead of complicated procedures of protein extraction. They identified the sequence of few peptides from *milk proteins*. This strategy is suitable to characterize cultural heritage items since very small samples are needed, the protein integrity is not required and especially the fact that is "non-destructive" (since the trypsin solution necessary for protein digestion can easily be removed, without damaging the artwork specimen).

Kuckova and collaborators (2007) were among the first authors who succeded in developing a method for binding identification on the painting "Sitting nude and grotesque masque" by Edvard Munch through MALDI-TOF MS. They collected a peptide signal database from proteinaceous binder model samples and then they applied the method to the famous painting discovering that the binder used was the whole egg. These authors found some constraints mainly concerning the animal species attribution of the *collagen glues*. These and other hurdles due to aging processes (that generated post-translational modifications, PTM, giving rise to unknown peaks in the MS spectrum) and the difficulty to ascribe some proteins to fish glue (since no fish proteins were present in the genomic and proteomic databanks at that time) have been overcome more recently (Leo et al. 2011; Dallongeville et al. 2011 and 2013). As far as PTM following aging are involved, Leo and co-workers (2011) analyzed Pisa cemetery frescoes by CHIP/Q-TOF MS and found that the major modification is consistent with deamidation of asparagine and glutamine. Other changes include an overall

403 decrease of lysine, methionine and tyrosine as well as the oxidation of serine, cysteine and
404 phenylalanine resulting in the presence of amino malonic acid in the samples. Since asparagine
405 deamidation occurs 10 time faster than glutamine deamidation, these authors have suggested
406 measurement of glutamine deamidation as the most reliable strategy to detect aging in ancient
407 historical items. The research group coordinated by Caroline Tokarski set up a method for typing the
408 animal glues at the species level (Dallongeville et al. 2011). In a first experiment, commercial
409 reference glues from different animal origin (cow, rabbit and fish) were analyzed by Fourier
410 Transform Ion Cyclotron Resonance Mass Spectrometry (FTICRMS). This strategy allows
411 identification of peptides from animal species whose sequences are not available in proteomic
412 databases. Fifteen bovine-, three rabbit- and three fish- specific peptides were identified. Then, the
413 same experiment was repeated on a complex mixture in order to detect possible interfering effects
414 due to linseed oil and lead white, i.e. the most common matrix in which the glue can be present.
415 Overcoming these difficulties has allowed identification of peptides in a more than one-hundred-year-
416 old binder sample of "Colle à Doreurs" (patented in 1868) and a gilt sample of the 18th century (Fig.
417 6) located in the St Maximin Church of Thionville (France). In these specimens, two rabbit collagen
418 specific peptides and 13 specific peptides from bovine collagen were found, respectively. It is
419 interesting to underline that gilding can be obtained by three ways: oil gilding (based on resins),
420 ground gilding (based on egg white or animal glue) and water gilding (based on gypsum, refined clay
421 and bovine or rabbit glue). In this paper, it was then possible to establish that the technique used was
422 water gilding.

423 The same research group studied a 17th century colored glaze belonging to the Holy Ghost
424 Altarpiece of the St Michael Church in Mondsee (Austria) (Dallongeville et al. 2013). It is known
425 that, prior to application of colored glazes on the silver leaf, ancient polychrome objects were treated
426 with a protective glue coating to improve optical properties. On the hypothesis of animal glue, the
427 authors first tried to detect proteins into the thin layers between the silver leaf and the glaze by SYPRO
428 Ruby staining. Once the proteins were actually detected, they characterized them by LC-MS/MS
429 achieving identification of seven peptides specific to fish collagen proteins. Fish glues were originally
430 prepared using the membrane of the swim bladder of sturgeons. In this work, peptide identification
431 was based upon similarities with peptides from other fish species. The advantage of using tandem
432 MS instead of PMF rely on the fact that even proteins modified by amino acid substitution, oxidation,
433 deamidation or proteins present in complex mixtures (as frequently occurs in ancient samples) can be
434 identified. Furthermore, without an ascertained genome sequence present in the database this remain
435 the only mean for identifying proteins of unknown animal origin.

436

CONCLUSIONS

Compared to DNA-based methods, proteomic studies take advantage from the higher stability of proteins and lesser sample contamination by operators. These are very important features when studying archaeological and historical specimens. Some constraints/hurdles concern the difficulties of finding all protein sequences in databases, especially for species (like fish species) whose genome is unavailable. Successful results have been obtained also in complex matrices such as "colle à doreurs" and sometimes the combination with immunological methods helps in clarifying the role of some proteins allowing deciphering the cause of death or existing pathologies in ancient samples.

Very promising strategies, such as the EVA film technology and the tribo-electric methodology, will constitute in the future the most suitable approaches to study cultural heritage samples avoiding the risk to damage such delicate and precious items. The two strategies can be employed for detecting a large variety of molecules (e.g. proteins, dyes, DNA) and can potentially become officially accepted methods employed by Museums and public libraries to inspect Cultural Heritage. The overall results underline the importance of such a biotechnological approach for acquiring information on ancient civilizations, dietary habits and human evolution.

COMPLIANCE WITH ETHICAL STANDARDS

Funding: this work was financially supported by "Ricerca Locale-ex 60%" of the Turin University.

Conflict of interest: all the Authors declare that they do not have any conflict of interest.

Ethical approval: this article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

Asara JM, Schweitzer MH, Freimark LM, Phillips M, Cantley LC (2007) Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry. Science 316:280-285.

Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S, Wood J, Earn DJ, Herring DA, Bauer P, Poinar HN, Krause J (2011) A draft genome of *Yersinia pestis* from victims of the Black Death. Nature 478:506-510.

466 Brown KA (2001) Identifying the sex of human remains by ancient DNA analysis. Ancient Biomol
467 3:215.
468 Brown S, Higham T, Slon V, Pääbo S, Meyer M, Douka K, Brock F, Comeskey D, Procopio N,
469 Shunkov M, Derevianko A, Buckley M (2016) Identification of a new hominin bone from Denisova
470 Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. Sci Rep 6:23559.

471 Buckley M, Collins M, Thomas-Oates J, Wilson JC (2009) Species identification by analysis of bone
472 collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid
473 Commun Mass Spectrom 23:3843-3854.

474 Buckley M, Collins MJ (2011) Collagen survival and its use for species identification in Holocene-
475 lower Pleistocene bone fragments from British archaeological and paleontological sites. Antiqua 1:1-
476 7.

477 Buckley M, Fraser S, Herman J, Melton ND, Mulville J, Pálsdóttir AH (2014) Species identification
478 of archaeological marine mammals using collagen fingerprinting. J Archaeol Sci 41:631-641.

479 Buckley M, Warwood S, van Dongen B, Kitchener AC, Manning PL (2017). A fossil protein chimera;
480 difficulties in discriminating dinosaur peptide sequences from modern cross-contamination. Proc R
481 Soc B 284:20170544.

482 Buckley M, Whitcher Kansa S, Howard S, Campbell S, Thomas-Oates J, Collins M (2010)
483 Distinguishing between archaeological sheep and goat bones using a single collagen peptide. J
484 Archaeol Sci 37:13-20.

485 Colombini MP, Modugno F (2004) Characterization of proteinaceous binders in artistic paintings by
486 chromatographic techniques. J Sep Sci 27:147-146.

487 Corthals A, Koller A, Martin DW, Rieger R, Chen EI, Bernaski M, Recagno G, Dávalos LM (2012)
488 Detecting the immune system response of a 500 year-old Inca mummy. PLoS One 7:e41244.

489 D'Amato A., Zilberstein G., Zilberstein S, Compagnoni BL, Righetti PG (2018) Of mice and men:
490 traces of life in the death registries of the 1630 plague in Milano. J. Proteomics pii: S1874-
491 3919(17)30411-6. doi: 10.1016/j.jprot.2017.11.028. [Epub ahead of print]

492 Dallongeville S, Koperska M, Garnier N, Reille-Taillefert G, Rolando C, Tokarski C (2011)
493 Identification of animal glue species in artworks using proteomics: application to a 18th century gilt
494 sample. Anal Chem 83:9431-9437.

495 Dallongeville S, Richter M, Schäfer S, Kühlenthal M, Garnier N, Rolando C, Tokarski C (2013)
496 Proteomics applied to the authentication of fish glue: application to a 17th century artwork sample.
497 Analyst 138:5357-5364.

498 Dallongeville S, Garnier N, Rolando C, Tokarski C (2016) Proteins in Art, Archaeology, and
499 Paleontology: From Detection to Identification. Chem Rev 116:2-79.

500 Demarchi B, O'Connor S, de Lima Ponzoni A, de Almeida Rocha Ponzoni R, Sheridan A, Penkman
501 K, Hancock Y, Wilson J (2014) An integrated approach to the taxonomic identification of prehistoric
502 shell ornaments. PLoS One 9:e99839.

503 Demarchi B, Williams MG, Milner N, Russell N, Bailey G, Penkman K (2011) Amino acid
504 racemization dating of marine shells: a mound of possibilities. Quat Int 239:114-124.
505 Evershed RP, Turner-Walker G, Hedges REM, Tuross N, Leyden A (1995) Preliminary results on
506 the analysis of lipids in ancient bone. J Archaeol Sci 22:277-290.

507 Fiddyment S, Holsinger B, Ruzzier C, Devine A, Binois A, Albarella U, Fischer R, Nichols E, Curtis
508 A, Cheese E, Teasdale MD, Checkley-Scott C, Milner SJ, Rudy KM, Johnson EJ, Vnouček J,
509 Garrison M, McGrory S, Bradley DG, Collins MJ (2015) Animal origin of 13th-century uterine
510 vellum revealed using noninvasive peptide fingerprinting. Proc Natl Acad Sci USA 112:15066-
511 15071.

512 Gambino M, Cappitelli F, Cattò C, Carpen A, Principi P, Ghezzi L, Bonaduce I, Galano E, Pucci P,
513 Birolo L, Villa F, Forlani F (2013) A simple and reliable methodology to detect egg white in art
514 samples. J Biosci 38:397-408.

515 Hawass Z, Gad YZ, Ismail S, Khairat R, Fathalla D, Hasan N, Ahmed A, Elleithy H, Ball M, Gaballah
516 F, Wasef S, Fateen M, Amer H, Gostner P, Selim A, Zink A, Pusch CM (2010) Ancestry and
517 pathology in King Tutankhamun's family. JAMA 303:638-647.

518 Hollemeyer K, Altmeyer W, Heinzle E, Pitra C (2008) Species identification of Oetzi's clothing with
519 matrix-assisted laser desorption/ionization time-of-flight mass spectrometry based on peptide pattern
520 similarities of hair digests. Rapid Commun Mass Spectrom 22:2751-2767.

521 Kuckova S, Hynek R, Kodicek M (2007) Identification of proteinaceous binders used in artworks by
522 MALDI-TOF mass spectrometry. Anal Bioanal Chem 388:201-206.

523 Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, Birolo L (2011) Deamidation
524 at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous
525 binders in mural paintings. Anal Chem 83:2056-2064

526 Leo G, Cartechini L, Pucci P, Sgamellotti A, Marino G, Birolo L (2009) Proteomic strategies for the
527 identification of proteinaceous binders in paintings. Anal Bioanal Chem 395:2269-2280.

528 Manfredi M, Barberis E, Gosetti F, Conte E, Gatti G, Mattu C, Robotti E, Zilberstein G, Koman I,
529 Zilberstein S, Marengo E, Righetti PG (2017) Method for Noninvasive Analysis of Proteins and Small
530 Molecules from Ancient Objects. Anal Chem 89:3310-3317.

531 Nielsen-Marsh CM, Richards MP, Hauschka PV, Thomas-Oates JE, Trinkaus E, Pettitt PB,
532 Karavanic I, Poinar H, Collins MJ (2005) Osteocalcin protein sequences of Neanderthals and modern
533 primates. Proc Natl Acad Sci U S A 102:4409-4413.

534 Ostrom PH, Schall M, Hasand G, Shen T, Hauschra PV, Strahler JR, Gage DA (2000) New Strategies
535 for characterizing ancient proteins using matrix-assisted laser desorption ionization mass
536 spectrometry. Geochim Cosmochim Acta 64:1043-1050.

537 Schmidt AL, Gilbert MTP, Cappellini E, Olsen JV (2011) Identification of animal species in skin
538 clothing from museum collections. Proceedings of the ICOM-CC 16th congress, Lisbon.

539 Smith PR and Wilson MT (1990) Detection of haemoglobin in skeletal remains by ELISA. J Archaeol
540 Sci 17:255-268.

541 Solazzo C, Fitzhugh WW, Rolando C, Tokarski C (2008) Identification of protein remains in
542 archaeological potsherds by proteomics. Anal Chem 80:4590-4597.

543 Solazzo C, Wadsley M, Dyer JM, Clerens S, Collins MJ, Plowman J (2013) Characterization of novel
544 α-keratin peptide markers for species identification in keratinous tissues using mass spectrometry.
545 Rapid Commun Mass Spectrom 27:2685-2698.

546 Teasdale MD, Fiddyment S, Vnouček J, Mattiangeli V, Speller C, Binois A, Carver M, Dand C,
547 Newfield TP, Webb CC, Bradley DG, Collins MJ (2017) The York Gospels: a 1000-year biological
548 palimpsest. R Soc Open Sci 4:170988.

549 Tokarski C, Martin E, Rolando C, Cren-Olivé C (2006) Identification of proteins in renaissance
550 paintings by proteomics. Anal Chem 78:1494-1502.

551 Toniolo L, D'Amato A, Saccenti R, Gulotta D, Righetti PG (2012) The Silk Road, Marco Polo, a
552 Bible and its proteome: a detective story. J Proteomics 75:3365-3373.

553 Tripković T, Charvy C, Alves S, Lolić AĐ, Baošić RM, Nikolić-Mandić SD, Tabet JC (2015)
554 Electrospray ionization linear trap quadrupole Orbitrap in analysis of old tempera paintings:
555 application to nineteenth-century Orthodox icons. Eur J Mass Spectrom (Chichester) 21:679-692.

556 Tuross N (1989) Albumin preservation in *Taima taima* mastodon skeleton. Appl Geochem 4:225-
557 229.

558 Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2016) Proteomic strategies for cultural
559 heritage: From bones to paintings. Microchem J 126:341-348.

560 Von Holstein ICC, Ashby SP, Van Doorn NL, Sachs SM, Buckley M, Meiri M, Barnes I, Brundle A,
561 Collins MJ (2014) Searching for Scandinavians in pre-Viking Scotland: molecular fingerprinting of
562 Early Medieval combs. J Archeological Science 41:1-6.

563 Wadsworth C, Buckley M (2014). Proteome degradation in fossils: investigating the longevity of
564 protein survival in ancient bone. Rapid Commun Mass Spectrom 28:605-615.

565 Wadsworth C, Procopio N, Anderung C, Carretero JM, Iriarte E, Valdiosera C, Elburg R, Penkman
566 K, Buckley M (2017) Comparing ancient DNA survival and proteome content in 69 archaeological
567 cattle tooth and bone samples from multiple European sites. J Proteomics 158:1-8.

568 Zilberstein G, Maor U, Baskin E, Righetti PG (2016) Maestro, Marguerite, morphine: the last years
569 in the life of Mikhail Bulgakov. J Proteomics 131:199-204.

570 Zilberstein G, Maor U, Baskin E, D'Amato A, Righetti PG (2017) Unearthing Bulgakov's trace
571 proteome from the Master i Margarita manuscript. J Proteomics 152:102-108.

FIGURE LEGENDS

Fig. 1. MALDI-TOF spectrum of tryptic peptides from pig collagen (unpublished data)

Fig. 2. *Mycobacterium tuberculosis*

Fig. 3. Structural features of beta (left side) and alpha (right side) keratin.

Fig. 4. Page of the "Bible of Marco Polo". This pocket Bible was made in the North of France in the 1230s and delivered the Mogul Emperor between the mid 13th and the mid 14th century. The Bible remained in China until the travel of the Jesuit father Philipp Couplet in 1685 who brought it to Florence as a gift for the Grand Duke of Tuscany, Cosimo III (modified from Toniolo et al. 2012).

Fig. 5. Proteins commonly identified in hydrolyzed protein extracts from egg (both white and yolk)-based painting by peptide mass fingerprint (Kuckova et al. 2007).

Fig. 6. Application of "colle à doreur".

Table 1 Approaches used to identify proteins in ancient samples.

Method	Protein	Sample	Reference
SDS-PAGE MALDI-TOF MS MALDI-PSD	Osteocalcin	Fossil bison bones	Ostrom et al. 2000
Edman degradation MALDI-TOF MS MALDI-PSD	Osteocalcin	Neanderthal bones	Nielsen-Marsh et al. 2005
MALDI-TOF	Collagen	Bones from32 mammalian species	Buckley et al. 2009
MALDI-TOF MS	Collagen	Sheep and goat bones	Buckley et al. 2010
MALDI-TOF MS	Collagen	Bone fragments	Buckley and Collins 2011
MALDI-TOF MS	Collagen	Bones from marine mammals	Buckley et al. 2014
LC/ESI/Ion-trap MS/MS	Collagen	Mastodon and Dinosaur bones	Asara et al. 2007
LC/ESI/Orbitrap MS/MS	Non-collagenous proteins (NCPs)	Bovine bones	Wadsworth and Buckley 2014
MALDI-TOF MS	Collagen	Hominid fossils	Brown et al. 2016
nLC/ESI/LTQ-Orbitrap MS/MS Amino acid racemisation	Fetuin-A, Prothrombin and other bone-related proteins	Cattle tooth and bones	Wadsworth et al. 2017
MALDI-TOF MS LC/ESI-FT MS/MS	Myoglobin and hemoglobin	Harbor seal muscle and blubber	Solazzo et al. 2008
MALDI-TOF nLC/ESI/Q-TOF MS/MS	Keratin	Marine mammals	Solazzo et al. 2013
µ-LC/ESI- LTQ-Orbitrap MS/MS	Human serum proteins Mycobacterium proteins Immune response to infectious disease proteins	Inca mummies	Corthals et al. 2012
MALDI-TOF MS	keratin	Oetzi mummy clots	Hollemeyer et al. 2008
MALDI-TOF MS	Collagen	Reindeer and red deer combs	Von Holstein et al. 2014
Amino acid composition Raman spectroscopy	-	Mollusk shells	Demarchi et al. 2014
nLC/ESI/LTQ MS/MS	Collagen	Parchment	Toniolo et al. 2012
GC HPLC Pyrolysis interfaced-GC	Egg proteins caseins	Paintings	Colombini and Modugno 2004
MALDI-TOF MS nLC/ESI/Q-q-TOF MS/MS	Proteins from egg white and yolk	Renaissance paintings	Tokarski et al. 2006

DOT-blot immunoassay CHIP/Q-TOF MS	Proteins from egg white and yolk	13th century gilding samples	Gambino et al. 2013
MALDI TOF nano-LC/ESI- LTQ-Orbitrap MS/MS	Egg yolk proteins Bovine collagen	19th century icons	Tripkovic et al. 2015
Direct tryptic cleavage CHIP/Q-TOF MS nLC/ESI-Q Trap MS/MS	Milk proteins	Giotto's painting	Leo et al. 2009
MALDI TOF MS	Whole egg proteins	The painting "Sitting nude and grotesque masque" of Edvard Munch	Kuckova et al. 2007
CHIP/Q-TOF MS	Collagen	Pisa cemetery frescoes	Leo et al. 2011
nLC/ESI-FTICR MS	Rabbit and bovine collagen	18th century gilt samples	Dallongeville et al. 2011
SYPRO Ruby staining nLC/ESI-FTICR MS	Fish collagen	17th century colored glaze	Dallongeville et al. 2013
Triboelectric protein extraction MALDI TOF MS	Calfskin uterine proteins	13thcentury medieval manuscripts	Fiddyment et al, 2015
Triboelectric protein extraction MALDI TOF MS	Calfskin and sheepskin proteins	York Gospels (medieval manuscript)	Teasdale et al, 2017
EVA film nLC-ESI/LTQ Orbitrap MS	Nephrotic syndrome protein markers	Last draft of Michail Bulgakov's "Master and Margarita"	Zilberstein et al, 2017
EVA film LC-ESI/TripleTOF system	Milk and egg proteins	16 th century frescos	Manfredi et al, 2017
EVA film nUHPL-ESI/Orbitrap Fusion trihybrid MS	Yersinia pestis proteins	Register of death at Milan archives (1630)	D'Amato et al, 2018

Fig. 1

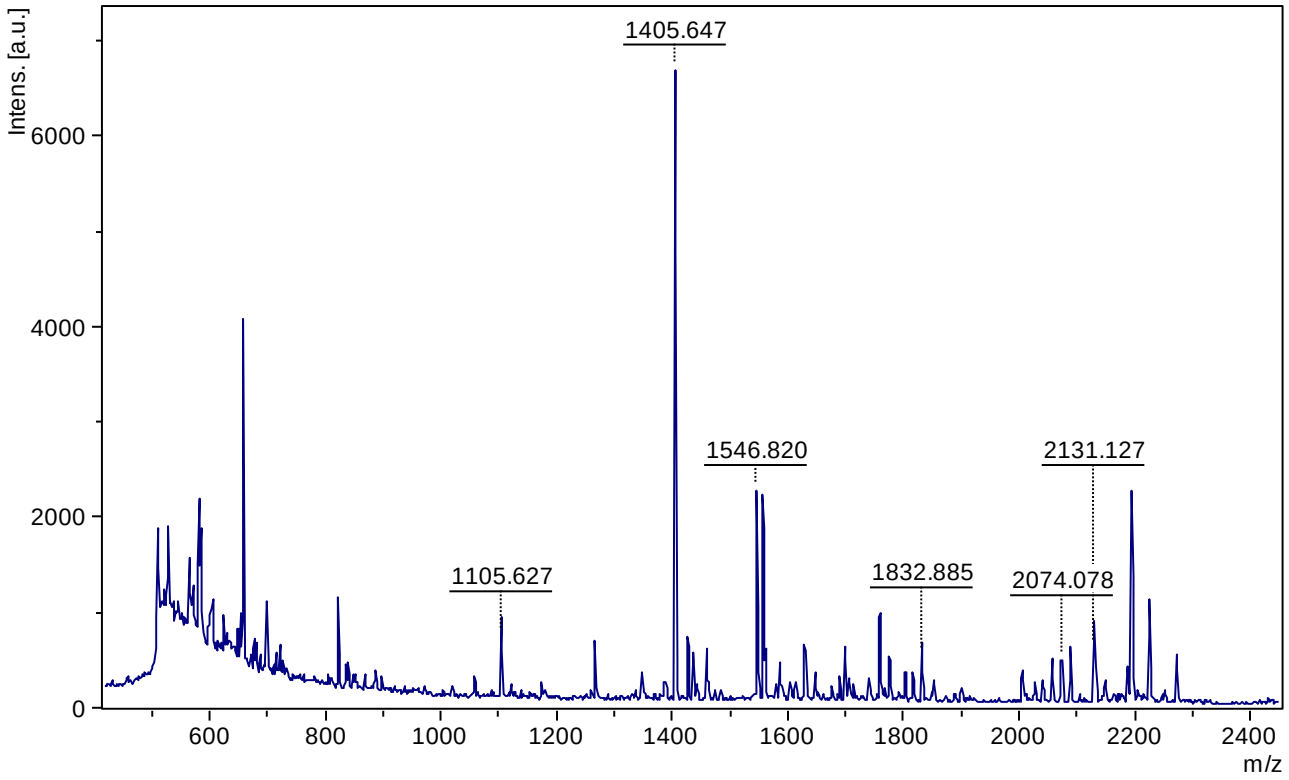
Intens. [a.u.]
6000
4000
2000
0
1405.647
1546.820
2131.127
1105.627
1832.885
2074.078
600
800
1000
1200
1400
1600
1800
2000
2200
2400
m/z

Fig. 2

Fig. 3

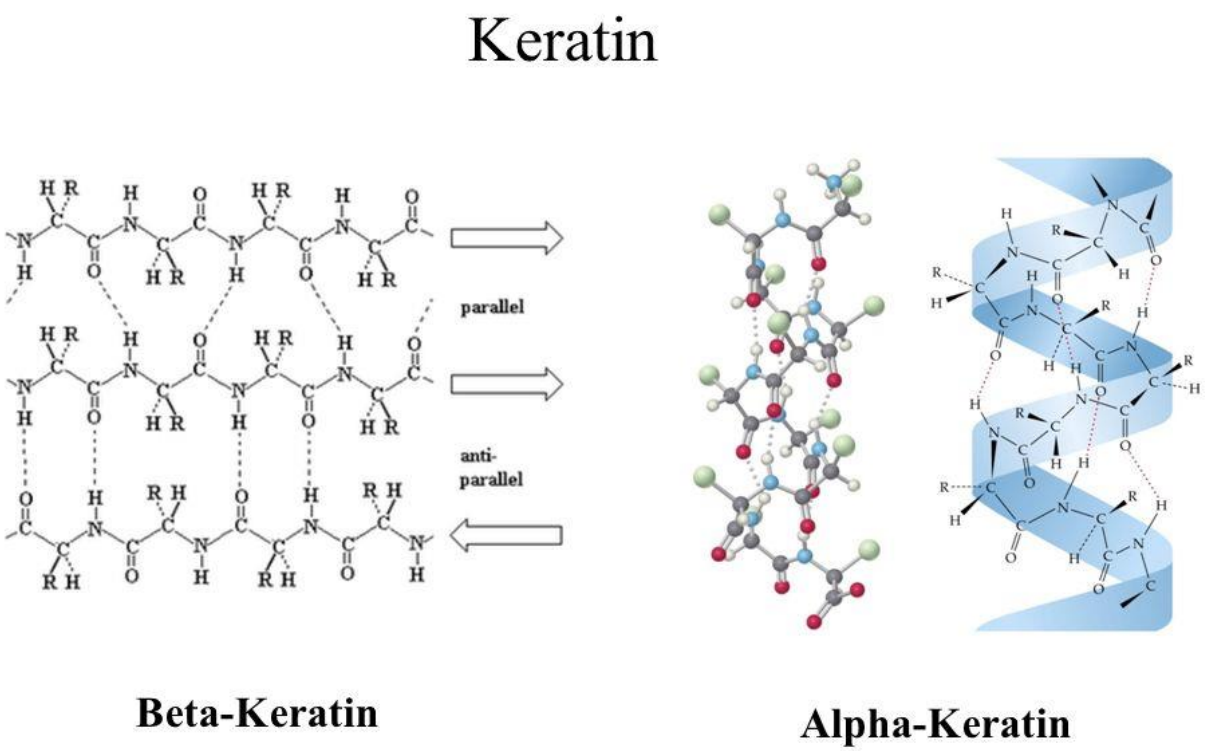
Keratin
parallel
anti-
parallel
Beta-Keratin
Alpha-Keratin

Fig. 4

Fig. 5

Protein	Sequence	Position
Low-density lipoprotein receptor-related protein I	KPEHELFLVYGK	523-534
	DSKRGKIER	675-683
	SDEKQSYCSSRKCK	2548-2561
Vitellogenin I	QFSSRPAYRR	355-364
	LTELLNSNVRLR	831-842
	LVTFEDPER	1058-1066
Vitellogenin II	LSSKLEISGLPENAYLLK	54-71
	ILGIDSSMFKVANK	523-534

Protein	Sequence	Position
Ovalbumin-related protein X	VKVYLPQMK	123-131
Ovomucin α-subunit	CMYDTCNAEK	620-629
	HCKSAAPVPVR	2030-2040
Ovomucoid	VEQGASVDKR	137-146

Fig. 6