Università degli Studi di Napoli Federico II

Ph.D. Thesis in Chemical Science

XXIX Course

PROTEOMICS AND MASS SPECTROMETRY FOR ARCHAEOLOGY AND PALEONTOLOGY

Addolorata De Chiaro

Tutor

Prof.ssa L. Birolo

Supervisor

Prof. A. Vergara

Coordinator

Prof. L. Paduano

Index

Abstract

The preservation of cultural heritage is central for protecting a sense of who we are, a meaningful reference in our culturally diverse world. (1)

Efforts to preserve resources on cultural heritage have gained new momentum throughout the world nowadays. The protection of our cultural heritage is a cultural but also economical process. (1)

It is nowadays broadly recognized that chemical analysis of materials can contribute to this process with important and often crucial information that are helpful for the study, preservation, or restoration of any kind of object of cultural heritage ranging from works of art to archaeological findings. The identification of materials in and on archaeological objects, monuments or artworks can open a window to the past for archaeologist and historians, and simultaneously helps in paving a safe way to the future, by allowing the conservators to make informed decisions regarding suitable preservation strategies. (2)

What has became increasingly important today is: WHAT (identification) has been found and WHERE (provenience). (3)

To start with, it is important to record what has been found, before you can delve on the questions about the provenance of the artefacts or the technology of how they were made, (4) and, in this perspective is also important to understand the chemical composition of each remain. (3)

Most artefacts are composed of complex mixtures of different molecules having a wide range of chemico-physical properties. Most of them are at a trace level, whilst others are in high amounts. (4) The main challenge in this type of work is the incredible variety and heterogeneity of the materials encountered. (5)

Some of them are facing deterioration, corrosion or destruction due to human activity as well as natural causes. Over the centuries, these factors increase and it is becoming more and more important to plan appropriate protection measures to save and preserve timeless evidences of human being (6).

If we are able to identify natural substances and their degradation products, we can shed light on the nature and the origin of the material employed and on the state of conservation, as well as on the degradation process itself.

Furthermore, analyses need to be as less invasive as possible. Moreover, samples are complex and heterogeneous. Therefore, it is essential to develop analytical protocols, which are minimally invasive and extremely high sensitive. (7)

To address these demanding analytical challenges, it is necessary to develop and use powerful and versatile analytical methodologies such as those based on mass spectrometry. These techniques offer unique means for advanced and microdestructive characterization of a wide spectrum of materials in archaeological findings, monuments, and works of art. (4)

My PhD project was focused on the development and optimization of chemical and proteomic analytical strategies based on advanced mass spectrometric techniques for archaeological remains.

The study of archaeological remains ranges from the molecular characterization of vessels or potteries' content to the analysis of organic remains such as human and animal tissues, for example bones, but also painting binders or vessels decorations.

After a brief introduction, a description of the Archaeological materials herein analysed is reported in Chapter 2, since this is the basis of the scientific interest in the search of new investigation procedures.
Since many of the challenges in this field arise from the complex composition and alteration pathways of the materials considered, this Chapter also provides the information of organic materials used, focusing the attention not only on protein-based materials.
To understand the potential of proteomics and mass spectrometric techniques applied on Archaeological objects, a description of the common methods is given in Chapter 3. Moreover, to help the understanding of the proteomic strategies developed and the various mass spectrometric methods applied, in this Chapter the main concepts in mass spectrometric instrumentation are also discussed.
Chapter 4 describes the development of a unique method that could be considered, efficient and applicable to different kinds of samples that contain proteinaceous materials.
In the chapters 5 to 8, a description of different study cases demonstrate the success of the experimental approaches.

The project was based on the identification and molecular characterization (by looking at chemical modification) of proteins and other organic molecules in archaeological finds like amphorae, tools and bones after optimization of analytical protocols and mass spectrometry approaches.

Specifically, the study was focused on the aging and degradation processes, defining the molecular markers of conservation state in human bones by studying different typologies of burials. Specific attention was paid to human bones from the excavations of Pompeii and Herculaneum. Moreover, the project was focused also to the characterization of embalming materials used in animal mummification process in Sudan. Another case of study deals with the identification of the animal species of a textile membrane thread from a drape found in the tomb of Emperor Henry VII of Luxembourg.

As a further section of my project, I worked to the development of a dedicated peptide database and targeted MS method (Multiple Reaction Monitoring, MRM), for the identification of binders in paintings as described in Chapter 9. The need to develop this method comes from the intrinsic analytical disadvantages due to the complexity of these unusual samples, in terms of chemical composition, presence of contaminating protein-based materials, as well as the high levels of molecular damage found in ancient samples. This method can be considered the first example of the application of targeted proteomics to samples in the field of cultural heritage that will allow an increase in sensitivity of protein detection in complex and rather contaminated samples.
The improvement in selectivity and sensitivity achieved with the targeted approach in respect to classical profiling experiments could pave the way to the possible application of this most advanced analytical strategy in the field of archaeology and artworks.

References

1. EkwelemVO,OkaforVN,UkwomaSC (2011) Preservation of Cultural Heritage: The Strategic Role of the Library and Information Science Professionals in South East Nigeria Library Philosophy And Practice562
2. Zhao W, Cioffi R, Maglioccola F, Catuogno R, Colangelo F, Russo L (2010) New technology and materials on the conservation of cultural heritage. Diagnosis for the conservation and valorization of cultural heritage (1) 319-321.
3. Gunneweg J. (2016) Clay, Technology and Qumran Pottery: What to verify and publish ? Researchgate.
4. Colombini MP, Modugno F (2009) Organic mass spectrometry in art and archaeology. Wiley, New York
5. Jakes KA (2002) Archaeological Chemistry. American Chemical Society.
6. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: Investigating the longevity of protein survival in ancient bone. Rapid Commun. Mass Spectrom (28) 605–615.
7. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.

1. Introduction

Culture can be defined as: the arts, beliefs, customs, institutions, invention, language, technologies, values and other products of human work and thought considered as a unit, especially as far as a particular time or social group is concerned.Culture produces similar behaviour and thought among most people in a particular society.

This definition of culture is adopted by Linton (1945) who says:

"*the culture of a society is the way of life of its members; the collection of ideas and habits which they learn; share and transmit from generation to generation (world Book Encyclopedia 2004) A people's cultural heritage, therefore, is their way of life and, in a broad sense, their traditional behaviour including the ideas, acts and artefacts which are passed on from one generation to another one (Banjo, 1997).* "

The preservation of cultural heritage is central for protecting a sense of who we are, a meaningful reference in our culturally diverse world. However, Sekler (2001) states that:

"there are many ways in which a cultural identity is formed and maintained. Much of the process has to do with the intangible cultural heritage of a body of traditions and usages, rites, poetry, song, and dance. A great deal of all these are passed on orally through generations. Consequently, its survival is always threatened…"

World Bank (2001) has warned that:

"Cultural matrixes contain elements of the human collective memory – language, believes and transmitted from generation to generation. Cultural references and signs are essential to the formation of national, group, and individual identities."

UNESCO (1995) argued convincingly that the prevalent model of development based solely on the narrow yardstick of economic growth is outmoded… the report concluded that, above all, cultural diversity is here to stay. It is a manifestation of the limitless creativity of the human spirit. Its aesthetic value can unfold in multiple ways and stimulate the production and marketing of new and unique products.

UNESCO has been extremely active in identifying world heritage sites all over the world. These sites are considered to be important for the world's cultural or natural heritage.

Efforts to preserve resources on cultural heritage have gained new momentum throughout the world nowadays. Protecting cultural heritage is economical, as well as historical and also a cultural process. Cultural heritage is based on the aspects of our past that we cherish, and they want to keep and pass on to future generations and outside world. However, the economic benefits of preservation are secondary to the intrinsic value of that heritage which is been preserved.

As rightly observed by Sekler (2001),

"tangible cultural heritage has the great advantage over its intangible counterpart, such that with proper care it will remain authentic over centuries. As long as historic monuments remain without falsification and misleading imitations, they will, even in a neglected state, create a sense of continuity that is an essential part of cultural identity". (1)

It is nowadays broadly recognized that chemical analysis of materials contributes important, and often crucial information for the study, preservation, or restoration of objects of cultural heritage. By identifying materials in and on archaeological objects, monuments or artworks, a window to the past is opened for archaeologist and historians, simultaneously paving a safe way to the future, by allowing conservators to make informed decisions regarding suitable preservation strategies.

The main challenge in this type of work is the vast diversity and variety of materials encountered, which include:

pigments, mineral or organic, natural or synthetic; metals or metal alloys and their corrosion products; stone and glass; bio-organic materials such as wood, leather, paper, parchment, but also oil, protein or sugar-based binding media and glues and synthetic or natural varnish coating and associated by-products arising during their aging.(2)

Some of them are facing deterioration, corrosion or destruction due to human made as well as natural causes. There are many external factors responsible for the damaging of the world's heritage materials, such as climate changes, light, pollution, armed conflicts, natural disaster, over-development and uncontrolled tourism. Over the years, these factors increase and it is

becoming more and more important to act appropriately measures to save and preserve these timeless achievements of human being (3).

Furthermore, given the value and sensitivity of works of art, analysis needs to be less invasive than possible, because they are often, carried out on the object itself or, in any cases, on small samples. (4)

Chemical analysis of samples provide very useful information but their scientific investigation requires quite challenging approaches, because these samples are complex and their materials heterogeneous.

The main challenges of the analysis are:

- Microscale: the amount of sample available for analysis are in the order of micrograms (max milligrams). This is because these studies involve invaluable works of art or archaeological finds that must be preserved.
- Intrinsic heterogeneity: the samples of art are very heterogeneous as they can contain a wide range of both organic and inorganic compounds (5).
- Extraction of the molecules: the samples are usually in powder or solid form. Sometimes, the analysis is not immediate but it is necessary to make the molecules available for the analysis.
- A minimally invasive analysis.
- Degradation of the original materials affected by aging and pollution.

Therefore, it is essential to develop analytical protocols, which are minimal invasive and extremely high sensitive.

In addressing these types of demanding analytical challenges, it has been necessary to develop and use powerful and versatile analytical instruments and methodologies as mass spectrometric and separation techniques, implemented separately or in combination. These techniques offer unique means for advanced and microdestructive characterization of a wide spectrum of cultural heritage's materials in archaeological findings, monuments, and works of art. (4)

1.1 References

1. EkwelemVO, OkaforVN, UkwomaSC (2011) Preservation of Cultural Heritage: The Strategic Role of the Library and Information Science Professionals in South East Nigeria. Library Philosophy And Practice 562.
2. Anglos D, Karydas AG (2009)Non-destructive and microanalytical techniques in art and cultural heritage. Analytical and Bioanalytical Chemistry (395) 1947-1948.
3. Zhao W, Cioffi R, Maglioccola F, Catuogno R, Colangelo F, Russo L (2010) New technology and materials on the conservation of cultural heritage. Diagnosis for the conservation and valorization of cultural heritage (1) 319-321.
4. Colombini MP, Modugno F (2009) Organic mass spectrometry in art and archaeology. Wiley, New York
5. Leo G, Cartechinini L, Pucci P, Sgamellotti A, Marino G, Birolo L (2009) Proteomic strategies for the identification of proteinaceous binders in paintings. Analytical and Bioanalytical Chemistry (395) 2269-2280.

2. Archaeological remains, from historical to artistic objects.

Archaeology is the study of man's past by scientific analysis of the material remains of his cultures

There are many kinds of archaeology, including archaeological science.

Archaeological science is a general term for laboratory methods in archaeology that includes both instrumental and noninstrumental areas such as faunal analysis, archaeobotany, human osteology, and even some aspects of stone and ceramics analysis. In some cases, these methods can also be applied in the field. (1)

Archaeometry is a specialized branch of archaeological science that involves the measurement of the physical or chemical properties of archaeological materials in order to solve questions about chemical composition, technology and chronology. Sometimes, it is described as "instrumental" archaeology. Archaeometry also includes areas like dating methods, remote sensing, ancient DNA. The term *molecular archaeology* is sometimes used to refer to the organic component of archaeological chemistry and particularly to the investigation of ancient DNA in plant and animal remains including human ones. (1)

However, the survival of proteins in fossil tissues has been of great interest for their potential in recovering phylogenetic information. In particular, proteins are abundant reserve of ancient biomolecules. (2)

Archaeological chemistry is a subfield of archaeometry and involves the investigation of organic and inorganic composition- elements and isotopes, molecules and compounds- of archaeological materials. Archaeological chemistry is primarily concerned with (a) characterization- measuring the chemical composition of a variety of prehistoric materials, and (b) identification- determining the original material of an unknown item. Information on identity and composition is used for different purposes such as: (a) authentication- verifying the antiquity of an item, often associated with works of art, (b) conservation- determining the optimal methods for preserving and protecting archaeological items. (1)

What becomes important today is: WHAT (identification) was found and WHERE (provenience). (3) The purpose of Archaeology and its intrinsically connected branch of Archaeometry can be summarized in the below W scheme.

WHICH products were manufactured?

WHEN has been an object made?

WHAT raw materials were involved to make artefacts?

WHY do certain artefacts appear or disappear?

WHERE have been artefacts made?

WHO were the producers and consumers?

WHERE were the goods transported TO?

What tools were artefacts made WITH?

What borders did the objects or trades men travel WITHIN?

To start with, it is important to record what has been found, before you can delve on the questions about the provenance of the artefacts or the technology of how they were made, (3) but it's also important to understand the chemical composition of each remain.

Most artefacts are composed of complex mixtures of different molecules having a wide range of chemico-physical properties. Most of them are at a trace level, whilst others are in high amounts. (4)

This chapter is focused on the materials that can be found in an archaeological object. Being able to identify natural substances and their degradation products is a challenge. If we manage to do so, then we can shed light on the nature and the origin of the material employed and on the state of conservation.

2.1 Artefacts

Humans have always used the natural materials around us to produce functional objects and works of art. Paintings and other objects that are part of our cultural heritage, including textiles, books, sculptures, archaeological objects, furniture and the organic residues found in association with them (e.g. cosmetics, medicines, perfumes, food), contain a wide variety of organic materials from natural to synthetic. Since ancient times, natural organic materials

have been employed as paint binders, adhesives, waterproofing materials and so on, as reported in classical literature by Plinius the Elder and Vitruvius. Archaeological excavations often bring to light a wide variety of objects and materials that have been collected, processed and used by humans over time. (4)

The main organic materials encountered in archaeological objects, are reported in Table 2.1. The table summarizes the organic materials and how they were once used.

TABLE2.1: Organic materials found in archaeological objects and contest how they were once used.

Category	Organic materials	Context
Proteins	Egg, milk and casein, animal glue, silk, wool, vegetable proteins, human and animal tissues	Paint binders, adhesives, textiles, parchment, tissues, vessels, other tools.
Glycerolipids	Animal fats, vegetable oils	Paint binders, varnishes, illuminants, ingredients of cosmetic and pharmaceutical preparations, vessels, other tools.
Waxes	Animal waxes,vegetable waxes, fossil waxes	Paint binders, coatings, sealants, writing tablets, sculptures, ingredients of cosmetic and pharmaceutical preparations, vessels, other tools.
Natural resins	Plant resins, animal resin, from thermal treatment of plant resins or wood	Varnishes, coatings, waterproofing materials, paint binders, ingredients of cosmetic and pharmaceutical preparations, vessels, other tools.
Polysaccharide materials	Starch, cellulose, plant gums	Paper, paint binders, adhesives, vessels, other tools.
Bituminous materials	Bitumen, asphalt	Moulding materials, adhesive, pigment, vessels, other tools.

Therefore, the study of archaeological remains ranges from the analysis of vessels or potteries' content to analysis of organic remains. Furthermore, it deals with analyses and molecular characterization of human and animal tissues, for example bones, but also binders characterization in the sarcophagus or vessels decorations.

2.2 Proteinaceous materials

Proteins represents promising molecules for the identification of archaeological and fossil samples of unknown origin, both human and animal, and food and goods too.

Both vegetable and animal proteins are encountered in archaeological objects as organic residues of commodities, for example food, human or animal tissues, or paint binders too. (4)

2.2.1 Animal tissue

Proteins represent a promising means for tissue identification in samples degraded by heat or time. (5) It is widely accepted that proteins have potential to survive significantly longer periods of time than DNA. (6)

Among tissues, bones are most interesting. Indeed, bones can survive in the burial environment for millions of years (7) and can provide direct information about the individuals during their life and post mortem. Bone contains both organic (mainly proteins) and inorganic components, with the most abundant protein being type I collagen. (8) The presence of the hydroxyapatite (mineral) crystals, which embed and protect the protein, contribute to the stability and preservation of bone over geological timescales (9, 10).

The application of proteomics to the study of ancient bones holds great potential for increasing our understanding of many aspects of "paleobiology" such as evolution, habits and ecology of ancient organisms.

Collagen, the dominant protein in tissues, is well preserved and highly conserved throughout the evolution of species due to its essential biological function as well as its peculiar sequence, strictly connected to its structural role. (5).

A collagen molecule is made up of three polypeptide strands (alpha chain), all in a left-handed helix conformation. Together, the three alpha chains are supercoiled in a right-handed coiled coil, which is stabilized by numerous hydrogen bonds between the chains. In turns, these tropocollagens aggregate further into fibrous right-handed coils and larger fibrillary structures. (11)

Vertebrates have 28 distinct collagen types that have their own specialized function in different tissues. However,Collagen types I to IV account about 90% of all collagens. Collagen type I is the most abundant collagen and is present in connective tissues, including skin, bone and cartilage. It consists of two alpha 1(I) chains and one alpha 2(I) chain. Collagen type II is the main component of cartilage and consists of three equal alpha 1(II) collagen chains. (11).

The amino acid sequence of collagen chains consists of monotonously repeating triplets of sequence Gly-X-Y where X is often Proline and Y is often hydroxyproline (post-translational modification of proline). (12)

These three aminoacids play an important role in the structure of collagen: the large side chain of proline and hydroxyproline force the chain in sharp twists, while glycine, without

side chain, is required at every third position (ca. 1/3 of the total sequence). The assembly of the triple helix puts this residue at the interior of the helix, which allows for a very close and tight wrapping of the three chains. (11)

Although collagen is the main protein found in bone (about 90% of the organic fraction), bone does contain a number of non-collagenous proteins, with osteocalcin being the second most abundant bone protein. Other proteins in bone include blood proteins. (13)

2.2.2 Tools, food and vessels content

Pottery is one of the most important sources of information for the archaeologist. Potsheds are one of the most common finds on archaeological sites of all periods. (14)

Large earthen containers were used for the storage of food such as butter, molasses, jam, cheese, (15) or other substances like for example cosmetic and pharmaceutical preparations. Some of these matrices contain proteins and the analysis of residues on ceramic samples is important because it may provide information on origins, uses and habits of the people and commercial trades with other populations. However, other types of molecules, such as oils, can be found in the ancient vessels, and the combined analysis of all substances may provide information about the artefacts. (14)

Besides, identification of proteins in ancient tools is important too. These analysis allow to understand a lot of information, similarly to the case of vessels, and may define the specific historical and social use of the analysed tool. In this case, other types of molecules can be found.

2.2.3 Paint binders

In paintings, animal proteins such as egg, casein and glue, are frequently used as binders for pigments. Egg can be used as whole egg or egg yolk. Moreover, one of the most important vegetable proteins is garlic (Allium sativum), which is used as an adhesive in gildings. (4)

2.3 Others archaeological organic materials

Organic materials in archaeological contexts may be the same or belong to the same classes, but can also vary considerably, often presenting different ageing pathways and chemical environments, thus requiring different analytical approaches.

Archaeological organic samples involve a range of critical characteristics, making their study an analytical challenge. Their investigation is complicated by the ageing and degradation of the molecules present in the organic substances. (4,16-19) This means that in order to

establish the origin of the natural substances in the sample and to understand the alteration processes that have modified the materials it is essential to identify the molecular composition as well as to reveal the presence of molecular markers. (20)

Most organic materials encountered in cultural heritage are macromolecular. In some cases, they are natural or synthetic polymers (such as proteins, plant gums, vinyl, and acrylic resins), others undergo oligomerisation or cross-linking reactions as an effect of exposure to light and air (such as natural resins or drying oils). Organic materials in the cultural heritage are also polar and have a low volatility. The answers clearly depend on the classes of organic materials found in the sample and the nature of the sample itself. Thus, some of the following chapter discuss the recent advances in GC/MS for the characterisation and identification of organic materials in archaeological contexts. (20)

Lipids of animal and plant origin as well as plant resins are the most commonly encountered substances in archaeological findings. In fact, owing to their lower susceptibility to structural modification and degradation by chemical and microbiological attacks as well as hydrophobicity, they may survive better than carbohydrates, proteins, and nucleotides in archaeological environments. (16-19) A similar argument can be made for the chemical components of alcoholic beverages and wines whose water solubility limits their persistence in archaeological environments. The detection of wine and other alcoholic beverage residues has always been of great interest in the archaeological community, and although some papers have appeared in the literature, (21-23) it still remains an analytical challenge. (20)

Oils and fats are mixtures of triglycerides, also known as triacylglycerols. They are basically esters of glycerol with fatty acids, and contain smaller amounts of other compounds, which include sterols and vitamins. The physical and chemical properties of individual oils and fats are determined by the nature and proportions of fatty acids that enter into the triglycerides composition. Animal and dairy fat like plant oils are dominated by triacylglycerols, with steroids as minor components, cholesterol and its esters being the most significative. The triacylglycerols of animal fats differ from plant oils since they contain more saturated fatty acids and consequently are solid at room temperature. (4)

Vegetable oils and dairy and animal fats were used extensively in ancient times in cookery, for lighting, and as ingredients of cosmetics, balms and medications. (16,24) Olive, almond, balanos, castor, coconut, linseed, moringa, palm, poppy, radish, safflower, and sesame oils were well known oleiferous species in the Mediterranean. (16,25) Data on the use of oils are

derived from papyri and from texts written by Theophrastus, Dioscorides and Pliny, which help to clarify the identification of plants cultivated for their oily seeds. (25)

The similarities in the composition of many vegetable oils used in ancient times and the way they might have been mixed together, means that oils exhibit complex molecular patterns that usually prevent us from identifying the original botanical source. Nevertheless, there are some vegetable oils that have a very specific composition. (4,20)

2.3.1 Embalming materials

The ancient Egyptians believed that the preservation of the physical body was essential for the continuation of life after death. Mummies are one of the most important sources of information about this practice. (26)

The material used for embalming have always been a matter of controversy. (27-29) Natural resins, gums waxes, oils and other natural substances, (30,31) together with various embalming practices, have been utilized over the centuries. (22)

The dark colour of many mummies led to the assumption that petroleum bitumen (or natural asphalt) was ubiquitous in mummification. (32)

These materials are chemically very complex and the composition of fossil organic matter depends on the kind of organism from which the deposit has formed. (31) Bitumen generally contains a complete range of homologous series of n-alkanes, and the isoprenes pristane and phytane. The hopanes and Steranes are biomarker compounds which can be related to their original biogenic precursors. (4)

Moreover, bitumen, asphalt, and other fossil organic materials such as coal, lignite and peat are found as natural deposits and have practically always been used in arts and handicrafts. They were used in medicines and cosmetics, as pigments and as adhesives. (31, 33,34)

Critically, the dark colour of balms can be simulated by heating/aging mixtures of fats, resins and beeswax known to be used in balms.

Natural waxes are highly heterogeneous lipid materials containing esters of long chain carboxylic acids, which are solid at room temperature and highly hydrophobic. Waxes of animal origin (beeswax, Chinese wax, lanolin, spermaceti wax), vegetable origin (carnauba, candelilla, esparto wax, Japan wax) and fossil waxes (paraffin wax, montan wax, ceresine) (31) have been used for many other purposes such as sealants, surface coatings and polishes,

casting and modelling materials, ingredients cosmetics, and lighting candles. Beeswax, obtained from the hives of bees, is the most commonly used natural wax for manufacturing works of art. Since prehistory, beeswax has been used as a waterproofing and sealing agent.

The Egyptians used it in balms for mummies, and also in ship building, to polish the surface of paintings, for lighting, and to make statues and writing tablets. (31,35,36) It was used by the Greeks and Romans to waterproof stone surfaces, as a protective agent and as a varnish. (37) Until the Middle Ages beeswax was used as a binder in a painting technique referred to as the encaustic technique. (37,38) Between the seventeenth and twentieth centuries the ceroplastic technique was developed for the realization of anatomical sculptures and botanical models. (29)

Furthermore, coniferous resins, like Colofony, were used in the embalming process. A typical molecule of aged archaeological coniferous resins is Abietic acid. (39) Although Egyptians use these resins in embalming process at least as early as 2200 bc, their use becomes most apparent in later periods. The increasing use of coniferous resin suggest that the embalmers may have become aware of the ability of specific natural products to inhibit microbial degradation by means of mechanisms analogous to their protective roles in the plants from which they derived. (16,27)

Other than Colofony, Pistacia resin was used in Egyptian contexts. This resins was a valued commodity in ancient Egypt and regarded as a luxury item because it was imported and difficult to obtain. (16) Indeed, its use in mummification has rarely been observed. The presence of Pistacia resin in the balm was indicated by the presence of triterpenoids oleanonic and isomasticadienoic acid and a range of oxidation products. (40)

2.4 References

1. Jakes KA (2002) Archaeological Chemistry. American Chemical Society.
2. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: Investigating the longevity of protein survival in ancient bone. Rapid Commun. Mass Spectrom (28) 605–615.
3. Gunneweg J. (2016) Clay, Technology and Qumran Pottery: What to verify and publish ? Researchgate.
4. Colombini MP, Modugno F (2009) Organic mass spectrometry in art and archaeology. Wiley, New York.
5. Buckley M, Collins C, Thomas-Oates J (2009). Species identification by analysis of bone collagen using matrix-assisted laser desorption / ionisation time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry (23) 3843-3854.
6. San Antonio JD, Schweitzer MH, Jensen ST, Kalluri R, Buckley M, Orgel J (2011) Dinosaur peptides suggest mechanism of protein survival. Plos One (6) 1-7.
7. Collins MJ, Riley MS, Child AM (1995) A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. Journal of Archaeological Science (22): 175-183.
8. Rich A, Crick FH (1961). The molecular structure of collagen. Journal of Molecular Biology, (3) 483-506.
9. Turner-Walker G, (2008) The Chemical and Microbial Degradation of Bones and Teeth. Advances in Human Palaeopathology.
10. Covington AD, Song L, Suparano O, Koon HE, Collins MJ (2010) Link-lock: the mechanism of stabilising collagen by chemical reactions. World Leather (25) 35-43.
11. Nelson DL, Cox MM, (2004) Principles of Biochemistry. W. H. Freeman(Ed), New York,.
12. Lewandowski CM (2015) Organic Mass Spectrometry in Art and Archaeology. The effects of brief mindfulness intervention on acute pain experience: An examination of individual difference (1).
13. Van Doorn NL. Wilson J, Hollund H, Soressi, Collins MJ (2012) Site-specific deamidation of glutamine: A new marker of bone collagen deterioration. Rapid Commun. Mass Spectrom (26) 2319–2327.
14. Gibson AM, Woods A (1997) Prehistoric Pottery for the Archaeologist. Leicester Univ Pr.

15. Fındık NÖ, Akyol AA, Sar N (2014) Archaeometric analyses of Hasankeyf unglazed ceramics. Mediterranean Archaeology and Archaeometry, (14) 261-271
16. Serpico M, White R (2000) Oil, fat and wax, in Ancient Egyptian Materials and Technology. P. Nicholson and I. Shaw (Eds), Cambridge University Press, Cambridge
17. Pollard AM, Heron C (1996) Archaeological chemistry. Royal Society of Chemistry, Cambridge
18. Evershed RP (2000) Modern analytical methods in art and archaeology. Ciliberto , Spoto(eds),Wiley, Canada.
19. Colombini MP, Modugno F, Ribechini E (2012) Archaeometric data from mass spectrometric analysis of organic materials: proteins, lipids, terpenoid resins, lignocellulosic polymers, and dyestuff. John Wiley & Sons, Inc.
20. Bonaduce I, Ribechini E, Modugno F, Colombini MP (2015) Analytical Approaches Based on GasChromatography Mass Spectrometry(GC/MS) to Study Organic Materials in Artworks and Archaeological Objects Topics in Current Chemistry (374) 1-37.
21. Stern B et al (2008) New investigations into the Uluburun resin cargo. J Archaeol Sci (35) 2188–2203.
22. McGovern PE et al (2013) Beginning of viniculture in France. Proc Natl Acad Sci (110) 10147–10152
23. McGovern PE et al (2004) Fermented beverages of pre-and proto-historic China. Proc Natl Acad Sci USA (101) 17593–17598.
24. Colombini MP, Modugno F, Ribechini E (2005) Organic mass spectrometry in archaeology:evidence for Brassicaceae seed oil in Egyptian ceramic lamps. Journal of Mass Spectrometry (40) 890–898.
25. Sandy DB (1989)The production and use of vegetable oils in Ptolemaic Egypt. Bulletin of theAmerican Society of Papyrologists (6).
26. Colombini MP, Modugno F, Silvano F, Onor M (2000)Characterization of the Balm of an Egyptian Mummy from the Seventh Century B.C. Studies in Conservation, (45) 19-29.
27. Lucas A (1962) Ancient Egyptian Materials and Industries, Edward Arnold (ed), London.
28. Rullkotter J, Nissenbaum A (1988) Dead Sea asphalt in Egyptian mummies: molecular evidence. Naturwissenschaften (75) 618-621.

29. Nissenbaum A (1992) Molecular archaeology: organic geochemistry of Egyptian mummies. Journal of Archeological Science (19) 1-6.
30. Ohary M (1982) Plants of the Bible. Cambridge University Press.
31. Mills JS,White R (1987)The Organic Chemistry of Museum Objects. Butterworths, London.
32. Clark KA, Ikram S, Evershed RP (2016) The significance of petroleum bitumen in ancient Egyptian mummies. The royal society.
33. Regert M, Rolando C (2002) Identification of archaeological adhesives using Direct Inlet Electron Ionization Mass Spectrometry. Analytical Chemistry (74) 965–975.
34. Evershed RP, Vaughan SJ, Dudd SN, Soles JS(1997) Fuel for thought? Beeswax in lamps and conical cups from late Minoan Crete. Antiquity (71) 979–985.
35. White R (1978)The application of gas-chromatography to the identification of waxes. Studies inConservation (23) 57–58.
36. Bayer HG (1982) The bust of Nefertiti, Analytical Chemistry. (54) 619–628.
37. Chen JCT, Amar AP, Levy ML, Apuzzo MLJ(1999) The development of anatomical art and sciences: The ceruplastica anatomical models of la Specola. Neurosurgery (45) 883-891.
38. Connan J, Dessort D (1991) Du bitume dans des baumes de momies e`gyptiennes (1295 av. J.C. - 300ap. J.C.): determination de son origine et e´valuation de sa quantite´. Comptes Rendue Academie des Sciences de Paris (312) 1445–1452.
39. Jones J, Higham TFG, Oldfield R, O'Connor TP, Buckley SA (2014) Evidence for Prehistoric Origins of Egyptian Mummification in Late Neolithic Burials. PLoS ONE (9) 1-13.
40. Clark AK, Ikram S, Evershed RP (2013) Organic chemistry of balms used in the preparation of pharaonic meat mummies. Pnas (110) 20392–20395.

3. Mass Spectrometry in Cultural Heritage Field

Mass spectrometry can be used to analyse a wide range of both organic and inorganic compounds. The main advantage of the technique is its capability of measuring masses with high accuracy, with high sensitivity. Mass spectrometry has been widely applied to a number of 'omics' fields, and is based on the production of gas phase ions and on their separation according to their mass-to-charge (*m/z*) ratio. (1)

It was introduced at the end of the nineteenth century and for a long time it was confined to the physical sciences, being used to study elementary particles of matter. In the 1940s, MS found applications in chemistry in the characterization of hydrocarbon. However, due to the lack of separation methods, the results were difficult to interpret. Owing to technological innovation and advances in instrumentation, mass spectrometers became more and more sophisticated, and versatile tools for different applications in chemistry. At the end of the 1970s, wide classes of volatile organic molecules with low to medium molecular weight could be characterized by this methodology. The introduction of soft ionization techniques marked the beginning of a new era for MS. In fact, they allowed MS to extend its applications to wide classes of nonvolatile, polar, thermally unstable and high molecular weight analytes. This opened up new horizons for MS in many unexpected fields, such as biology, biomedicine and biotechnology, in which this methodology had not previously found any possible application. (1)

The growing interest for the identification and characterization of polar and large compounds prompted the development and the introduction of new ionization techniques, such as electrospray ionization (ESI), (2) and matrix assisted laser desorption ionization (MALDI), (3) thus establishing new MS based approaches for studying large molecules, polymers and biopolymers, such as proteins, carbohydrates, nucleic acids. Parallel to the development of mass spectrometric instrumentation and methodologies, the improvements of separation techniques, such as gas chromatography (GC), high performance liquid chromatography (HPLC), and of their coupling with MS allowed the study of complex mixtures. Nowadays, MS plays an important role in many fields of scientific and technological research for identification, structural characterization and quantitative determination of wide classes of

compounds. Its unique capabilities, such as high sensitivity, high selectivity, accuracy in molecular weight determination, and ability to analyze complex mixtures, give this methodology a central role in the solution of problems that cannot be handled by other techniques. (1,4,5)

3.1 Features of Mass Spectrometry

MS has characteristic features, in particular:

Sensitivity.

MS is an extremely sensitive method. Nowadays, substances within a wide range of molecular weight can be measured [at the attomol (10_18 mol) or subattomol level]. (4,5)

Specificity and selectivity.

In most of cases, a mass spectrum is produced by just one molecule and a given molecule produces only one mass spectrum. It follows that a mass spectrum is unequivocally related to the chemical structure of the compound that produced it. (4,5)

Resolution

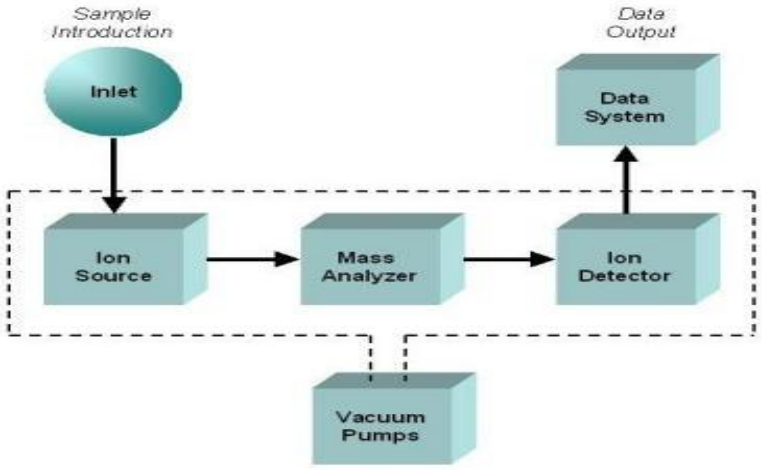
Resolution or resolving power is the ability of a mass spectrometer, and in particular of its analyzer system, to separate ions with different *m/z* ratios. When resolution refers to two peaks it can be defined by the Equation R=m/Δm where m is the nominal m/z value of the ions to be separated and Δm is their *m/z* difference calculated by considering the exact masses. Two peaks are considered to be resolved also in the presence of a valley, of a specified fraction of either peak height, between them. Resolution can be also determined by considering a single peak by using its width Δm at a given height. If the peak width is taken at 50% of its height, the resolution is defined as the full width at half maximum. The possibility to have well resolved peaks allows the measurement of the accurate mass of an ion, i.e. *m/z* values significant to four decimals, and to calculate its elemental composition. Of course, from the definition, resolution varies as a function of m. (4,5)

3.2 The Mass Spectrometer

The first part of every instrument is the inlet system that allows the introduction of the sample, generally molecules, into the mass spectrometer. There are different ways to introduce the sample, depending on its purity and properties. The ionization process occurs in the ion source by using an ionization technique. There are different ionization techniques. The ions thus produced are accelerated and driven from the source to the analyzer that separates them according to their *m/z* ratios. Finally, the detector reveals ions. All mass spectrometers have different stages of pumping in order to maintain the analyzer and detector

regions under high vacuum, i.e. 10^{-7}–10^{-8}torr or higher. Depending upon the ionization technique, the inlet system and the ion source must be/must not be/may be under vacuum. A schematic diagram of a mass spectrometer is shown in Figure 3.1.The mass spectrum is the result of a MS experiment. It is a plot of relative abundances of the ions produced inside the mass spectrometer against their m/z ratios. Once a mass spectrum has been obtained, it is possible to perform a library search, in different databases installed in the local computer or in remote servers through the Internet that can help in identification of unknowns. (1)

Figure 3.1: Diagram of a mass spectrometer.



3.2.1 Inlet System and Sample Introduction in a Mass Spectrometer

Generally, it is not possible to introduce a material, as it is, into a mass spectrometer: an appropriate sample preparation procedure, basically consisting of analytes extraction from the substrates and in their preliminary purification, precedes the mass spectrometric analysis. There are several devices for introducing a sample into a mass spectrometer. Their choice is based on the purity and chemico-physical properties of the sample. The simplest way is the introduction of a sample directly in the ion source through a direct inlet whose design depends upon the ionization technique used. If different analytes are introduced directly into the mass spectrometer at the same time the result will be a mass spectrum formed by the superimposition of each mass spectrum of each component of the mixture. This will be not easily interpretable and will be meaningless when the goal is to identify, characterize and quantify analytes that are present in complex mixtures. This will be the case in the analysis of samples from cultural heritage materials, such as ceramics, paintings, monuments, etc. In these cases, separation methods are definitively required prior the mass analysis so that components can arrive in the ion source in a fractionated way. It follows that MS must be

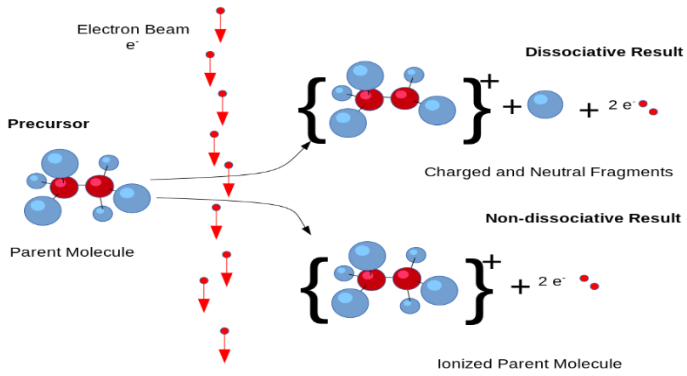
coupled to a separation technique for analysis of mixtures. This coupling is generally on-line and it allows the combination of complementary features: high sensitivity, selectivity and specificity offered by MS with high performance in separation. (1,4,5)

3.2.2 The Ion Source and Ionization Techniques

Through a direct inlet ora separation technique the sample is introduced into the ion source (the region in which the ionization occurs, i.e. the mass spectrometer region where a molecule is transformed into an ion). The design of the ion source is dependent on the ionization technique. A wide variety of ionization techniques is nowadays available and one can choose the most suitable on the basis of different factors, such as the chemico-physical properties of the analyte under investigation, its molecular weight, polarity, etc. Similarly, molecules can be ionized by the formation of negative ions due to single or multiple proton abstraction. *Electron Ionization* (EI) and *Chemical Ionization* (CI) are the most commonly used ionization techniques for the analysis of volatile, nonpolar, low molecular weight (up to 700–900 Da) and thermally stable compounds. These techniques can be also useful for studying non highly volatile molecules, but in these cases, preliminary chemical reactions, such as derivatization, have to be performed in order to obtain derivatives with increased volatility. The introduction and development of more recent ionization techniques, such as *Electrospray Ionization* (ESI) and *Matrix Assisted Laser Desorption Ionization* (MALDI), have allowed the study of nonvolatile and thermally unstable compounds. (4,5)

Electron Ionization: EI is probably the most frequently used ionization technique in the field of cultural heritage analyses. In EI the ionization of a molecule is produced by electrons. The EI process might be simplified as follows: an electron beam with an appropriate energy passes in proximity to the gas phase molecules and transfers part of its energy to them. If the transferred energy is high enough, an electron is extruded from the molecule. The EI source is formed by a tungsten or rhenium filament, an anode, an ion repeller, a magnet, and a series of lenses for extracting, focusing and accelerating the ions formed (Figure 3.2).

Figure 3.2: Scheme of Electron Ionization (EI) source.



The EI technique is well established, sensitive and suitable for qualitative and quantitative analysis of organic molecules. (1)

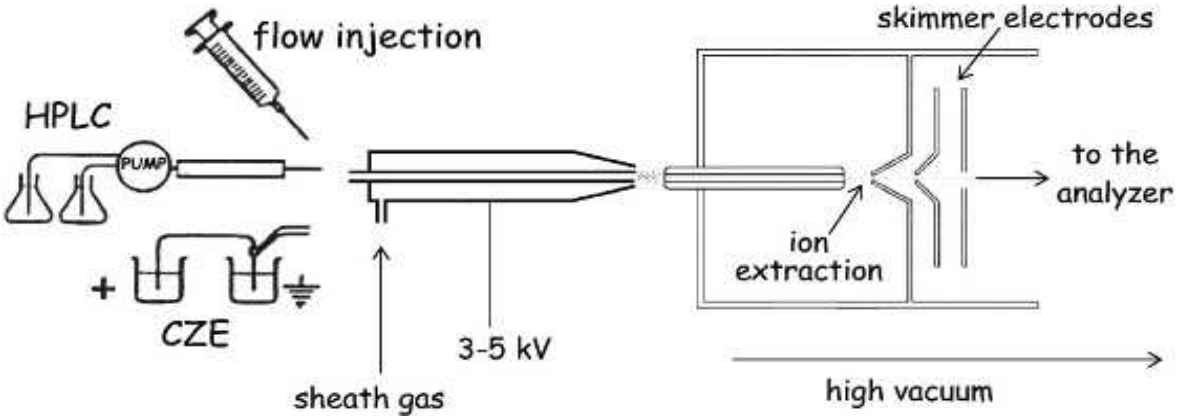
'Soft' Ionization Techniques: The term soft is opposite to hard and both these terms are used to refer to the amount of energy deposited onto the molecule during the ionization process. In the case of EI (a large amount of energy is deposited onto the molecular ion. It causes abundant fragmentation and structural information can be obtained. However, if the amount of excess energy is too large, the molecular ion might be undetectable thus not allowing molecular weight determination of the analyte. In soft ionization methods the excess energy deposited onto the ionized molecule is very small and stable even-electron ions are formed. This leads to easy determination of the molecular weight of the analyte, but as fragmentation is absent or it occurs to a very low extent, structural information is missing in the mass spectrum. However, one can obtain structural information by causing ion fragmentation out of the source by means of tandem mass spectrometry experiments (see below). Owing to soft conditions, the mass spectra obtained by this kind of ionization techniques are also characterized by the presence of adduct ions, i.e. ionic species formed by weak interactions between the ions and other chemical species. Soft ionization methods could be classified into different groups among which spray ionization techniques and desorption ionization techniques. (4,5)

Spray Ionization Techniques: Spray ionization techniques operate at atmospheric pressure: a solution spraying from a capillary is ionized at atmospheric pressure and the ions produced are driven into the high vacuum of the mass analyzer where they are separated. The use of spray and desorption ionizations does not require volatilization of the sample before ionization. This means that all these techniques can ionize nonvolatile, polar and large to very

large compounds. The most important techniques belonging to this class are electrospray ionization (ESI), atmospheric pressure chemical ionization. (4,5)

The term electrospray reveals important features of the process: first of all, the ionization occurs on a spray. This is produced by forcing a liquid to pass through a capillary. The first part of the ionization's term contains further important information: an electric field (3–5 kV) is applied to the capillary and it causes ionization. The overall effect is the formation of small liquid-charged droplets from which gas phase ions are formed. At the exit of the capillary, the liquid assumes a characteristic shape, called a Taylor cone, from which the jet of charged particles emanates. During their travel from the exit capillary to the entrance of the high vacuum region, that are 1–3 cm apart, the charged droplets, each containing ions and solvent molecules, undergo evaporation of the solvent with a consequent reduction of their size and increase of the field at their surface. This causes a 'Coulomb explosion', due to high charge density, and ion escape occurs. The ions so generated are transferred to the high vacuum region of the mass spectrometer and are driven to the mass analyzer. ESI can be used to produce protonation or deprotonation of the analyte. A requisite of a molecule to be ionized by ESI is its polarity: nonpolar compounds cannot be ionized by this technique. ESI mass spectra are characterized by protonated/deprotonated molecules and adduct ions. A key feature of the electrospray process is the formation of multiply charged ions. An empirical rule states that a molecule generally takes a charge every thousand units of molecular weight. This allows the measurement of very high molecular weight molecules at low *m/z* values. (4,5)

Figure 3.3: Electrospray Ion (ESI) source.

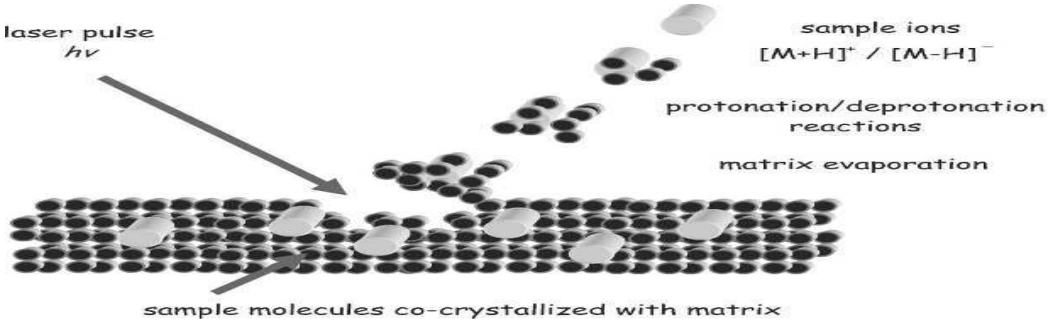


Desorption Ionization Techniques: The term 'desorption ionization' indicates those ionization techniques in which the production of ions is based on a desorption process. This consists of the rapid addition of energy to a sample in a condensed phase (i.e. liquid or solid) with subsequent production and emission of stable ions in the gas phase. These are generally even electron species that fragment only to a limited extent. The development of desorption methods has amplified the impact and utility of MS in a lot of fields, such as biology, biochemistry and proteomics. Over the years, a lot of desorption ionization techniques have been introduced to MS, most of them are actually no longer used. (4,5)

During the development of MS, a lot of studies have been devoted to the use of laser light as an energy source for ionizing molecules. As a result, in the mid 1980sMALDI (3) was introduced and soon applied to the study of large molecules. (6)

The MALDI source may operate under high vacuum or under atmospheric pressure. A lot of features of MALDI are conveyed by its name: it is a desorption ionization, produced by a laser beam, and assisted by a matrix.To generate gas phase, protonated molecules, a large excess of matrix material is co-crystallized with analyte molecules. The resulting solid is then irradiated by nanosecond laser pulses, usually from small nitrogen lasers with a wavelength of 337 nm. The matrix is typically a small organic molecule with absorbance at the wavelength of the laser employed and the matrices differ from each other in the amount of energy they impart to the biomolecules during desorption and ionization and hence the degree of fragmentation (unimolecular decay) that they cause. Generation of ions is believed to arise through ion/molecule reactions in the gas phase. Generally, $[M+H]^+$ ions are preferentially formed in the positive ion mode, and $[M-H]^-$ ions in the negative ion mode. The presence of singly charged signals in MALDI spectra allows to correlate each peak one-to-one with analytes, thus allowing the analysis of complex mixtures and big molecules (like whole proteins). (4)

Figure 3.4: Schematic view of MALDI.



3.2.3 The Analyzer

Once the ions are formed in the ion source, they are accelerated towards the mass analyzer where separation according to their *m/z* ratio occurs. Before describing the main analyzers, it is better to consider one feature common to all of them: the resolution. The choice of an analyzer depends on the kind of studies and applications that will be carried out: *m/z* range, resolution, mass accuracy, scan speed, number of mass separations, are some of the parameters that have to be considered. (1,4)

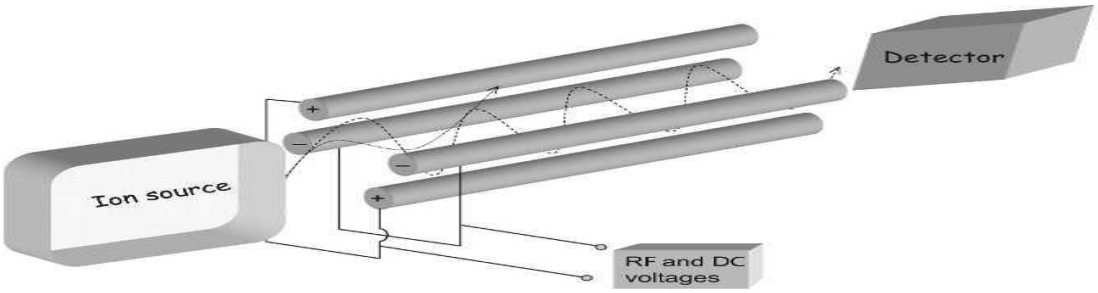
There are four basic types of mass analyser currently used in proteomics research. These are the quadrupole (Q), ion trap (IT), time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR) analysers. (7) They are very different in design and performance, but also for their resolution, sensitivity and accuracy and for the types of experiments that are able to perform. These analysers can be stand alone or, in some cases, put together in tandem to take advantage of the strengths of each.

The Quadrupole Mass Filter

The quadrupole mass filter (Q, Figure 3.5) is composed of four parallel cylindrical, or ideally hyperbolic, rods to which a direct current (DC) and a radiofrequency (RF) are applied. The separation of ions according to their *m/z* ratio occurs by scanning simultaneously DC and RF, but maintaining their ratio to be constant. For a given value of DC and RF only ions with a given *m/z* value will be able to pass through the entire quadrupole and they will be transmitted to the detector. By changing DC and RF another family of resonant ions will be transmitted and so on. A complete scan allows all the *m/z* values in a given range to be analysed. (1,4)

Commercially available instruments usually have mass/charge limits ranging from 0 to 4000 *m/z* and at best are normally set to resolve the various ^{13}C isotope peaks for a singly charged ion, although the resolution may be intentionally degraded to improve sensitivity. In ESI, multiple charging enables quadrupole mass measurement of molecules >100,000 Da, if the molecule can be charged sufficiently. (1,4,5)

Figure 3.5: The quadrupole analyzer.



Time of Flight Analyzer

TOF analyser working under high vacuum (10^{-8}-10^{-9} atm) and ions can be separated during their flight in a field-free drift region basing upon the time they use to pass through the tube. Their different velocities depend on their *m/z* ratios. All ions entering the TOF tube have a fixed kinetic energy, which is proportional to the applied accelerating voltage and the charge. This implicates that the higher is the mass of the ion the lower is its velocity and the longer it takes before the ion reaches the detector. A detector at the end of the tube records the time of flight of the ions.

Although the first generation of TOF analyzers were formed by just a drift tube and had low resolution, many improvements have been made over the years. In particular, the introduction of a 'reflectron', located at the end of the flight tube, improved theTOF performances a lot. In fact, once accelerated in the source, ions with the same *m/z* value do not have exactly the same kinetic energy. Thus they arrive at the detector at different times, causing peak broadening and a consequent low resolution.

The reflectron allows ions to slow down and reverse their flight path to a second detector. Ions with lower kinetic energy do not penetrate the reflectron as deep and thus turn around faster, catching up with ions of slightly greater kinetic energy that penetrate the reflectron deeper. Thereby the flight times of ions with identical *m/z* values, but different kinetic energy values, will be corrected when the ions arrive to the detector. Commercial TOF based instruments have high resolution of about 10,000 or greater and high accuracy characteristics. (1,4)

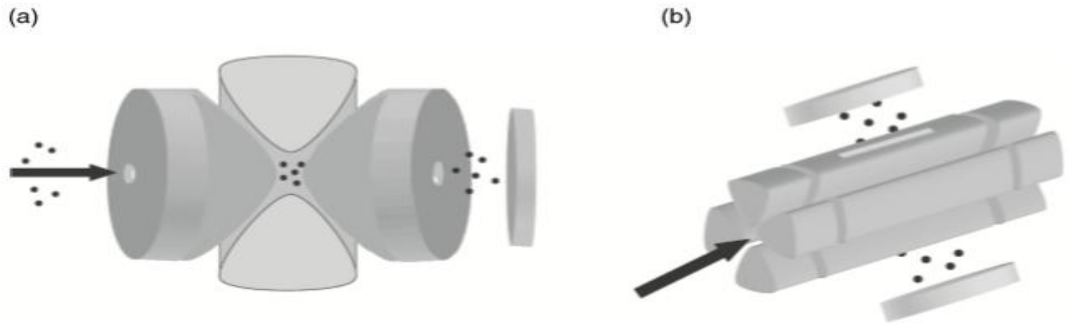
TOF analysers are typically used in combination with MALDI (MALDI-TOF MS instruments); however, both MALDI sources and TOF analysers can be used in different configurations. (4)

Three- and Two-Dimensional Ion Traps

The quadrupole ion trap (QIT) is based on the same principle as the quadrupole mass filter, except that the quadrupole field is generated within a three-dimensional trap.

According to the movement of ions inside them, ion traps can be divided into two groups: three- and two-dimensional (linear) ion traps (Figure 3.6). In the former, ions, that are confined and maintained near the centre of the ion trap, have complex movements in three dimensions. However, in linear ion traps, ion movements are almost in a two-dimensional plane, like a snake on land. A three-dimensional ion trap is formed by three electrodes, two end caps and a ring electrode; in the linear ion trap different electrodes form each of the four edges. The ions introduced into the trap are maintained on stable orbits applying DC and RF potential to the electrodes. To separate ions, according to their *m/z*, a RF scan is made. While the RF amplitude increases, ions with higher *m/z* values are destabilized and they leave the ion trap moving towards the detector. Ion trap mass spectrometers are widely used and can be easily interfaced with all the ionization techniques. One of the most attractive aspects of ion trap mass spectrometers is represented by the possibility of carrying out multi mass separation steps with only one analyzer. (1,4)

Fig 3.6: Three-dimensional (a) and linear (b) ion trap mass analyzers. Arrows indicate the entrance of ions from the source. Detectors are depicted as plates.



The advantages of the ion-trap mass spectrometer include compact size and the ability to trap and accumulate ions (typically the ion trap can hold up about 105–106 ions before coulombic repulsions significantly affect their trajectories and reduce the mass resolution) to increase the signal-to-noise ratio of a measurement. Mass spectrometer equipped with QIT technology has good sensitivity but low resolution and mass accuracy. Due to its small trapping volume, QIT has a limited trap capacity; when the number of ions in the trap is too high, we assist to deterioration in the mass spectrum and loss of dynamic range due to space-charging

interactions. To avoid these effects, the number of ions introduced into the trap has to be strictly controlled. (8)

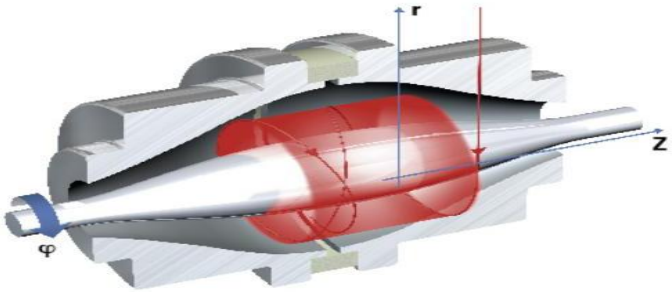
Analyzers with Fourier Transform Treatment of the Signal

FT-ICR and the Orbitrap belong to this group. In these analyzers the *m/z* values of the ions are not directly measured, but they are obtained by Fourier transform treatment of the signal. (9)

The Orbitrap analyzer, was invented by Alexander Makarov. Its name recalls the concept of trapping ions. Indeed, ions are trapped in an electrostatic field produced by two electrodes: a central spindle-shaped and an outer barrel-like electrode. Ions are moving in harmonic, complex spiral-like movements around the central electrode while shuttling back and forth over its long axis in harmonic motion with frequencies dependent only on their *m/z* values. The latter can be determined by Fourier transform treatment of the signal. Note that, in contrast to other trapping analyzers, neither RF nor a magnet field are used. This means that the Orbitrap is simple to use and easily maintained. The Orbitrap allows very high resolution to be achieved (the resolving power in commercial instruments is 100000, rivalling that of FT-ICR instruments) and routine mass measurement accuracies less than 2 ppm. It finds applications in many fields, such as biology, proteomics, food chemistry and cultural heritage. (9)

The FT-ICR analyzer is based on the interaction of a charged species with an external magnetic field produced by superconducting magnets. A charged species subject to a magnetic field covers a circular trajectory. The rotational frequency of a certain ion on this orbit is called the ‘natural cyclotronic frequency’ and it is dependent on the magnetic field and on the *m/z* ratio of the ion. An additional RF is applied and the net result is a spiral motion. To obtain the mass spectrum the RF is scanned while the magnetic field is maintained constant. In FT-ICR mass spectrometers, ion isolation and detection occur in the same region. In fact all ions coming from the source are simultaneously excited applying a RF pulse of alarge wide band. The resulting signal is then converted by Fourier transform to obtain the mass spectrum. The main feature of this analyzer isits extremelyhigh resolution.However, owing to the difficulty of operation and its very high cost, the FT-ICR analyzer cannot be used for routine purposes. (9)

Figure 3.7: Orbitrap mass analyser.



3.2.4 Detector

Once ions have been produced and analysed they must be detected. Indeed, the detector is the final part of a mass spectrometer. At the very sunrise of MS, detectors were composed of a fluorescent screen or a photographic plate; the modern instruments are equipped with detectors able to transform the signal produced by the ion beam into an electric current that is transmitted to the data system. At the moment, there is a variety of different types of detectors whose choice is determined by the instrument design and applications. (10)

3.3 Tandem Mass Spectrometry

Fragmentation of a molecule, after ionization in the mass spectrometer, can be used to obtain structural information. The advent of tandem mass spectrometry (MS/MS) made it possible to analyze complex peptide mixtures without the need to separate the individual components, because a mass window narrow enough to exclude all but one of the mixture components can be specified in the first of the two mass spectrometers of a tandem instrument configuration. In fragmentation mode, the precursor ion selected, upon dissociation, most commonly induced by collision with an inert gas (collision- induced dissociation, CID, also referred to as collisionally activated dissociation, CAD), produces a product ions spectrum in which all fragment ions are derived from the selected precursor ion (Figure 3.8). Thus, it is possible to analyze and obtain structural information on peptides that can be read to give the sequence of the peptides. (5,11)

Whereas it is very easy to acquire MS/MS data for a single peptide, the interpretation of such data is far from trivial. (11

Peptides consist of amino acid residues linked together by amide bonds. Under CID, peptides fragment mostly along the peptide backbond, often with transfer of one or two hydrogens to create stable ion structures. Two main nomenclature systems for peptide fragments have been

proposed, both utilizing the letters of the alphabet to label ions. N-terminal fragment ions (those containing the N-terminus of the peptide) are labeled with the first few letters of the alphabet (a, b, c), whereas C-terminal fragment ions (which contain the C-terminus of the peptide) are labeled with the last few letters of the alphabet (x,y,z). For some of the fragment ions, hydrogen transfers from other amino acids occuras indicated in Figure 3.8. CID fragment ions are numbered sequentially, beginning with the C-terminal amino acid. Thus for a tetrapeptide, a cleavage of the bond between the carbonyl carbon of the second amino acid of the peptide and the amine nitrogen of the third amino acid would generate two fragments. For a singly protonated precursor ion, if the charge were retained within the first two amino acids, then a b, ion would be produced along with a neutral fragment; the latter, containing the C-terminus of the peptide, is not detectable with the mass spectrometer. Conversely, if the charge were retained within the last two amino acids of the tetrapeptide, then the above fragmentation would yield a y2 fragment ion and a neutral species containing the peptide N-terminus. A multiply protonated precursor could, depending on where the charges are localized along the peptide chain, yield two charged fragments from a single bond cleavage. A significant difference between high and low energy CID is that the former process generates all the types of the fragment ions described in Figure 3.8, although not necessarily in every spectrum, because the fragmentation process and the type of fragment ions produced are also affected by the type of amino acids present in each peptide. In low energy CID spectra, c, x, z, and a-type fragment ions are not commonly observed. (11)

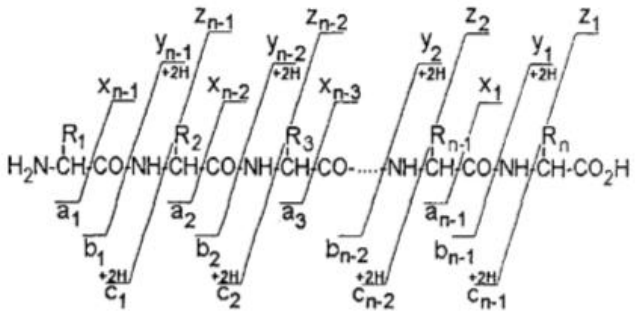


Figure 3.8: Nomenclature for peptide CID fragment ions. Typically, all of the above fragment ion types are observed in high-energy CID tandem mass spectra (although usually not in the same spectrum), but only b, y, and, less frequently, a and z, fragments are observed in low-energy CID spectra.

3.4 Quadrupole-time of flight (QqTof) mass spectrometer

QqTOF tandem mass spectrometer can be described as a triple quadrupole with the last quadrupole section replaced by a TOF analyser. It has rapidly been embraced as powerful and robust instrument with unique capabilities. QqTOF combines the high performance of TOF devices, in both the MS and MS/MS modes, with the widely used techniques of electrospray ionization. These mass spectrometers are characterized by high sensitivity, high mass accuracy and high mass resolution for both precursor and product ions. (5,12)

The instrument consists of three quadrupoles, Q0, Q1 and q2, followed by a reflecting TOF mass analyzer with orthogonal injection of ions.

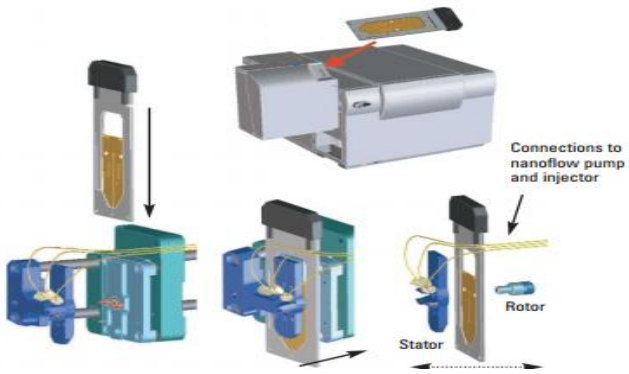
The resulting spectra benefit from the high resolution and mass accuracy of the TOF instruments, and also from their ability to record all ions in parallel, without scanning. Both Q0 and q2 are operated in the RF only mode: the RF field creates a potential that provides radial confinement of the precursor and/or fragment ions. Since the RF quadrupoles are normally operated at a pressure of several millitorr, they provide both radial and axial collisional damping of ion motion. Ions are thermalized in collisions with neutral gas molecules, reducing both the energy spread and the beam diameter and resulting in better transmission into and through both the Q (13) and TOF (14) analysers. For MS/MS, Q1 is operated in the mass filter mode to transmit only the parent ion of interest. The ion is then accelerated to an energy of between 20 and 200 eV before it enters the collision cell q2, where it undergoes collision induced dissociation (CID) with neutral gas molecules. The resulting fragment ions are focused by RF fields and re-accelerated to the required energy (usually several tens of eV per unit charge), and focused by ion optics into a parallel beam that continuously enters the ion modulator of the TOF analyzer. Initially the modulator region is field-free, so ions continue to move in their original direction in the gap. A pulsed electric field is applied at a frequency of several kHz across the modulator gap, pushing ions in a direction orthogonal to their original trajectory into the accelerating column, where they acquire their final energy of several keV per charge. From the accelerating column, ions arrive in the field-free drift space, where TOF mass separation occurs. The ratio of velocities (or energies) in the two orthogonal directions is selected such that ions reach the ion mirror and then the TOF detector naturally, without requiring an additional deflection in the drift region, which could affect the mass resolution. (15)

The ESI Q-TOF mass spectrometer generally can also be equipped with nano-HPLC-CHIP, a different and innovative technology compared to the classic LC-MS systems. In classical

LC/MS instruments a precolumn (o enrichment column) where the samples is washed from salts and detergents and concentrated composes the liquid chromatography. After the reverse-phase separation, another post-column capillary transports the analyte to the electrospray. The passage through all the column and the volume of elution in the capillary volume involves a broadening of chromatographic peaks and the decrease of resolution that affect the sensitivity of more dilute species. (16,17)

The HPLC-Chip is made of sandwiched polyimide films. In particular, the laser ablation is used to generate grooves of specific dimensions which can be filled with packing material, and used as pre-column, reverse phase separation nanocolumn and nanospray tip. In this way diffusion in dead volumes cannot affect the broadening of peaks and the sensitivity is strongly enhanced. One end of the chip tapes off into a polyimide nanospray emitter (tip). A stator and a rotor fix the chip when it is placed into the HPLC-Chip Cube and the rotation of the latter allows the switch between loading and analysis (Figure 3.9). (16,17)

Figure 3.9 The HPLC-Chip Cube automatically loads the chip, establishes leak tight connections and positions the chip orthogonal to the MS inlet simply by clicking on the operate command in the ChemStation menu.
(http://www.chem.agilent.com/cag/EMEA/LCMS_Literature/Brochures/5989-5492EN.pdf)



During loading, the sample is transferred at a flow rate of 4 µl/min onto the enrichment column. Peptides are retained at the front end of the column while the loading solvent is directed to waste. During analysis, the outlet end of the enrichment column is connected to the analytical column; as the nanoflow gradient starts (400 nl/min), the retained peptides are eluted onto the analytical column where they are further separated by reversed phase chromatography. The peptides emerging from the tip enter the spray chamber as an aerosol thanks to a nebulizing gas. Droplets are led by an electric field as described before for an ESI source. (16,17)

3.5 MRM

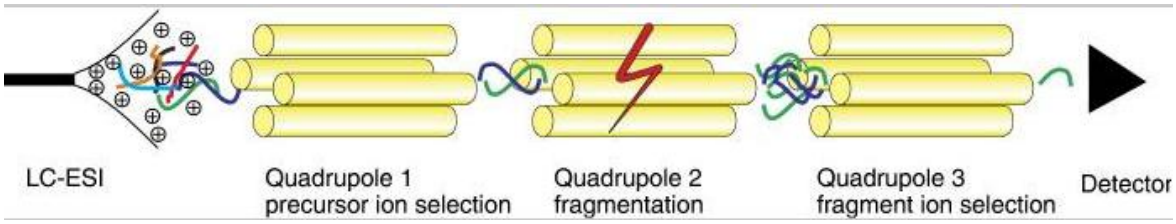
There are several emerging technologies that show great promise to handle ancient samples. (18)

Multiple reaction monitoring (MRM) is one of them and it ideally complements the discovery capabilities of proteomics by its unique potential of reliable detecting and quantification of analytes of low abundance in complex mixtures. (18)

In an MRM experiment, a predefined precursor ion and some of its fragments are selected by the two mass filters of a triple quadrupole instrument and monitored over time for precise quantification. In MRM the first and the third quadrupoles act as filters to specifically select predefined *m/z* values corresponding to the peptide ion and different specific fragments ion of the peptide, whereas the second quadrupole serves as collision cell. (19)

Despite the high specificity of MRM analysis, obtained by the two filters of mass of the QQQ, a particular transition (precursor/product) may not be specific to a molecule. In fact, in a complex mixture, non-specific signals can be derived from other molecules with transitions of similar masses, thus affecting the validity of the analysis. The acquisition of two or more parallel transitions to the same molecule (precursor → product1, product2) gives the analysis further selectivity. In fact, the typical elution time of the analyte in question, these transitions will produce a set of peaks "co-eluted" only if derived from the same molecule. Therefore, the acquisition of multiple transitions in parallel is most important in the MRM method set up. The total number of transitions per stroke LC-MS/MS is limited; commonly, only the best 2-4 transitions for each molecule are selected for quantitative analyzes. (19)

Figure 3.10: MRM analysis on QQQ MS. Several analytes are coeluting from the chromatographic system. The specific *m/z* selection in the first quadrupole filters out most coeluting ions. However, owing to identical mass, one interfering ion (blue) remains. In quadrupole 2, the analytes are fragmented. The *m/z* selection in the third quadrupole filters out all the fragments of the blue analyte and leaves only a set of fragments of the green analyte for specific detection.



3.6 Proteomics approach in Archaeology

The term “proteomics” was first coined in 1995 and was defined as the large-scale characterization of the entire protein complement of a cell line, tissue, or organism. Using the definition of proteomics, many different areas of study are now grouped under the rubric of

proteomics. These include protein-protein interaction studies, protein modifications, protein function, and protein localization studies to name a few. (1)

These ambitious goals will certainly require the involvement of a large number of different disciplines such as molecular biology, biochemistry, and bioinformatics. It is likely that in bioinformatics alone, more powerful computers will have to be devised to organize the immense amount of information generated from these endeavours.(20)

Although proteomics is typically considered to be associated with the study of living organism, its inherently multidisciplinary nature has recently led to the application of proteomic methods to oddly assorted areas; in particular, it has been proven to be an effective tool also for the scientific analysis of artworks. Firstly developed for the identification of proteins in painting binders, nowadays, these studies are not limited anymore to the identification of binders in paintings or artworks but have also been extended to archaeological remains and to the characterization of ancient proteins in what in now called paleoproteomics. (21)

Protocols routinely applied for typical modern samples still need to be fully adapted to take into account the low amount of proteinaceous material, the heterogeneity and the unusual physical state of the samples, as well as the high levels of damage found in ancient samples. Although analysis in proteomics are *per se* invasive, modern mass spectrometry instrumentation enable the characterization of proteins with extremely high sensitivity even in crude mixtures in which the dynamic range of components abundance exceeds 1000-fold and on very limited amount of sample, typically less than 10 μg. Modern instruments are therefore perfectly adequate to afford the minimal quantities of ancient samples. However, all the steps of the proteomic procedure need to be thought fully adapted, from the optimization of specific protocols for sample preparation to development of data analysis tools that can cope with ancient, damaged, samples. Moreover, although for merely identifying purposes, the detection of as few as two peptides is sufficient to properly pinpoint the protein, the characterization of the modifications induced by ageing and deterioration processes requires a deeper examination of most of the protein primary structure. (21)

The proteomic approach, used for archaeological objects, is composed of several steps including the hydrolysis of the proteins, the MS-based analysis of the resulting peptides, and data processing using bioinformatics tools.Tandem mass spectrometry (MS/MS) allows a deep structural approach with the accurate identification of the peptide backbone sequence

and its modifications. Thus, protein identification can be accurately performed on the basis of a few peptidic sequences. MS/MS allows the characterization of the amino acid sequences of peptides, providing thus an accurate identification of proteins. Additional information can also be obtained such as the biological origin of the proteins or information on sample degradation state or aging.The common strategy to interpreter tandem mass spectrometry data is matching the observed spectra with predicted mass values rather than the actual determination of peptide sequence. These peptides are predicted by theoretical tryptic cleavage of known protein sequences, as found in protein databases such as UniProt (http://www.uniprot.org/) or National Centre for Biotechnology Information (NCBInr: http://www.ncbi.nlm.nih.gov/protein/) consulted by dedicated search programs such as Mascot MS/MS Ions Search (Matrix Science). (22,23)

3.7 Aim of thesis

Recently, proteomics strategies and mass spectrometry techniques have been widely used in the field of cultural heritage with extremely promising results. Nowadays, these studies are not limited anymore to the identification of binders in paintings or artworks but have also been extended to archaeological remains.

My PhD project was aimed at the development and the optimization of biochemical and proteomic strategies based on advanced mass spectrometry techniques for archaeological remains.

The project was focused on the identification and molecular characterization (by looking at chemical modifications) of proteins and small molecules in archaeological finds like amphorae, tools and bones after the optimization of analytical protocols and mass spectrometry approaches.

Moreover, the study was focused on the study of aging and degradation processes, in the search of defining the molecular markers of conservation state. Moreover, the project was aimed also to the characterization of embalming materials used in the animal mummification process.

C**hapter 4** describes the development and the adaptation of classical proteomic strategies to different type of archaeological samples. The aim is to show how different analytical protocols can be used to obtaining desirable results. The use of specific samples pre-treatments and data analysis tools can indeed be crucial since it improves the quality of the experimental data that allow highlighting specific aspects of both identification and characterization.

C**hapter 5** was focused on the study of the degradation process of the proteins, at the molecular level, and was aimed at paving the way for in-depth paleo-proteomic characterization beneficiating from the advent of the latest proteomic platforms. The inspiring idea of this chapter is that different environmental conditions might be reflected in the nature and extent of chemical modifications in proteins. The difference in the level of deamidation, for instance, can constitute an important molecular marker.

C**hapter 6** was aimed on the characterization of a textile membrane thread, covered by wire, from a drape found in the tomb of Emperor Henry VII of Luxembourg. The techniques of productions of these membranaceous threads are not well known. These kind of studies hasn’t been made yet and today, the analysis of the organic materials in membranaceous

threads was limited to its classification in one of three categories: skin, parchment and animal gut, and it is based on its macroscopic appearance. Herein, proteomic apphoach was used to define the organism from which the connective tissue was used to prepare the membranaceous textile. The sample were provided by Dr. Ilaria Degano, Department of Chemistry and Industrial Chemistry, University of Pisa.

In C**hapter 7,** specific experimental procedures, based on mass spectrometric technologies have been developed to analyze the traces of different organic compounds that may be present in archaeological potsherds, with the aim of establishing a relationship between the molecules contained and the use of crockery and tools.

C**hapter 8** describes an analytical procedure for the extraction and identification of the main natural organic compounds present in animal mummy balms. Animal mummification was an enormous part of Egyptian culture. Indeed, Egyptians treated animals with great respect, regarding them both as domestic pets and representatives of the gods. Today, characterized animal embalming materials is a challenge and on this way, this chapter aims to focus on the characterization of the materials used in the mummification processes of 4 cats mummies, from the Museum "Umberto Scerrato", Napoli.

Chapter 9 deals with the first example of the application of targeted MS method (Multiple Reaction Monitoring, MRM) to samples in the field of cultural heritage.

Despite the tremendous success of standard based proteomics approaches, the demand of improvement in sensitivity of protein detection in highly complex and rather contaminated samples drove the development of a targeted assay, that was, in a first instance set up and applied to identify proteinaceous binders in paintings.

3.8 References

1. Colombini MP, Modugno F (2009) Organic mass spectrometry in art and archaeology. Wiley, New York
2. Yamashita M, Fenn JB (1984) Electrospray ion source. Another variation on the free-jet theme. J. Phys. Chem (88) 4451–4459.
3. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal. Chem. (60) 2299–2301.
4. Mann M,Hendrickson RC, Pandey A (2001) Analysis of proteins and proteomes by mass spectrometry. Annu. Rev. Biochem (70) 437-473.
5. Griffiths WJ, Jonsson AP, Liu S, Rai DK,Wang Y (2001) Electrospray and tandem mass spectrometry in biochemistry. Biochem J. (355) 545-561.
6. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, Matsuo T(1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. RapidCommun. Mass Spectrom. (2) 151–153.
7. Wysocki VH, Resing KA, ZhangQ, Cheng G (2005) Mass spectrometry of peptides and proteins. Methods (35) 211–222.
8. Payne AH, Glish(2005). GL Tandem mass spectrometry in quadrupole ion trap and ion cyclotron resonance mass spectrometers. Methods Enzymol (402) 109–148.
9. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG (2005) The Orbitrap: a new mass spectrometer. J. Mass Spectrom. (40) 430–443.
10. Koppenaal DW, Barinaga CJ, Denton MB, Sperline RP, Hieftje GM, Schilling GD, Andrade FJ, Barnes JH (2005)MS detectors. Anal. Chem., (77) 418–427.
11. Papayannopoulos IA, (1995) The Interpretation of Collision-Induced Dissociation Tandem Mass Spectra of Peptides. Mass Spectrometry Reviews, (14) 49-73.
12. Loboda V, Krutchinsky N, Bromirski M, Ens W, Standing, KG (2000) A tandem quadrupole / time-of-flight mass spectrometer (QqTOF) with a MALDI source : design and performance. Rapid Commun. Mass Spectrom.
13. Douglas DJ, French JB(1992) Collisional focusing effects in radio frequency quadrupoles. J. Am. Soc. Mass Spectrom. (3) 398–408.
14. Krutchinsky A, Chernushevich I, Spicer V, Ens W, Standing K (1998) Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer. J. Am. Soc. Mass Spectrom. (9) 569.

15. Guilhaus M, Selby D, Mlynski V. (2000) Orthogonal acceleration time-of-flight mass spectrometry. Mass Spectrom. Rev. (19) 65–107.
16. Brennen RA,Yin H,Killen KP (2007) Microfluidic gradient formation for nanoflow chip LC. Anal.Chem. (79) 9302-9309
17. Vollmer M, Van de Goor T (2007) HPLC-Chip/MS technology in Proteomic Profiling. Wiley-VCH
18. Elliott M, Smith D, Kuzyk M, Parker CE, Borchers CH (2009) Recent trends in quantitative proteomics. J Mass Spectrom. (44) 1637–1660).
19. Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol. Syst. Biol. (4) 222.
20. Graves PR, Haystead T J (2002) Molecular Biologist ' s Guide to Proteomics. Microbiol. Mol. Biol. Rev (66) 39–63.
21. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.
22. Corso G. et al. (2012) Characterization of pigments and ligands in a wall painting fragment from Liternum archaeological park (Italy). J. Sep. Sci. (35) 2986–2993.
23. Fremout W. et al. (2010) Tryptic peptide analysis of protein binders in works of art by liquid chromatography-tandem mass spectrometry. Anal. Chim. Acta (658) 156–162.

4.Proteomic strategies for cultural heritage: From bones to paintings

Nowadays the knowledge of the materials used in a work of art is of greatest significance, not only to give an insight in the historical context of objects and artists, but also to analyze degradation processes taking place in aged objects and to develop appropriate conservation and/or restoration treatments. (1)

Proteomics for cultural heritage, i.e. the identification of proteinaceous material used by artists in their masterpieces and found in archaeological remains, is still in its infancy, with the first paper dating back to the early 2000. (2) In particular, protocols routinely applied for typical modern samples still need to be fully adapted to heterogeneous materials in ancient samples of several cultural heritage objects. For this purpose is important to take into account the low amount of proteinaceous material, the heterogeneity and the unusual physical state of the samples, as well as the high levels of damage found in ancient samples.

Although analyses in proteomics are per se invasive, modern mass spectrometry instrumentations enable the characterization of proteins with extremely high sensitivity even in crude mixtures in which the dynamic range of components abundance exceeds 1000-fold, and on very limited amount of sample, typically less than 10 μg. Modern instruments are therefore perfectly adequate to afford the minimal quantities of ancient samples. However, all the steps of the proteomic procedure need to be thoughtfully adapted, from the optimization of specific protocols for sample preparation to the development of data analysis tools that can cope with ancient, damaged samples.

Moreover, although for merely identifying purposes the detection of as few as two peptides is sufficient to properly pinpoint the protein, the characterization of the modifications induced by ageing and deterioration processes requires a deeper examination of most of the protein primary structure.

The very critical step in the application of proteomic approach is related to the efficiency of either protein extraction from solid matrices or proteolytic digestion of substrates incorporated within the matrix itself. (3)

This chapter deals with some examples of the development and adaptation of classical proteomic strategies in the analysis of ancient samples to meet the different aims in the

cultural heritage field. Overall, it is intended to illustrate the logic that might address the choice of sample treatment and data analysis to get the most of the experiment.

4.1 Materials and method

4.1.1 Materials

Ammonium hydrogen carbonate (Ambic), Ethylenediaminetetra-acetic acid (EDTA); Tri(hydroxymethyl)aminomethane (TRIS), Urea, GuHCl and TPCK-treated trypsin were from Sigma; Formic acid and Acetonitrile (ACN) were purchased from Baker. Deionized water was obtained from Millipore cartridge equipment. Hydrochloric acid was purchased from Carlo Erba.

Models of paint layers were prepared with milk as binders and azurite ($Cu_3(CO_3)_2(OH)_2$), minium (Pb_3O_4), calcite ($CaCO_3$), and vermilion (HgS) as pigments on glass slides and skimmed milk as control without pigments. Paint replicas were left to dry at RT on the bench for one month. Bone sample was a fragment from human bone of the first century A.D.

4.1.2 Sample treatment

Urea pre-treatment: 10 µL of a solution of 6 M Urea was added to micro-samples (ca 300–800 µg) and incubated for 10 min at RT, followed by sonication for 20 min. Urea was then 6-fold diluted with water.

EDTA pre-treatment: about 100 µL of a solution of 0.5 M EDTA was added to the bone fragment for 10 days at RT, refreshing the solution every 2 days. After centrifugation for 2 min at 10,000 rpm in a benchtop microfuge, the Urea protocol described above was applied.

HCl pre-treatment: 50 µL of 0.6 M HCl was added to the bone fragment and incubated at 4 °C for 4 h. After centrifugation for 2 min at 10,000 rpm in a benchtop microfuge the supernatant was removed and washed with 20 µL of 10 mM Ambic. Washes were repeated for four times. 100 µL of 50 mM Ambic was added and sample was left at 65 °C for 3 h. After centrifugation at 10,000 rpm for 15 min the supernatant was removed.

GuHCl pre-treatment: 200 µL of 0.6 M HCl was added and incubated at 4 °C for 18 h. After centrifugation for 1 min at 14,000 rpm in a bench-top microfuge, and the acid-insoluble pellet washed three times with 200 µL of distilled water. The pellet was incubated at 4 °C for 72 h in a buffer containing 100 mM Tris and 6 M GuHCl at pH 7.4. The sample was then centrifuged for 1 min at 14,000 rpm in a benchtop microfuge. The supernatant was buffer-exchanged into 10 mM Ambic using 3 K molecular weight cut-off Amicon Ultra, centrifugal

filter unit.

4.1.3 Protein digestion and LC-MS/MS analysis

After any pre-treatment of the sample, enzymatic digestion was carried out as in the minimally invasive proteomic analytical procedure described by Leo et al. (4) Briefly, trypsin was added to a final concentration of 10 ng/µL to micro-samples (ca 300–800 µg) as directly suspended in 50 µL of Ambic 10 mM. After incubation at 37 °C for 16 h, the supernatants were recovered by centrifugation at 10,000 rpm, and the peptide mixture was filtered on 0.22 µm PVDF membrane (Millipore), concentrated and purified using a reverse-phase C18 Zip Tip pipette tip (Millipore). Peptides were eluted with 20 µL of a solution made of 50% Acetonitrile, 50% Formic acid 0.1% in Milli-Q water and analysed by LC-MS/MS. LC-MS/MS analyses were carried out on a 6520 Accurate-Mass Q-Tof LC/MS System (Agilent Technologies, Palo Alto, CA, USA) equipped with a 1200 HPLC System and a chip cube (Agilent Technologies). After loading, the peptide mixture was first concentrated and washed on a 40 nl enrichment column (Agilent Technologies chip), with 0.1% formic acid in 2% acetonitrile as eluent. The sample was then fractionated on a C18 reverse-phase capillary column (Agilent Technologies chip) at a flow rate of 400 nL/min, with a linear gradient of eluent B (0.1% formic acid in 95% acetonitrile) in A (0.1% formic acid in 2% aceto-nitrile) from 3% to 80% in 50 min. Peptide analysis was performed using data-dependent acquisition of one MS scan (mass range from 300 to 2000 m/z) followed by MS/MS scans of the three most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal surpassed the threshold of 50,000 counts. Double and triple charged ions were preferably isolated and fragmented.

4.1.4 Data handling

The acquired MS/MS spectra were transformed in Mascot Generic files (.mgf) format and used to query the SwissProt database 2015_04 (548,208 sequences; 195,282,524 residues), with Chordata as taxonomy restriction for protein identification in paint reconstruction samples and with *Homo sapiens* for bone samples.

A licenced version of Mascot software (www.matrixscience.com) version 2.4.0. was used with trypsin as enzyme; 3, as allowed number of missed cleavage; 10 ppm MS tolerance and 0.6 Da MS/MS tolerance; peptide charge from +2 to +3. No fixed chemical modification was inserted, but possible oxidation of methionines, formation of pyroglutamic acid from glutamine residues at the N-terminal position of peptides, and deamidation at asparagines and

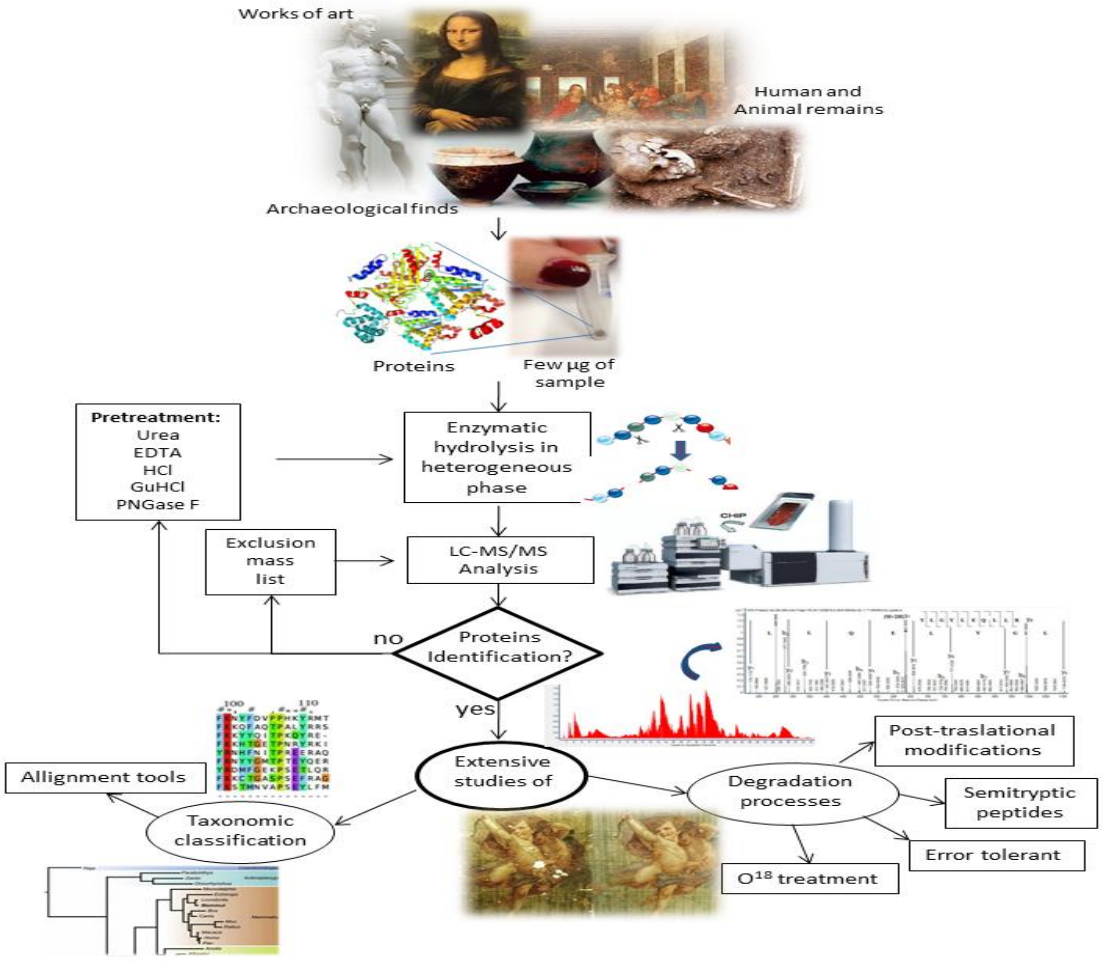
glutamines were considered as variable modifications. (5) When collagen proteins were identified, a further identification run was carried out, with the insertion of hydroxylation on lysine and proline as variable modifications, since more confident identifications are commonly obtained for these proteins by taking into consideration their extensive post-translational modifications. (5) Only proteins presenting two or more peptides were considered as positively identified. Individual ion score threshold provided by Mascot software to evaluate the quality of matches in MS/MS data was generally 31 for paintings and 43 for human samples. Spectra with Mascot score below 10 were rejected.

4.2 Results and discussion

4.2.1 The minimal protocol

The commonly used "bottom-up" approach to identify proteins is based on the enzymatic digestion of the proteins and can be directly performed in heterogeneous phase on a sample fragment followed by mass spectrometric analyses of the released peptides (Fig. 1).

Figure 4.1: General flow chart of proteomic investigations.



Even when the protein is embedded in a complex mixture such as that of a painting layer, few peptides released by the protease without any pretreatment of the sample are enough to identify the protein, (4) without significantly physically affecting the sample itself. This avoid the extraction of the whole protein from the sample using harsh methods, while digestion can rather be carried out by depositing on the surface of a small sample an aqueous neutral solution containing the enzyme that directly trims protruding peptides. The solution will then be gently removed and, once released, peptides can be analysed by mass spectrometric methods such as MALDI-TOF and LC-MS/MS, and database searches by bioinformatics tools such as Mascot (www.matrixscience.com) allow protein identification. This micro-invasive protocol (i.e. the intervention is intrinsically invasive but requires only a minimal quantity of material to work with) proved to be successful when applied to fragments of paintings from the collapsed vault of the Basilica di S. Francesco in Assisi, (4) and samples collected from the Camposanto Monumentale in Pisa, (6) and it was not significantly affected by the pigment that was present in the sample, i.e. the different metals do not actually affect the quality of the results. (3) It is worth mentioning that this protocol is closely similar to the procedure adopted when biocleaning of works of art is carried out, (7) and therefore the meaning of "destructive or invasive" approach should be resized and this can be safely considered as a minimally invasive or rather micro-invasive procedure.

Protein identification can fail because of overwhelming proteins from unavoidable contaminants. Samples coming from artwork have, indeed, an intrinsic contamination problem that originates from i.e. environmental exposure, restoration interventions, and so on, that cannot be overcome by just operating all the chemical manipulations in controlled conditions as in ordinary proteomic analysis. To circumvent unavoidable problems arising from "historical" contaminations, an exclusion list of the peptides which, in a first LC-MS/MS run allowed for identification of keratins or other protein contaminants and that in a further subsequent run have to be ignored by the mass spectrometer for fragmentation, can be adopted. It is a sort of instrumental trick to avoid "waste" of the mass spectrometer's time in fragmenting "useless" peptides derived from trypsin autodigestion or common protein contaminants. (4) The list is created *ad hoc* from the raw data of the LC-MS/MS analysis and a second analysis is carried out with the same LC-MS/MS method, but for the addition of the exclusion list.

4.2.2 Sample pre-treatments

It is evident from what outlined above that accessibility of proteases to the protein in the sample is likely the most important aspect for a successful proteomic experiment in cultural heritage application (obviously beside the actual presence in the sample of proteinaceous material). Whatever reduces the proteolysis yield and, therefore, determines an inefficient and poor production of peptides, would sensibly affect protein identification. For instance, the persistence of stabilizing interactions such as those occurring in structured proteins can greatly impair the efficacy of the enzymatic digestion. By reasoning as in classical biochemical experiments, where denatured, unfolded proteins are digested much more efficiently than structured, folded proteins, we introduced a denaturing step before trypsin digestion to "open" residual structural elements in proteins in the perspective that flexibility of the polypeptide substrate is an absolute requirement for the protease to properly hydrolyse peptide bonds. In the hypothesis that treatment with traditional protein denaturing agents such as urea or guanidinium chloride will make digestion sites more amenable to protease attack even in an unusual "dehydrated, non-soluble" physical state experienced by proteins in works of art or archaeological remains, we tested the denaturing pre-treatment in some cases where the minimally invasive approach described above failed. As an example, Table 1 reports a comparison of the results obtained by the minimally invasive approach, and the analysis carried out on the same pictorial sample after pre-treatment with 6 M urea, followed by dilution and tryptic digestion.

Table 4.1: Proteins identified in the paint replica containing minium and milk by LC-MS/MS. Aliquots were treated in heterogeneous phase with trypsin with the minimally invasive protocol either without any sample pre-treatment or with Urea pre-treatment and the resulting peptide mixtures were analyzed by LC-MS/MS. Proteins were identified in the Uniprot database with Mascot MS/MS Ion Search software, with Chordata as taxonomic restriction, with methionine oxidation, formation of pyroGlu at the N-terminus of Gln, and deamidation (N, Q) of peptides as variable modifications. Individual ion scores >31 indicate identity or extensive homology. Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html).

Sample	Protein (UniProt Accession number)	Minimally invasive protocol			Urea pre-treatment		
		Sequence coverage (%)	Protein score	n° of peptides	Sequence coverage (%)	Protein score	n° of peptides
Milk and minium	alpha-S1 casein (P02662)	68	548	11	75	828	22
	alpha-S2 casein (P02663)	33	259	8	54	753	13
	beta-casein (P02666)	49	298	6	79	536	11
	kappa-casein (P02668)	22	166	5	74	319	8
	beta-lactoglobulin (P02754)	38	308	6	36	309	7

The results clearly show that the pre-treatment with urea improves the quality of identification. We successfully applied the urea pre-treatment protocol to some gilding samples, (8) to ink samples from the Qumran archaeological site, (9) and to a small stone flake from Sibudu Cave dated 49,000 years ago. (5) This urea pre-treatment protocol is slightly more aggressive and therefore more invasive than that described above, and it should be used as a second attempt, when the above protocol with only protease in bicarbonate buffer fails or when more peptides or higher quality spectra are required. It is worth noting that it can be also directly applied on the same sample after the first attempt with the protease in the simple ammonium bicarbonate buffer has failed, thus not reducing the sample size. In principle, after the removal of the bicarbonate buffer containing the protease from the solid sample, the urea pretreatment can be carried out without any further step in between.

Moreover, it was very recently demonstrated (10) that the introduction of a deglycosylating step with PNGaseF before the digestion with the protease, greatly improves proteins identification when egg containing samples are analysed. This further demonstrates that accessibility of peptidic linkages to proteases is the absolute requirement for a reliable identification.

While two or very few peptides can be enough to merely identify materials, more extended sequence coverage could be needed for more detailed analyses, such as species discrimination, and/or conservation state evaluation. In fact, if the goal is to discriminate the organism of origin of proteins that are highly conserved throughout the evolution and among species, it is important to identify proteotypic peptides, i.e. peptides that are unique to the protein sequence specific for an individual organism and not in common to other species.

This can be easily explained with the example of milk proteins: among the peptides that have been identified in the analysis of an ancient food residue by Hong et al., (11) some are shared by goat/ sheep and cattle sequences while others can be used to discern the origin of milk in the ancient residue. Similarly, detection of β-lactoglobulin in dental calculus is per se a direct evidence of milk consumption and can constitute a signature for adoption of dairying habits. Most interestingly, identification of specie-specific peptides allowed discrimination of the origin of dairy product whether they are cattle, sheep or goat dairy product. (12) Specie-specific information that are peculiar of proteomics are therefore extremely useful for delineating domestication timelines and paleodietary habits. To meet aims like these, however, the mild protocols described above might be not always adequate and alternative procedures might be required, to obtain a higher number of peptides and consequently higher sequence coverage, thus increasing the probability of detecting proteotypic peptides.

This is definitively the case of collagen from bones, where an in depth analysis might be needed to gain extremely important information that can be used for a wide range of purposes, from taxonomic analyses, (5,13,14,15) to the characterization of degradation processes. (16) Proteins and collagen in particular, are indeed gaining momentum and are now supporting DNA in evolutionary studies, since they are more stable than nucleic acids, can persist much longer, and, moreover, the survived molecules bear the signature of time.

As commented by the authors of a tremendous paper in Nature this year (15,17) “ancient proteins could now prove as revolutionary as DNA for studying the tree of life”, since we could be able to find proteins that are orders of magnitude older than the oldest DNA discovered so far. Intrinsic collagen protein stability and its entrapment within the hydroxylapatite protective cage make the collagen the longest surviving protein in ancient bone. (18) Consequently and adversely, collagen is highly resistant to extraction and it can be quite challenging to obtain a good sequence coverage that would allow highly reliable sequence comparisons for taphonomic studies (i.e. studies of decaying organisms over time). Procedures need to be optimized for protein extraction from ancient bones, and several

procedures have been proposed (17-19) and references therein). Table 2 reports the sequence coverage of human COL1α1 (collagen alpha-1(I)) and COL1α2 (collagen alpha-2(I)), the two chains constituting type I collagen, as obtained in the analysis of a human bone dated the first century A.D., using selected different extraction procedures.

Table 4.2: Sequence coverage (%) of Collagen alpha-1(I) and Collagen alpha-2(I) from the digestion of an ancient human bone with different pretreatment protocols followed by LC-MS/MS analysis. Sequence coverage was calculated on the sequence of the mature form of the protein. Proteins were identified in the UniProt database with Mascot MS/MS Ion Search software, with *Homo sapiens* as taxonomic restriction, with methionine oxidation, formation of pyroGlu at the N-terminus of Gln, deamidation (N, Q), and hydroxylation (K, P) of peptides as variable modifications.

Protein (UniProt Accession number)	**Minimally invasive protocol** (%)	**Urea** (%)	**EDTA** (%)	**HCl** (%)	**GuHCl** (%)
Collagen alpha-1(I) (P02452)	24.0	60.0	75.0	37.5	63.4
Collagen alpha-2(I) (P08123)	22.8	59.4	72.0	27.5	52.2

It can be noted that even with the minimally invasive protocol, where the bone fragment is simply deepened in a trypsin solution, the number of peptides detected can be high enough to obtain good sequence coverage (24% and 23% for COL1α1 and COL1α2, respectively. However, when decalcification is carried out with EDTA 0.5 M, nearly the whole sequences can be verified, reaching 75% and 72% of sequence coverage for mature COL1α1 and COL1α2, respectively.

However, while decalcification with EDTA seems to be the ultimate choice at this stage for collagen protein sequence coverage, it might not be the best choice when non-collagenic proteins (NCPs) are searched. (18,20) NCPs have potentially higher phylogenetic value than collagen I because of their greater sequence variation, but they are less resistant than collagen I, constitute a relatively small fraction of the total protein content of bones, and are more soluble than collagen and might then be lost in demineralization steps. The choice of the protocol to be used for bone treatment will ultimately depend on the protein/s of interest but also on the analytical application. In fact, the extraction protocol used might also influence the results in the following characterization of the proteins recovered from bones.

Obtaining as many peptides as possible and the choice of the appropriate protocol can be also extremely important to study the molecular modifications occurring post-mortem. It is still an

open question whether deamidation of glutamine (Q) and asparagine (N) could be used as a dating technique in ancient sample, but it is widely accepted that deamidation can be considered as a biomolecular marker of deterioration and natural ageing of proteins in artistic and archaeological materials. (6,16,21,22) Detection of deamidation can offer interesting prospects in the evaluation of the conservation state of work of arts and archaeological remains.

4.2.3 Data handling

Development of bioinformatics tools that can address the specific issues, such as the identification of ageing signatures in proteins of ancient samples or to handle extinct species when genomic data are available for the extant ones, is another challenge in the field of proteomics for cultural heritage.

A main difficulty in identifying ancient proteins is indeed the paucity of ancient genomic data. In shotgun proteomics, proteins are routinely identified by matching experimental MSMS spectra of enzymatic digests of protein samples to simulated spectra from protein databases, usually derived from genomic sequences. Identification in database search procedures rely on how well experimental spectra fit to theoretical spectra obtained from sequences that are present in databases. While this procedure works for most of the identifying purposes such as whether egg, milk or animal glue has been used as binder in a painting (4,23) or organic component in food residues, (11,24) the issue might be more complex when trying to assess the species of origin and/or in the case of extinct organisms. Standard database search algorithms fail to identify peptides that are not exactly contained in a protein database, such as those arising from unreported mutations occurred throughout evolution. Good quality fragmentation spectra are in principle discarded because of even point mutation that makes them non-ascribable to known sequences. Identification of the specific species will then depend on the detection of two or few more peptides that are conserved between the ancient protein and its modern counterpart, despite possible differences in other peptides, which will be ignored in a standard database searches. This is an intrinsic limit of the otherwise powerful proteomic approach that will mask novel sequences or points of divergence with reported sequences.

The tremendous attractive possibility of proteomics of reaching much further back in time to gain information on more ancient samples than genomics, thanks to the higher stability of proteins in respect to nucleic acids, and to characterize in depth diagenetic alterations directly on surviving molecules, can be therefore limited in principle by the lacks of corresponding

DNA information.

However, many software tools have been developed for the automated identification of peptides by *de novo* sequencing directly from the MS/MS spectrum of peptides. Representative de novo sequencing software packages include PEAKS (25), PepNovo (26), NovoHMM, (27) and Lutefisk, (28) and very recently, Welker et al. (13) extensively and successfully used PEAKS to overcome the absence of corresponding genomic data to resolve the evolutionary history of Darwin's south American ungulates.

Moreover, the lacks of corresponding DNA information can also be partially overcome with bioinformatic tools such as the error tolerant search utility in Mascot, which allows for single substitution in peptides in respect to sequences which are present in databases, taking for granted that any obtained match has to be manually confirmed afterwards. These approaches allow taxonomic attribution of extinct species, in specimens that no longer can yield DNA (14) and can be combined to a wide panel of other bioinformatics tools such as generation *ad hoc* of peptide databases as experimented by Waters and collaborators (29) that used a collagen sequences database to analyse the mass spectrometric data obtained from a mastodon bone, and classical homology search tools like BLAST. (15)

Error tolerant searches can also be used to detect diagenetic modifications, (30) among which the observation of semi-tryptic peptides accounting for partial hydrolysis of the polypeptidic chain within the samples. (12,16) High occurrence of partial hydrolysis is indeed expected as degradation effect. Table 3 reports a comparison of the results obtained on several pictorial models and on a sample from the Camposanto Monumentale in Pisa when identification is carried out, on the same set of experimental data with Mascot search programme allowing for semi-tryptic cleavages and only for specific tryptic cleavages: the abundance of non-tryptic termini suggests hydrolysis as a likely and expected effect of protein degradation.

Table 4.3: Proteins identified in the paint replica containing different inorganic pigments mixed with milk as binder, and in the sample from the Camposanto Monumentale of Pisa, by trypsin digestion in heterogeneous phase and LC-MS/MS analysis. Proteins were identified in the UniProt database with Mascot MS/MS Ion Search software, with *Chordata* as taxonomic restriction, with methionine oxidation, formation of pyroGlu at the N-terminus of Gln, and deamidation (N, Q) of peptides as variable modifications. Individual ion scores >31 indicate identity or extensive homology. Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html).

Sample	Protein (UniProt Accession number)	Tryptic identification			Semitryptic identification		
		Sequence coverage (%)	Protein score	n° of peptides	Sequence Coverage (%)	Protein score	n° of peptides (specific tryptic peptides)
Casein and minium	alpha-S1 casein (P02662)	76	641	16	77	1496	27 (16)
	alpha-S2 casein (P02663)	56	512	17	56	603	21 (17)
	beta-casein (P02666)	53	426	9	69	1104	14 (9)
	kappa-casein (P02668)	24	242	5	50	520	8 (5)
Casein and cinnabar	alpha-S1 casein (P02662)	69	666	15	69	1718	31 (15)
	alpha-S2 casein (P02663)	52	530	12	53	792	14 (12)
	beta-casein (P02666)	79	465	12	79	1070	21 (12)
	kappa-casein (P02668)	46	345	6	53	728	8 (6)
Casein and $CaCO_3$	alpha-S1 casein (P02662)	64	483	12	64	1072	21 (12)
	alpha-S2 casein (P02663)	52	356	6	52	453	13 (6)
	beta-casein (P02666)	47	266	7	61	732	21 (7)
	kappa-casein (P02668)	26	214	3	47	363	6 (3)
Casein	alpha-S1 casein (P02662)	71	589	14	72	1339	23 (14)
	alpha-S2 casein (P02663)	53	512	14	53	560	18 (14)
	beta-casein (P02666)	79	445	12	79	1058	19 (12)
	kappa-casein (P02668)	44	272	7	53	567	9 (7)
Sample from Camposanto Monumentale	alpha-S1 casein (P02662)	50	565	11	56	1377	31 (11)
	alpha-S2 casein (P02663)	27	252	7	29	374	10 (7)
	beta-casein (P02666)	71	432	8	72	1601	32 (8)
	kappa-casein (P02668)	22	141	3	46	391	10 (3)
	Collagen alpha-1(I) chain (P02453)	36	854	25	42	879	28 (25)
	Collagen alpha-2(I) chain (P02465)	23	522	21	25	558	22 (21)

Moreover, a clever use of the "open mass" modification search utility in the Protein Prospector database search, as recently suggested by Hill et al., (16) in the analysis of fossil bones, allowed to identify extensive surviving galactosylation and glucosyl-galactosylation of hydroxylysine residues in collagen. This bioinformatics strategy of analysis offers the potentiality to detect unexpected modifications and certainly will hold interesting prospects in the characterization of molecular details of degradation processes.

The modifications that occur on proteins in natural environment over time are still under investigation and delineating both *in vivo* and diagenetically derived alterations will provide important information on the physiology and/or phylogenies of organisms, as well as on the ageing mechanisms. These can be function of specific environmental factors, and their knowledge will lead to a more conscious preservation of ancient samples.

Moreover, the intrinsic damaged nature of the ancient proteins can intuitively be expected to be the first evidence of authenticity of a sample, that ruling out possible contaminations from modern materials. (31)

While much effort has been already devoted to the development of identification tools, a lot of work has still to be done to understand and characterize the whole range of modifications occurring upon ageing on deteriorating proteins in samples that have been exposed to a wide spectrum of different environmental conditions, thus contributing to what can be called the field of paleoproteomics.

4.3 Conclusions

The physical state of the samples and the degradation processes undergone during ageing, provide unusual challenges that require to be counteracted by specific adaptations of the classically adopted protocols used in the analysis of biological samples. Experiments to improve and develop new protocol have been described in this chapter. A denaturing step with 6M urea was used before trypsin digestion to "open" residual structural elements in proteins in the perspective that flexibility of the polypeptide substrate is an absolute requirement for the protease to properly hydrolyze peptide bonds. Results obtained in comparison with minimally invasive approach and the analysis carried out clearly indicate that the denaturating pre-treatment with urea improves the quality of the identifications.

Since more extended sequence coverage could be needed for detailed analyses, such as species discrimination, and/or conservation state evaluation, 5 different protocols whit or without any pretreatment were tested. Protocols with EDTA pretreatment seems to be the best choice to improve sequence coverage.

Furthermore, it was demonstrated that a clever use of bioinformatics tools such as operating Mascot in an error tolerant search mode, can allow to identify degradation processes such as random proteolysis.

4.4 References

1. en Have ST, Ahmad Y, Lamond AI (2011) Proteomics — current novelties and future directions. J. Anal. Bioanal. Tech. (1) 1–6.
2. Ostrom PH, Schall M, Gandhi H, Shen TL, Hauschka PV,. Strahler JR, et al. (2000) New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. Geochim. Cosmochim. Acta (64) 1043–1050.
3. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.
4. Leo G, Cartechini L, Pucci P, Sgamellotti A, Marino G, Birolo L (2009) Proteomic strategies for the identification of proteinaceous binders in paintings. Anal. Bioanal. Chem. (395) 2269–2280.
5. Villa P, Pollarolo L, Degano I, Birolo L, Pasero M, Biagioni C, et al. (2015) A milk and ochre paint mixture used 49,000 years ago at sibudu, South Africa. PLoS One (10)
6. Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, et al. (2011) Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. Anal. Chem. (83) 2056–2064
7. Ranalli G, Alfano G, Belli C, Lustrato G, Colombini MP, Bonaduce I, et al. (2005) Biotechnology applied to cultural heritage: biorestoration of frescoes using viable bacterial cells and enzymes. J. Appl. Microbiol. (98) 73–83.
8. Gambino M, Cappitelli F, Cattò C, Carpen A, Principi P, Ghezzi L, et al. (2013) A simple and reliable methodology to detect egg white in art samples. J. Biosci. (38) 397–408.
9. Rasmussen KL, Tenorio AL, Bonaduce I, Colombini MP, Birolo L, Galano E, et al. (2014) Reply to Ira Rabin's comment on our paper Rasmussen et al. J. Archaeol. Sci. (43)155–158.
10. Vinciguerra R, Galano E, Vallone F, Greco G, Vergara A, Bonaduce I, et al. (2015) Degly-cosylation step to improve the identification of egg proteins in art samples. Anal. Chem.
11. Hong C, Jiang H, Lü E, Wu Y, Guo L, Xie Y, et al. (2012) Identification of milk component in ancient food residue by proteomics. PLoS One (7) 1–7.

12. Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, Trachsel C, et al. (2014) Direct evidence of milk consumption from ancient human dental calculus. Sci. Rep. (4) 7104.
13. Rybczynski N, Gosse JC, Harington CR, Wogelius R, Hidy AJ, Buckley M (2013) Mid-Pliocene warm-period deposits in the high Arctic yield insight into camel evolution. Nat. Commun. (4) 1550
14. Buckley M, (2013)A molecular phylogeny of plesiorycteropus reassigns the extinct mammalian order "Bibymalagasia". PLoS One (8)
15. Welker F, Collins MJ, Thomas JA, Wadsley M, Brace S, Cappellini E, et al. (2015) Ancient proteins resolve the evolutionary history of Darwin/'s south American ungulates. Nature (522) 81–84.
16. Hill RC, Wither MJ, Nemkov T, Barrett A, D'Alessandro A, Dzieciatkowska M, et al. (2015) Preserved proteins from extinct bison latifrons identified by tandem mass spectrometry; hydroxylysine glycosides are a common feature of ancient collagen. Mol. Cell. Proteomics (14) 1946–1958.
17. Marchant J, (2015) Mystery of Darwin's "strange animals" solved. Nature 1–3
18. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. Rapid Commun. Mass Spectrom. (28) 605–615.
19. Jiang X, Ye M, Jiang X, Liu G, Feng S, Cui L, et al. (2007) Method development of efficient protein extraction in bone tissue for proteome analysis. J. Proteome Res. (6) 2287–2294.
20. Cleland TP, Voegele K, Schweitzer MH(2012) Empirical evaluation of bone extraction protocols. PLoS One (7) 1–9.
21. Wilson J, Van Doorn NL, Collins MJ (2012)Assessing the extent of bone degradation using glutamine deamidation in collagen. Anal. Chem. (84) 9041–9048.
22. Solazzo C, Wilson J, Dyer JM, Clerens S, Plowman E, Von Holstein I (2014) Modeling Deamidation in Sheep α-Keratin Peptides and Application to Archeological Wool Textiles. Anal. Chem. (86) 567–575.
23. Tokarski C, Martin E, Rolando C, Cren-Olivé C (2006) Identification of proteins in renaissance paintings by proteomics. Anal. Chem. (78) 1494–1502.

24. Buckley M, Melton ND, Montgomery J (2013) Proteomics analysis of ancient food vessel stitching reveals N4000-year-old milk protein. Rapid Commun. Mass Spectrom. (27) 531–538.
25. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, et al. (2003) PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun. Mass Spectrom. (17) 2337–2342.
26. Frank A, Pevzner P (2005) PepNovo: de novo peptide sequencing via probabilistic network modeling. Anal. Chem. (77) 964–973.
27. Fischer B, Roth V, Buhmann J, Grossmann J, Baginsky S, Gruissem W, et al. (2005) A hidden Markov model for de novo peptide sequencing. Anal. Chem. (77) 7265–7273.
28. Taylor JA, Johnson RS Sequence (1997)database searches via de novo peptide sequencing by tandem mass spectrometry. Rapid Commun. Mass Spectrom. (11) 1067–1075
29. Waters MR, Stafford TW, McDonald HG, Gustafson C, Rasmussen M, Cappellini E, et al. (2011) Pre-clovis mastodon hunting 13,800 years ago at the Manis site, Washington. Science (334) 351–353.
30. Cappellini E, Jensen LJ, Szklarczyk D, Ginolhac A, Da Fonseca RR, Stafford TW, et al. (2012) Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. J. Proteome Res. (11) 917–926.
31. Buckley M, Walker A, Ho SYW, Yang Y, Smith C, Ashton P, et al. (2008) Comment on “protein sequences from mastodon and tyrannosaurus rex revealed by mass spectrometry,” . Science (319) 33.

5. Burial under the ashes of the Vesuvius eruption

Among the archaeological remains, bones are most interesting. (1) The application of proteomics to the study of ancient bone holds great potential for increasing our understanding of many aspects of "paleobiology" such as the evolution, the habits and the ecology of ancient organisms.

The survival of proteins in fossil bones has been of great interest for their potential in recovering phylogenetic information for over 30 years. (2) Bones, indeed, can be an abundant reservoir of ancient biomolecules due to natural resistance to post-mortem decay arising from a unique combination of mechanical, structural, and chemical properties. While fossil bones represent the standard source of ancient DNA molecules, collagen is the dominant protein in mineralised tissue, and it is known to persist with extraordinary longevity, thus adding it to the list of ancient biomarkers amenable to molecular characterization, and opening access to a reservoir of previously unavailable genetic information for phylogenetic inference. (1,2)

Collagens are well preserved and highly conserved throughout the evolution of species due to theirs essential biological function as well as their peculiar sequence, made of one third of glycine, that have to be strictly conserved due to structural role. Sequences differ by only few amino acids among species that are nevertheless enough for MS techniques to allow to discriminate the species of origin. MS screening of species-specific peptides from collagen in bones has therefore the potentially to be used as a low-cost, rapid alternative to DNA sequencing for taxonomic attribution of morphologically unidentifiable samples. (3)

Moreover, the effect of aging, environmental and storage conditions at the molecular level can be analysed and studied as chemical modifications induced in proteins, possibly leading to the discovery of potential biomarkers to be searched for in ancient samples. In particular, degradation is visible at amino acid residue level and deamidation of asparagine and glutamine, a spontaneous non-enzymatic postsynthetic modification of protein (resulting in a mass shift of +0.984 Da), has been detected as an important biomolecular marker of the deterioration and natural aging of proteins in artistic and archaeological materials. (4,5)

Moreover, *in vitro* and *in vivo* studies have moreover established that nonenzymatic deamidation of asparaginyl and glutaminyl residues acts as a molecular clock to time biological processes such as the regulation of protein turnover and tissue aging. (6)
Proteomic strategies are herein extensively used as the approach for the characterization of the aging and deterioration phenomena occurring to proteinaceous materials. (7).
This project was focused on the degradation of the proteins, at the molecular level, due to the storage conditions and was aimed at paving the way for in-depth paleo-proteomic characterization, taking advantage from the advent of the latest proteomic platforms that should nevertheless be adapted to ancient samples.

5.1 Structure of bones

The complex collagen structure gives bone its strength and therefore ensures its preservation within the burial environment, and its prevalence in archaeological scientific research. (8-11)

In order to understand the mechanical properties of this material it is important to consider its component phases and the structural relationships between them. Bone have five hierarchical structural levels. (12)There are many factors other than age which can contribute to the level of preservation of organic material in bone; these include conditions of the burial environment such as: temperature, pH and fluctuation of the water table. (13) In order to understand the diagenetic process of collagen, degradation of the sub-nanostructures of bone needs to be described. Living bone material is mainly composed of; water, organic, and inorganic fractions. The inorganic fraction contains mineral salts called hydroxyapatite (HAP). The organic fraction includes proteins, lipids and water. (12)

The importance of bone mineral in the role of preservation of the organic fraction is highlighted by the presence of the hydroxyapatite (mineral) crystals, which embed and protect the protein, and contribute to the stability and preservation of bone over geological timescales. (11,14) The presence of mineral is thought to have a powerful stabilizing effect, protecting bone against both microbial and chemical degradation. (11)

Collagen makes up approximately 90% of the protein matrix. (11,15) This fibrous protein is made up of three polypeptide chains of similar size. Type I collagen is the main protein found in bone material and is heterotrimeric, consisting of two identical chains identified as α 1, and a third identified as α-2. (16-18) The peptide sequence of collagen is repetitive, with every third residue being Gly. (19) Other residues in the sequence are described in their relation to Gly. It is common for x and y positions within collagen to contain high levels of proline (Pro)

and lysine (Lys). (20) This repeating sequence helps to establish the topology required to form the triple helix characteristic of collagen. (21) Gly, is an important part of the chain. This amino acid is the smallest in size and makes up approximately a third of the protein, enabling assemblage of adjacent chains, as well as being essential for the formation of the twist which is needed to create the helical shape. (16) Collagen is exposed to extensive post-translational modifications. (22) These modifications can occur through both enzymatic and non-enzymatic degradation. One relevant example of a post-translational modification, which frequently occurs in collagen, is the hydroxylation of proline. Interactions within the triple helix are enhanced by the hydrogen bonding between the amide group and the hydroxyl group of hydroxyproline (Hyp). During enzymatic degradation, an enzyme called lysyl oxidase forms cross-linked bonds between the polypeptide chains. (23)

The shape and morphology of each bone will vary depending on its role within a given biological system; however the basic chemical structure remains the same, with type I collagen being ubiquitous throughout the animal kingdom. (21) Procollagen is the immature form of the protein; at this stage both the N and C terminal propeptidesare still attached at the end of the triple helix. These propeptides have different roles in collagen fibril formation. The main role of the C-propeptide is to ensure that the procollagen remains soluble during its trafficking through the cell, whereas the N-propeptide ultimately influences the shape and diameter of the fibril. (16) In order to form the microfibril, the C and N propeptidesare removed enzymatically by procollagen C and N proteinases. Five collagen helices can then begin to fold, align and twist into a microfibril. The way in which these collagen helices line up is crucial to the formation of the microfibril.

5.2 The 79 AD eruption of Mount Vesuvius

Vesuvius, on the west coast of Italy, is one of the most studied volcanoes in the world, because of its easy accessibility, and the first well-documented historic explosive eruption of 79 AD. The volcano is classed as a complex stratovolcano because its eruptions typically involve explosive eruptions as well as pyroclastic flows. (24) The eruption destroyed Pompeii, Herculaneum, Oplonti and Stabiae and caused the death of Pliny the Elder among many other people. Before the eruption, earthquakes occurred for some time, but they were disregarded by local inhabitants because of their familiarity with the phenomenon. As the younger Pliny testified "*for several days before (the eruption) the earth had been shaken, but this fact did not cause fear because this was a feature commonly observed in Campania*". (25)

The main phases of the eruption have been described by Pliny the Younger who observed the eruption from a distance of more than 25 km, taking into consideration also the observations of contemporary witnesses and closer view accounts, such as his uncle, Pliny the Elder, who died while rescuing the inhabitants of the area. The beginning of the eruption is uncertain: the two Plinys observed the cloud at the seventh hour of the day (1 PM). We must presume that the eruption began earlier to allow the arrival, at about the same hour, of a messenger sent from the Vesuvian area.The eruptive cloud was directly observed from Misenum at a distance of 21 km, so that they could fully appreciate its total extent and behavior. During the night of the first day of the eruption, and in the morning of the next day, the houses of Misenum were shaken by earthquakes that caused a lot of panic. In the morning of the second day of eruption, Pliny the Younger observed the development of pyroclastic flows descending down the flanks of Vesuvius and flowing on the sea. " *From the other side, black and horrible clouds, broken by sinuous shapes of flaming winds, were opening with long tongues of fire ... .After a little while descended onto the land, opened the sea, covered Capri and prevented the sight of Misenum ...*" The sequence of events described by the Younger Pliny fits well the geologic record of the eruption (Lirer et al., 1973; Sigurdsson et al., 1985).(25)

The cities of Pompeii and Herculaneum, were destroyed and many of its inhabitants died during the AD 79 catastrophic eruption in just 24 hours. Two thousand of people died, and the cities were abandoned for many years. When a group of explorers rediscovered the site in 1748, they were surprised to find that–underneath a thick layer of dust and debris–Pompeii was mostly intact. The buildings, artifacts and skeletons left behind in the buried city have showed us most about everyday life in the ancient world.(24)

Owing to their different locations Pompeii and Herculaneum were buried in different ways and this has affected the preservation of materials at site. Herculaneum was a small seaside town whereas Pompeii was the industrial hub of the region. Work continues at both sites and recent excavations at Herculaneum have uncovered beautiful and fascinating artefacts. These artefacts include finely sculpted marble reliefs, intricately carved ivory panels and fascinating objects found in one of the main drains of the cities. (24)

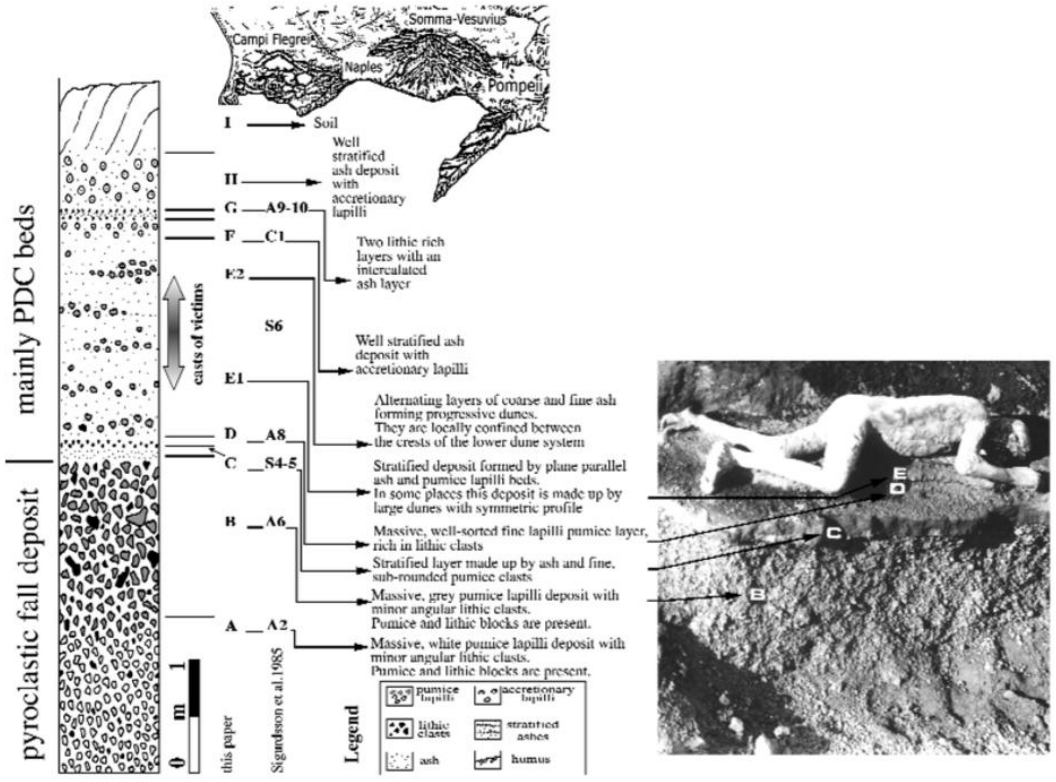
5.2.1 The effects of the 79 AD eruption on Pompeii

The process of Pompeii's destruction and burial started with the accumulation of a thick pumice lapilli deposit (layers A and B in Figure 5.1) resulting from the column fallout. The rate of deposition in the city ranged from 15 cm per hour in open areas to 25/30 cm per hour

in places accumulating additional pyroclasts rolling from the steeper roofs. Within six hours from the beginning of the eruption the roofs and part of the walls of the buildings had collapsed under the pumice load. By the morning of 25 August most structures were seriously damaged; the pumice fall deposit, generally 3 m thick, totally buried the lower part of the buildings. The percentage of victims (38%) found in this deposit at Pompeii is anomalously high while a 4% of deaths were caused by tephra fallout in the last four centuries during explosive eruptions. (26) This high percentage of deaths is possibly due to the attempt of some people to take shelter into buildings where roofs and walls collapsed under the load of the pyroclastic material. The small percentage of people found dead outdoor within the pumice fall deposit was probably killed by the crumbling roof tiles or by the largest lithic fragments following ballistic paths. The first pyroclastic density current (PDC) flowed through the city depositing the basal ash layer C (fig. 5.1) and caused irrelevant damages. Based on the evidence that all of the human remains lie above this deposit, it can be deduced that people were not killed by the earlier PDC. (27) The inhabitants survived also to the successive fallout phase that emplaced the lithic-rich layer D (fig. 5.1) and some were able to walk outdoor during the emplacement of the basal part of the unit E. In fact, some of the victims were found several centimetres above the base of this unit. Possibly, the parental pyroclastic current ran over the city with a low-temperature, dilute making few centimetres of ash. The rear part of the current had a non-uniform behaviour in terms of concentration, possibly due to the canalization of the basal part of the current along the longitudinal walls of the buildings. Inside these areas the current showed a greater destructive power, flattening most of the (especially transversal) walls, standing out of the pumice fallout deposit, in its north-south path. In the areas outside the channels the current had essentially a depositional behaviour engulfing the city and suffocating the inhabitants. These opposed behaviour of the PDC in very close areas testify the different kinetic conditions of Pompeii inhabitants and hence the different physical integrity of their bodies.Observations on objects, cloths, frescoes and skeletons rule out the possibility that burn injuries contributed to kill Pompeii inhabitants, as recently proposed for Herculaneum inhabitants. (28,29) Furthermore, the proposed non-uniform behaviour of PDCs, due to the interaction with the urban structures, justifies the different state of destruction of the buildings throughout the city. A final phreatomagmatic phase, punctuated by two minor lithic fall episodes, emplaced the upper part of the succession (F to H units). Field features, such as the presence of accretionary lapilli in the upper part of the ash and pumice deposit and the lack of high temperature evidences in the buildings,

support the idea of low emplacement temperature for the pyroclastic currents during the final phase of the eruption. (29)

Figure 5.1:Stratigraphic section of the 79 AD deposits at Pompeii with lithofacies characteristics and emplacement mechanisms reported for each layer. The maximum thickness for each unit is represented. On the right, a human cast near the casa delCriptoportico (photo from the Archives of the SoprintendenzaArcheologica di Pompei) stays above a well-exposed stratigraphic sequence (unit B to E1). Above: Campanian Plain with Somma-Vesuvius volcanic complex and the city of Pompeii. Pompeii map and location of regions (roman numbers). Unornamented areas represent unexcavated regions.



5.2.2 The effects of the 79 AD eruption on Herculaneum

The town of Herculaneum, lying at the foot of Mount Vesuvius on a cliff overlooking the sea, was buried by a succession of pyroclastic surges and flows (currents of volcanic ash and hot gases generated by collapse of the eruptive column) during the plinian eruption of AD 79. The skeletons of 80 of 300 people who had taken refuge in 12 boat chambers along the beach have now been unearthed from the first surge deposit. Mastrolorenzo et al. have investigated how these people were killed by this surge, despite being sheltered from direct impact, after its abrupt collapse (emplacement) at about 500 ^{0}C on the beach. The victims' postures indicate that they died instantly, suggesting that the cause of death was thermally induced fulminant shock1 and not suffocation, which is believed to have killed many of the inhabitants of Pompeii and of Herculaneum itself. The first surge was generated 12 hours after the eruption started. Unlike the subsequent surges, it billowed through the evacuated

town of Herculaneum without damaging it and without leaving any deposit or even disturbing small and fragile heat-resistant objects. The surge advanced as a deflating current with low momentum until it reached the 20-metre cliff drop, when its basal, denser component emplaced abruptly on the beach, halted probably by hydraulic-jump effects, bursting into the waterfront chambers and enveloping the people hiding inside. (29)

Therefore, during the eruption Pompeii was covered by a layer "only"6 meters deep consisting of ash and cinder. The ash cases contained skeletons that were kept intact. Herculaneum instead, was destroyed and buried by lava and mud during the eruption and bodies experienced different temperatures and buried conditions than bodies in Pompeii. The bodies were completely burnt.

The inspiring idea of this work is that different types of conservation might be reflected in the nature and extent of chemical modifications. The difference in the level of deamidation, for instance, can constitute an important molecular marker.

Figure 5.2: *Vesuvius from space*

5.3 Aging and modification induced in the fossil bones

Post-translational modifications (PTM) of proteins are part of what makes proteomics so much challenging. They are chemical or enzymatic modifications of amino acids after protein synthesis that change their molecular weights, the fundamental physical property measured by mass spectrometry.

The information that can be obtained in the study of chemical modifications of proteins are very useful to investigate several questions such as the conservation state of a material, the

effect of aging, or the effect of environment (such as pollution, humidity, effect of restoration procedures).

More recently, a particular focus was given to deamidation of both glutamine and asparagine that has been described as a molecular clock that regulates the timing of *in vivo* processes.

In order to properly look at deamidation in collagen, the structure of the microfilbril may be an important factor to consider, when we look at the ability of residues to undergo deamidation. The position of a residue, not only within the polypeptide but also within the microfibril, will influence its steric propensity to modification.

Despite the large amount of clinical based research which has linked deamidation with aging, comparatively few studies have looked at deamidation occurring in protein post-mortem. Postmortem deamidation has been investigated in a range of biological material such as: keratin in wool, (30) collagen in bone, (31-33) protein binders in painting, (7) and keratin in mummified skin. (34) However there have been no robust studies, which have evaluated deamidation in a range of dated material. Whilst analysing collagen peptides for the purpose of species identification, van Doorn et al., (2012) reported that there may be a correlation between deamidation and the thermal age of bone. However, large levels of variation were observed in samples from the same site, with the same thermal age.

5.3.1 Deamidation

Deamidation is a spontaneous postsyntetic modification that plays an important role in protein degradation.

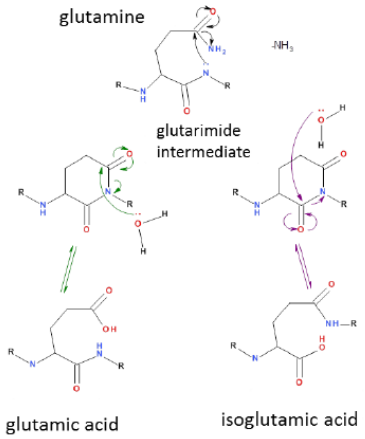
It occurs in Asn and Gln amino acids resulting, respectively, in aspartic/isoaspartic acid and glutamic/γ-glutamic acid.Conversions of Asn and Gln to aspartic and glutamic acids can be detected by MS techniques and MS/MS methods, although the modification results in a mass shift of only 0.984 Da. There are two alternative pathways of deamidation, either via internal cyclisation or via direct sidechain hydrolysis. (6)

Mechanisms of glutamine deamidation

Deamidation of glutamine induces a negative charge followed by removal of ammonia (NH_3) and addition of water (H_2O) resulting in a mass increase of +0.984 Da. This mass changing is readily measurable by mass spectrometry. (7) Glutamine deamidation has been found to occur at a slower rate than asparagine in both synthetic polypeptides and in longer constricted peptide chains. (35) Glutamine can undergo deamidation via a six-membered glutarimide intermediate.

Although it is thought that glutamine has a significantly longer half-life than asparagine, the exact rate of glutamine deamidation is unknown.

Figure 5.3: Glutamine undergoes deamidation via a 6 membered glutarimide intermediate, by which two isomers can form, glutamic or isoglutamic acid.

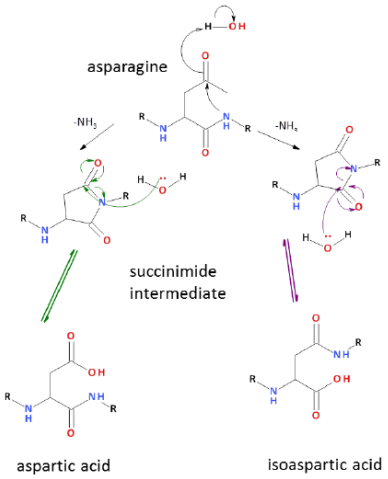


Mechanisms of asparagine deamidation

Asparagine can deamidate via a five membered succinimide intermediate. Asparagine residues have rapid deamidation rates in comparison to those measured in glutamine residues. This would suggest that asparagine would not be an ideal residue choice for measuring long term collagen degradation, since it might occur too rapidly. (35)

It should be noted that deamidation of residues in short synthetic peptides maybe increased by increased flexibility, which may not be comparable to larger intact proteins. (35)

Figure 5.4 Asparagine undergoes deamidation via a 5 membered succinimide intermediate, by which two isomers can form, aspartic acid and isoaspartic acid



5.4 Material and methods

5.4.1 Reagents

Ammonium hydrogen carbonate (AMBIC), Ethylenediaminetetraacetic acid (EDTA); Tri(hydroxymethyl) aminomethane (TRIS), Urea, GuHCl and TPCK-treated trypsin and Water-^{18}O were from Sigma; Formic acid and Acetonitrile (ACN) were purchased from Baker. Deionized water was obtained from Millipore cartridge equipment. Hydrochloric acid was purchased from Carlo Erba.

5.4.2 Samples

Samples analyzed were fragments of human bones of the first century A.D from Pompei obtained from Professor M. Cipollaro from the "Seconda Università di Napoli", and from Ercolano and Scalandrone bay also dated back to the first century AD, from Professor P.Petrone from Università "Federico", Napoli.

5.4.3 Sample treatment

Urea pre-treatment: 10 µL of a solution of 6M Urea were added to micro-samples and incubated for 10 min at room temperature, followed by sonication for 20 min. Urea was then 6-fold diluted with water.

EDTA pre-treatment: about 100 µL of a solution of 0.5M EDTA were added to the bone fragments for 10 days at room temperature, refreshing the solution every 2 days. After centrifugation for 2 min at 10,000 rpm in a benchtop microfuge, the Urea protocol described above was applied.

HCl pre-treatment: 50 µL of 0.6M HCl were added to the bone fragments and incubated at 4°C for 4 hours. After centrifugation for 2 min at 10,000 rpm in a benchtop microfuge the supernatant was removed and it was washed with 20 µL of 10mM Ambic. Washes were repeated for four times. 100 µL of 50mM Ambic were added and sample was left at 65°C for 3 hours. After centrifugation at 10,000 rpm for 15 minutes the supernatant was removed.

GuHCl pre-treatment: 200 µL of 0.6M HCl were added and incubated at 4°C for 18 hours. After centrifugation for 1 min at 14,000 rpm in a benchtop microfuge and the acid-insoluble pellet washed three times with 200 µL of distilled water. The pellet was incubated at 4 °C for 72 h in a buffer containing 100mM Tris and 6M GuHCl at pH 7.4. The sample was then centrifuged for 1 min at 14,000 rpm in a benchtop microfuge. The supernatant was buffer-

exchanged into 10mM Ambic using 3 K molecular weight cut-off Amicon Ultra, centrifugal filter unit.

Each pre-treatment was also carried out using water labeled with ^{18}O.

5.4.4 Protein digestion

After any pre-treatment of the sample, enzymatic digestion was carried out as in the minimally invasive proteomic analytical procedure. (7,36) Briefly, trypsin was added to a final concentration of 0.1 µg/µL in Ambic 10mM.

After incubation at 37°C for 16 hours, the supernatants were recovered by centrifugation at 10,000 rpm, and the peptide mixture was filtered on 0.22 µm PVDF membrane (Millipore), concentrated and purified using a reverse-phase C18 Zip Tip pipette tip (Millipore). Peptides were eluted with 20 µL of a solution made of 50% Acetonitrile, 50% Formic acid 0.1% in Milli-Q water and analysed by LC-MS/MS and/or MALDI-TOF.

5.4.5 Mass spectrometry analyses

MALDI-TOF MS analysis

MALDI-MS analyses were carried out on a 4800 Plus MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA) equipped with a nitrogen laser (337 nm). The peptide mixture was mixed (1:1, v/v) with a 10 mg/ml solution of a-cyano-4-hydroxycinnamic acid in acetonitrile/50mM citrate buffer (70:30 v/v). Mass calibration was performed using external peptide standards purchased from Applied Biosystems. Spectra were acquired using a mass (m/z) range of 400–4,000 and raw data were analyzed using Data Explorer Software provided by the manufacturer. Mass spectra were analysed by manual inspection against the theoretical list of peptides, providing a proteins identification.

NanoLC-MS/MS analysis

LC-MS/MS analyses were carried out on a 6520 Accurate-Mass Q-Tof LC/MS System (Agilent Technologies, Palo Alto, CA, USA) equipped with a 1200 HPLC System and a chip cube (Agilent Technologies). After loading, the peptide mixture was first concentrated and washed on a 40 nl enrichment column (Agilent Technologies chip), with 0.1% formic acid in 2% acetonitrile as eluent. The sample was then fractionated on a C18 reverse-phase capillary column (Agilent Technologies chip) at a flow rate of 400nL/min, with a linear gradient of eluent B (0.1% formic acid in 95% acetonitrile) in A (0.1% formic acid in 2% acetonitrile) from 3% to 80% in 50 min.

Peptide analysis was performed using data-dependent acquisition of one MS scan (mass range from 300 to 2000 *m/z*) followed by MS/MS scans of the three most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal surpassed the threshold of 50,000 counts. Double and triple charged ions were preferably isolated and fragmented.

The acquired MS/MS spectra were transformed in Mascot Generic files (.mgf) format and used to query the SwissProt database 2015_04 (548208 sequences; 195282524 residues), with *Homo sapiens* as taxonomy restriction.

A licensed version of Mascot software (www.matrixscience.com) version 2.4.0. was used with trypsin as enzyme; 3, as allowed number of missed cleavage; 10 ppm MS tolerance and 0.6 Da MS/MS tolerance; peptide charge from +2 to +3. No fixed chemical modification was inserted, but possible oxidation of methionines, hydroxylation on lysine and proline and deamidation at asparagines and glutamines were considered as variable modifications. Only proteins presenting two or more peptides were considered as positively identified. Individual ion score threshold provided by Mascot software to evaluate the quality of matches in MS/MS data was generally 34 for each samples. Spectra with Mascot score <10 were rejected.

5.5 Results

5.5.1 Analytical strategies to obtain the highest sequence coverage

In order to get the most from experiments on precious samples from Pompeii and Herculaneum, thoughtful considerations were made. While two or very few peptides can be enough to merely identify materials, more extended sequence coverage could be needed for more detailed analyses, such as the conservation state evaluation. The minimal protocols based on the enzymatic digestion of the proteins, directly performed in heterogeneous phase followed by mass spectrometric analyses of the released peptides, might be not always adequate. Alternative procedures are required, to obtain the highest number of peptides and thereof the highest sequence coverage. In the case of collagen from bones, the characterization of the degradation processes needs extensive sequence coverage, an in depth analyses might be needed to gain extremely important information that can be used for a wide range of purposes, from taxonomic analyses to diagenetic characterization.

The main purpose is to obtain information about the degradation and conservation state of the samples, studying deamidation of aminoacids Glutamine and Asparagine. In the particular

case of bones from Pompeii and Herculaneum, it is really interesting to investigate how the very peculiar type of burial has influenced the aging and the deterioration of these human remains.

To pursue a study on chemical modifications on specific aminoacids, it is necessary to successfully identify collagen in bones with a high *sequence coverage* (defined as the percentage of the proteins sequence covered by matched peptides). In this context, it is important to develop specific proteomic protocols that allow us to obtain optimal sequence coverage. In fact, while matching even few peptides (>2) is usually enough to identify a protein, the characterization of modifications of the sequence obviously requires to detect much of the sequence itself.

Three different pre-treatments of the bone samples were tested to improve sequence coverage. These pretreatment, which differ in the step of protein denaturation, were applied on two fragments of human bones, labeled 3A and 3D, from the archeological site of Pompeii, before the enzymatic digestion. The "EDTA pretreatment" is based on the use of the chelating agent Ethylenediaminetetraacetic acid (EDTA) which can remove Ca^{2+} ions present as structural part of the bone tissue by coordinating them, to make the access of the protease to proteins much easier. "UREA 6M pretreatment" and "HCl pretreatment" use respectively Urea and Hydrocloric acid to denature proteins and to make the hydrolysis sites more accessible for trypsin. Peptide mixtures obtained after enzymatic digestion were analyzed by mass spectrometry methodologies LC/MS-MS and typical proteins of bones, *Collagen alpha-1(I)* and *Collagen alpha-2 (I)*, were identified using the search engine *MASCOT*.

The sequence coverage of these two proteins were calculated for each aliquot of samples 3A and 3D, treated with the different protocols.

Table 5.1: Sequence coverage (%) of Collagen alpha-1(I) and Collagen alpha-2(I) from the digestion of both samples 3A and 3D with different pretreatment protocols followed by LC-MS/MS analysis. Sequence coverage was calculated on the sequence of the mature form of the protein. Proteins were identified in the UniProt database with Mascot MS/MS Ion Search software, with Homo sapiens as taxonomic restriction, with methionine oxidation, formation of pyroGlu at the N-terminus of Gln, deamidation (N, Q), and hydroxylation (K, P) of peptides as variable modifications.

	3A	**3A EDTA**	**3A HCl**	**3AGu HCl**	**3D**	**3D EDTA**	**3D HCl**	**3D GuHCl**
Collagen alpha-1(I)	68,4%	90,8%	44,4%	63.39%	57,4%	66,30%	63,90%	54.3%
Collagen alpha-2(I)	62,9%	76,1%	39,6%	52.20%	55,0%	48,90%	49,90%	50.2%

It can be noted that even with the minimally invasive protocol, where the bone fragment is simply deepened in a trypsin solution, the number of peptides detected can be high enough to obtain good sequence coverage (68% and 63% for COL1α1 and COL1α2, respectively).
Sequence coverage, obtained using GuHCl pretreatment, is comparable with the ones obtained using minimally invasive protocols, while data obtained with HCl pretreatment are even below expectation.
However, when decalcification is carried out with EDTA 0.5M, nearly the whole sequences can be nearly verified, reaching 91% and 76% of sequence coverage for mature COL1α1 and COL1α2, respectively. (36)
Therefore, decalcification with EDTA seems to be the best choice to map collagen with a really high sequence coverage.

5.5.2 Pretreatments and deamidation

It is widely accepted that deamidation can be considered as a biomolecular marker of deterioration and natural aging of proteins. Deamidation is also a delicate modification from a purely technical point of view, since it induces a mass shift of only 0.98 Da. Fragmentation spectra, however, not only clearly rules out any doubt about the modification but also localize the deamidation site within the peptide sequence.

However, deamidation is a delicate modification, since it is strongly influenced by several parameters such as pH and temperature.

The extraction protocol is therefore critical. Moreover, an extraction protocol as mild as possible can be crucial when we search for unpredicted modifications induced by

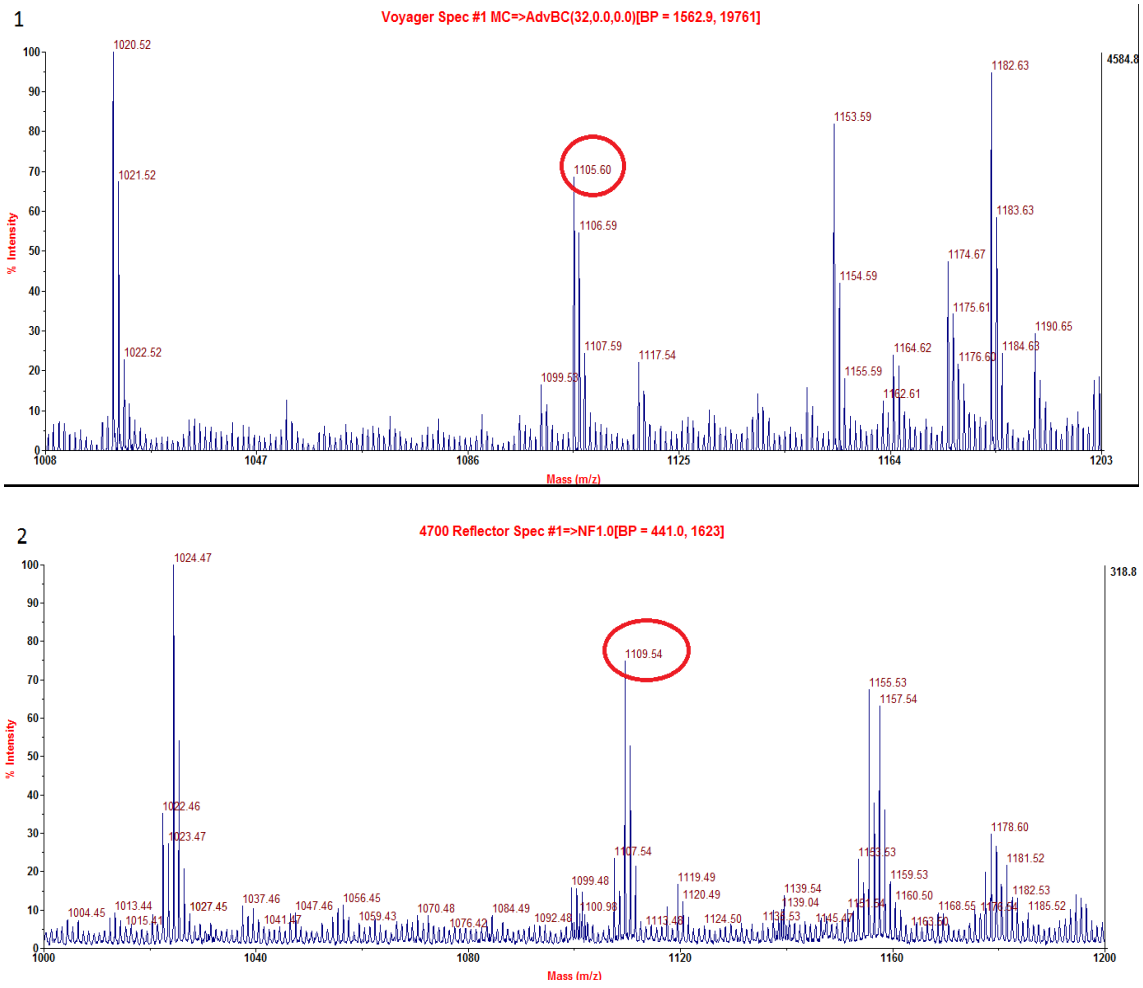
aging/conservation processes, since our knowledge of alterations occurring to proteins, because of diagenetic events as well as environmental factors is not yet complete.

Deamidation can occur also as a by-product of sample preparation. To consider deamidation as a signature of aging, its genuine pre- extraction origin must be verified. (37,38)

H_2 ^{18}O labelling can be used to assess any deamidation occurring during digestion process Every sample preparation step was carried out with and without H_2 ^{18}O and all pretreatments, reported above, were tested.

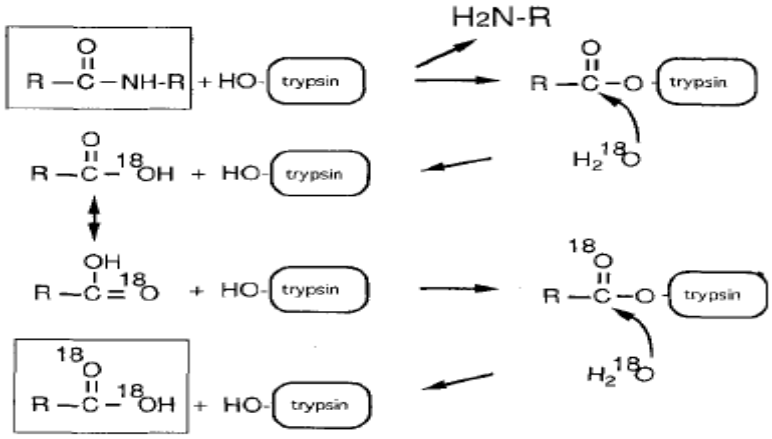
A most relevant example is shown below:

Figure 5.5: (1) MALDI-TOF spectra of the sample 3A. Peptide mixture was obtained by hydrolysis in heterogeneous phase after "UREA 6M" pre-treatment and followed by sample cleaning by reverse phase chromatography on ZipTip C18 cartridges. (2) MALDI-TOF spectra of the sample 3A. Peptide mixture was obtained by an hydrolysis in heterogeneous phase, preceded by "UREA 6M with water labeled with ^{18}O" pre-treatment and followed by reverse phase chromatography.



Signals marked with a red ring in figure 5.5 belong to the no deamidated peptide: R.GVQGPPGPAGPR.G of collagen alpha-1 (I). It is possible to observe a mass shift of 4 Da in the latter spectra due to the incorporation of two molecules of water labeled with ^{18}O, according to the hydrolysis reaction schematized in figure 5.6. (39)

Figure 5.6: scheme of enzymatic hydrolysis reaction with trypsin.

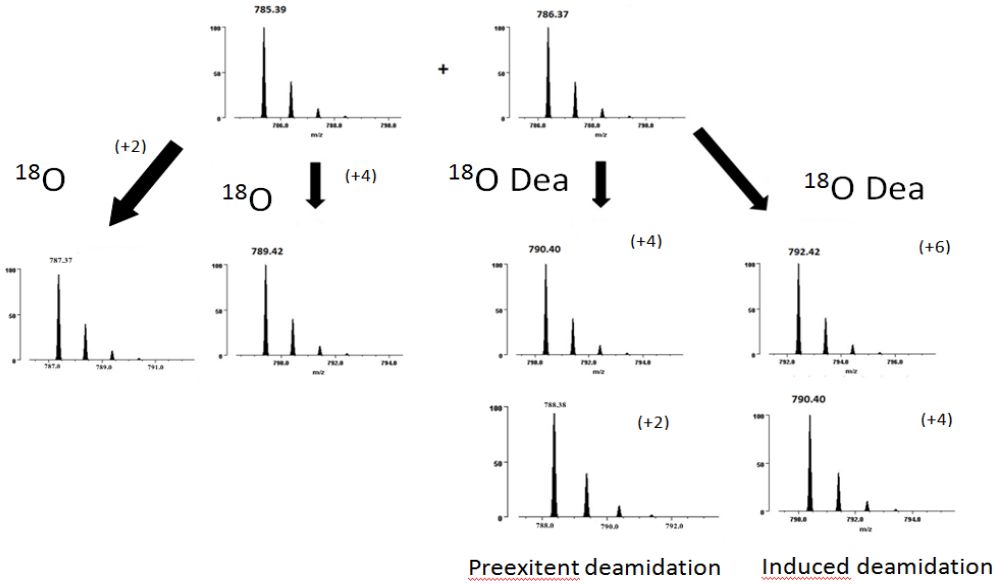


Enzymatic hydrolysis may cause mass shifts of 2 and/or 4 Da, depending on the number of water molecules incorporated during the reaction.

If the peptide is naturally deamidated, because of aging processes, we will also observe a typical pattern with the monoisotopic and the second peaks of comparable intensity in both spectra.

If deamidation is induced by treatments, we will notice additional mass shifts caused by incorporation of ^{18}O throughout the deamidation reaction. A scheme of these events is shown in figure 5.7, taking as example the peptide R.GDQGPVGR.T

Figure 5.7: scheme of isotopic patterns of the peptide R.GDQGPVGR.T identified in collagen alpha-1(I) in sample 3A by MALDI-TOF analysis.



For bone samples 3A and 3D treated with water labeled with ^{18}O and analyzed by MALDI-TOF methodology, the deamidation observed is preexistent in samples and not induced by sample treatment. Mass shifts are not consistent with incorporation of ^{18}O at the deamidation site during the treatment and for this reason, the protocols applied were considered as safe to study deamidation and its connection to aging.

5.5.3 Deamidation and burial under the ashes of the Vesuvius eruption

The aim of this chapter is to search for molecular evidences in the bones between the types of conservation. In particular, we focused our attentions on bone remains from Pompeii and Herculaneum.

As previously discussed, during the eruption in A.D. 79, Pompeii was covered by a layer "only"6 meters deep consisting of ash and cinder and the ash cases contained skeletons that were kept almost intact, that once hardened kept the shape of the body, allowing to prepare the plaster castes, whereas Herculaneum was destroyed and buried by lava and mud and the bodies were completely burnt.

Different type of conservation can be reflected in different chemical modifications. The difference in the level of deamidation can be an important marker of the stateo f conservation. Deamidation can result from aging, degradation and environmental storage conditions.

Furthermore, the study of others bone remains, from Scalandrone bay, that had not undergone the experience of the eruption, provide the comparison to another type of conservation.

The analysis of all the samples provided data to evaluate molecular features of bones conservation.

It is widely discussed above that deamidation can be considered as a biomolecular marker of deterioration and natural aging in proteins. Fragmentation spectra, obtained by LC-MS/MS analyses, not only clearly rule out any doubt about the actual presence of the modification but also localize the deamidation site within the peptide sequence.

The sites of deamidation observed in the different samples are shown in Tables A1 to A6 (Appendix A), and indicate that in the samples from Pompeii, only few deamidations can be observed, in respect to the extensive deamidation observed in the samples from Herculaneum and Scalandrone bay, where the majority of possible deamidation sites are indeed modified.

These data suggest that deamidation can actually be considered as an important marker of the conservation state: reflecting aging, degradation and environmental storage conditions as literature suggest.

The promising results were explored by analyzing on a statistical basis the 21 samples from Herculaneum, the 14 samples from Scalandrone bay and the only 2 samples from Pompeii. More sample from Pompeii will be needed to complete the analysis.

Percentage ratios between detected deamidation and detected sites (DD/DS) and ratio between detected deamidation and hypothetical deamidation sites (DD/HDS) for each samples, were reported in the appendix A (Tables A7 to A9). Average of these percentage ratios were calculated and are shown in tables 5.2 and 5.3.

Table 5.2: Average percentage ratios between detected deamidation and detected sites (DD/DS) and ratios between detected deamidation and hypothetical deamidation sites (DD/HDS) for samples from Herculaneum, Scalandrone and Pompeii in Collagen alpha 1(I).

Collagen alpha 1(I)	Average DD/DS (%)	Average DD/DHS (%)
Herculaneum	49.59	25.78
Pompeii	28.75	17.08
Scalandrone bay	49.40	27.53

Table 5.3: Average percentage ratios between detected deamidation and detected sites (DD/DS) and ratios between detected deamidation and hypothetical deamidation sites (DD/HDS) for samples from Herculaneum, Scalandrone and Pompeii in Collagen alpha 2(I).

Collagen alpha 2(I)	Average DD/DS (%)	Average DD/DHS (%)
Herculaneum	48.42	18.43
Pompeii	30.05	15.22
Scalandrone bay	62.52	20.50

It is possible to notice differences in the percentage of deamidated Asparagine and Glutamine identified in these tree different groups of samples; in particular, Pompeii samples show fewer deamidated sites.

Moreover, percentage ratios between detected deamidation and detected sites (DD/DS) and ratios between detected deamidation and total numbers of samples for each groups (DD/TS) were calculated and graphically reported. Data obtained are shown in figure 5.8 to 5.11.

Figure 5.8: Collagen alpha 1(I), the graph reports the percentage ratios between detected deamidation and detected sites (DD/DS) for each possible deamidation site. In light blue samples from Herculaneum, in green samples from Pompeii and in red samples from Scalandrone.

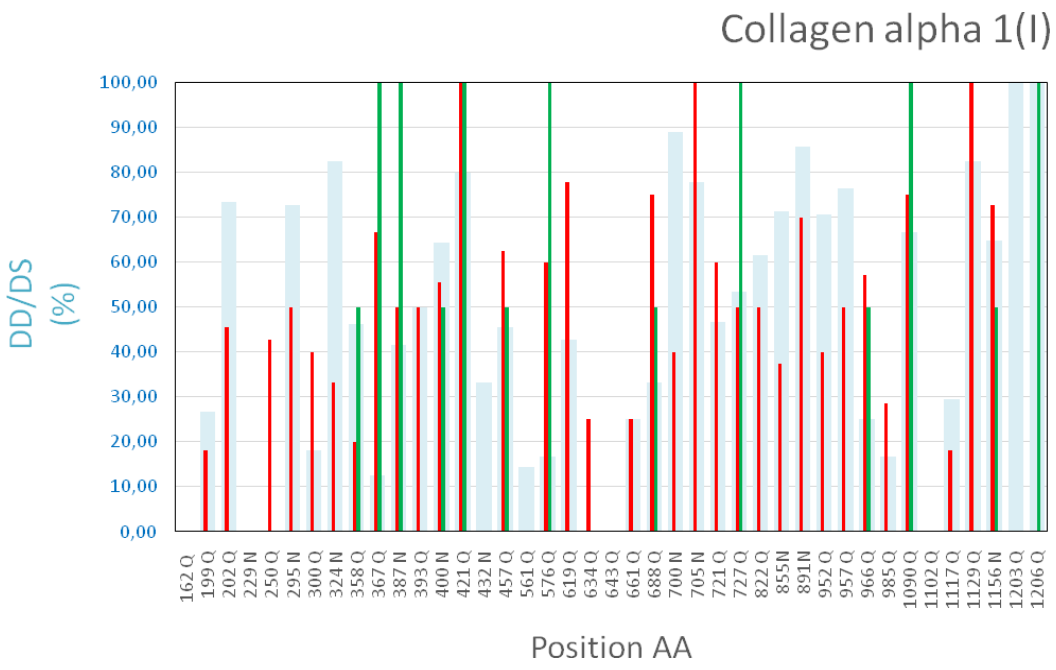


Figure 5.9: Collagen alpha 1(I),the graph reports the percentage ratios between detected deamidation and total numbers of samples for each groups (DD/TS) for each deamidation site. In light blue samples from Herculaneum, in green samples from Pmpeii and in red samples from Scalandrone.

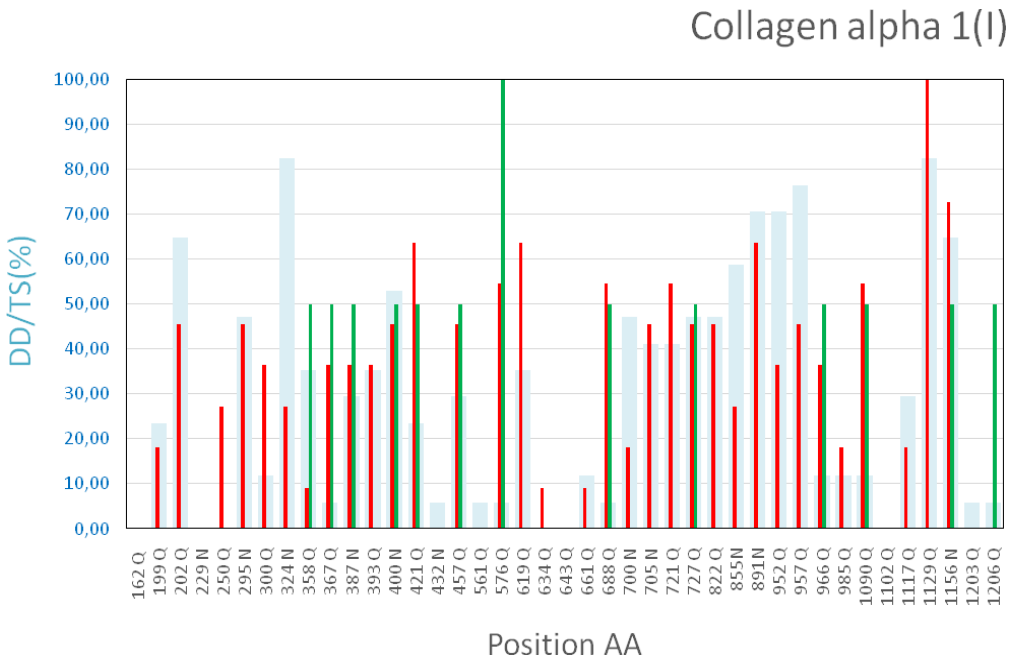


Figure 5.10: Collagen alpha 2(I), the graph reports the percentage ratios between detected deamidation and detected sites (DD/DS) for each possible deamidation site. In light blue samples from Herculaneum, in green samples from Pompeii and in red samples from Scalandrone.

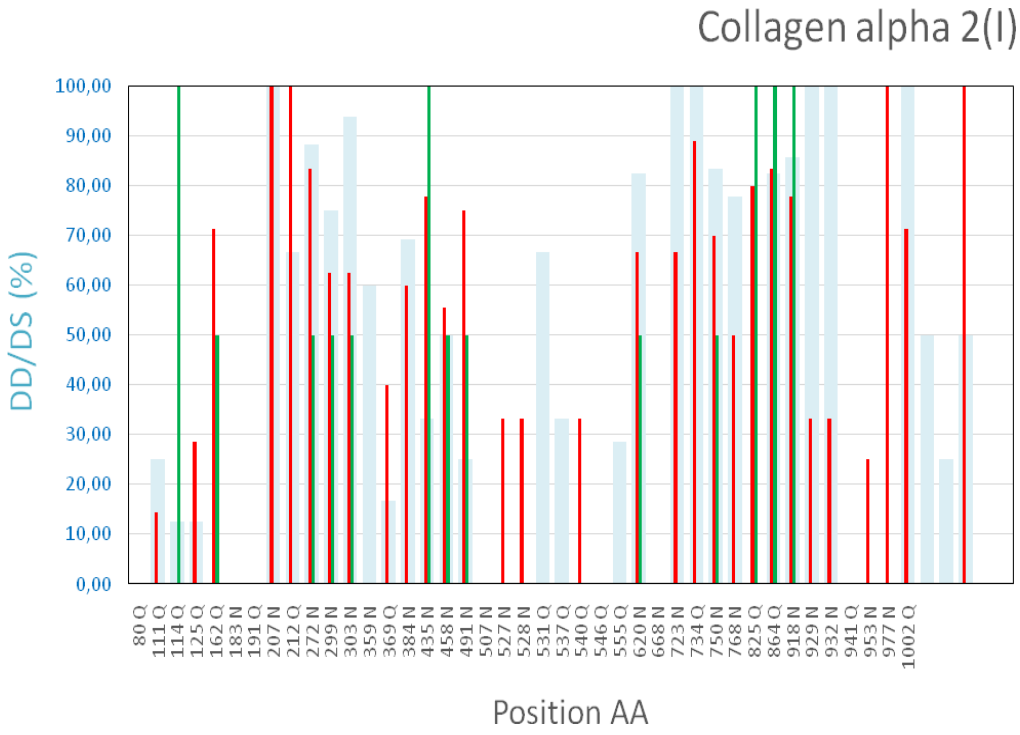
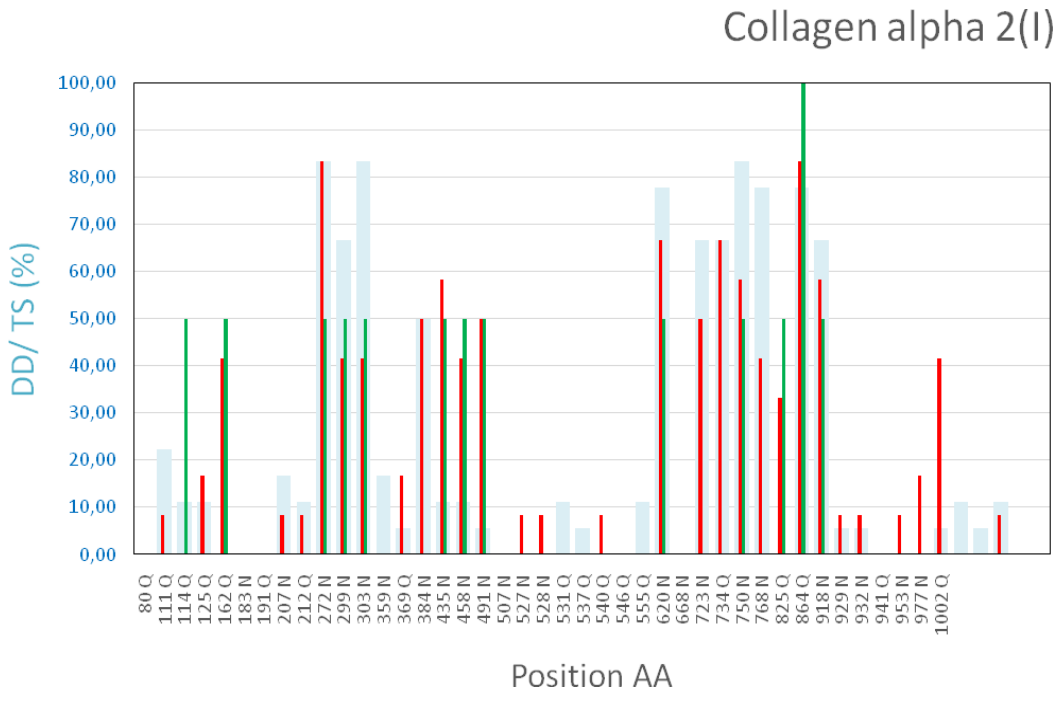


Figure 5.11: Collagen alpha 2(I),the graph reports the percentage ratios between detected deamidation and total numbers of samples for each groups (DD/TS) for each deamidation site. In light blue samples from Herculaneum, in green samples from Pmpeii and in red samples from Scalandrone.



In the graphs displayed above, the percentage ratio are reported as function of deamidation site. These data suggest that Pompeii samples contain fewer deamidated sites. Indeed, we can see that only few Asparagine and Glutamine are deamidated; Pompeii's percentage ratios, that seem to indicate a definitively lower deamidation, nevertheless require the analysis of more samples, to corroborate the hypothesis of considering deamidation as the actual molecular marker of better conservation state of the skeleton in Pompeii. Data showed for Herculaneum and Scalandrone bay samples, are instead significatively comparable.

Based on these data, it can be hypothesized that the reduced deamidation in the samples from Pompeii could be linked to its particular burial, where hot temperature were high enough to kill but not to burn, and hot ashes preserved bones from extensive degradation processes during centuries.

5.6 Conclusions

Among the archaeological remains, bones are most interesting. Bones, indeed, can be an abundant reservoir of ancient biomolecules due to natural resistance to post-mortem and collagen is the dominant protein in mineralised tissue. (1,2) The effect of aging, environmental and storage conditions at the molecular level can be analysed and studied as chemical modifications induced in proteins. In particular, degradation is visible at amino acid residue level and deamidation of asparagine and glutamine. (4,5)

This project was focused on the analysis of the degradation of the proteins, at the molecular level, due to the storage conditions and the main purpose was to obtain information about the degradation and conservation state of the samples, studying deamidation of aminoacids Glutamine and Asparagine as a possible reporter of degradation state.

To pursue a study on chemical modifications, it is necessary to successfully identify collagen in bones with a high *sequence coverage*. In this context, it was important to develop specific proteomic protocols that allow us to obtain optimal sequence coverage. In fact, while matching even few peptides (>2) is usually enough to identify a protein, the characterization of modifications of the sequence requires to detect as much as possible of the sequence itself.
Three different pre-treatments of the bone samples were tested to improve sequence coverage and the results obtained showed that decalcification with EDTA seems to be the best choice to map collagen with a really high sequence coverage.
Moreover, the extraction protocol is critical; deamidation can occur also as a by-product of sample preparation. To consider deamidation as a signature of aging, its genuine pre-

extraction origin must be verified. (37,38) H_2 ^{18}O labelling was used to assess any deamidation occurring during digestion process. Every sample preparation step was carried out with and without H_2 ^{18}O. The analysis demonstrated that the protocols applied were considered safe to study deamidation and its connection to aging.

The main issue of this chapter was to search differences between bones from Herculaneum and Pompeii at molecular level that could be linked to the different death and burial conditions that have been experienced by humans in the two cities.

The sites of deamidation observed in the different samples indicate that in the samples from Pompeii, only few deamidations can be observed, in respect to the extensive deamidation observed in the samples from Herculaneum and Scalandrone bay, our control of bones from deaths where the majority of possible deamidation sites are indeed modified.

Based on these data, it can be hypothesized that the reduced deamidation in the samples from Pompeii could be linked to its particular burial, which could have protected it from degradation processes during the centuries.

5.7 References

1. Hill RC, Wither MJ, Nemkov T, Barret A, D'Alesandro A, Dzieciatkowska M, Hansen (2015) KC.*Preserved proteins from extinct Bison latifrons identified by tandem mass spectrometry; hydroxylysine glycosides are a common feature of ancient collagen.* Molecular Cell Proteomics (14) 1946-1958.
2. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. Rapid Communication in Mass Spectrometry (28) 605-615.
3. Dobberstein RC, Collins MJ, Craig OE, Taylor G, Penkman KEH, Ritz-Timme S (2009) Archaeological collagen: Why worry about collagen diagenesis? Archaeological and Anthropological Sciences, (1) 31–42.
4. Solazzo C, Wilson J, Dyer JM, Clerens S, Plowman JE, Von Holstein I, Walton Rogers P, Peacock EE, Collins MJ (2014) *Modeling Deamidation in Sheep alpha-keratin peptides and application to archaeological wool textiles.* AnalyticalChemistry (86) 567-575.
5. Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, et al. (2011) Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. Anal. Chem. (83) 2056–206.
6. Robinson NE, Robinson B (2001) Molecular clocks. Proc. Natl. Acad. Sci. U. S. A. (98) 944–949.
7. Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, et al. (2011) Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. Anal. Chem. (83) 2056–2064
8. Collins MJ, Riley MS, Child AM (1995) A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. Journal of Archaeological Science (22) 175–183.
9. Collins MJ, Hiller J, Smith CI, Roberts JP, Prigodich RV, Wess TJ, Csapo J, Millard AR, Turner-Walker G (2002) The survival of organic matter in bone: a review. Archaeometry (44) 383-394.

10. Trueman CN, Martill DM The long-term survival of bone: the role of bioerosion. Archaeometry (44) 371-382.
11. Turner-Walker G (2008) The Chemical and Microbial Degradation of Bones and Teeth. Advances in Human Palaeopathology.
12. Rho J, Kuhn-Spearing L, Zioipos P (1998) Mechanical properties and the hierarchical structure of bone. Medical engineering and physics. (20) 92-102.
13. Gordon C., Buikstra JE (1981) Soil pH, Bone Preservation, and Sampling Bias at Mortuary Sites. American Antiquity (46) 566−571.
14. Covington AD, Song L, Suparano O, Koon HEC, Collins MJ, (2010) Link-lock: the mechanism of stabilising collagen by chemical reactions. World Leather, (25) 35–43.
15. Marieb M, Hoehn K, (2007) Human Anatomy and Physiology, Pearson Education.
16. Viguet-Carrin S, Garnero P, Delmas PD (2006) The role of collagen in bone strength. Osteoporosis International. (17) 319–336.
17. Whitford D, (2008) Proteins, structure and function. In John Wiley and Sons Ltd, 92–101.
18. Shoulders M D, Raines RT (2009) Collagen structure and stability. Annual Review of Biochemistry (78) 929–58.
19. Rich A, Crick FHC (1961). The molecular structure of collagen. Journal of Molecular Biology, (3) 483–506.
20. Berisio R, Vitagliano L (2002) Crystal structure of the collagen triple helix model 3. Protein science 262-270.
21. Orgel J, Miller A, Irving T (2001) The in situ supermolecular structure of type I collagen. Structure (9) 1061–1069.
22. Burjanadze (2000). New analysis of the phylogenetic change of collagen thermostability. Biopolymers (53) 523-528.
23. Gorres K, Raines R (2010) Prolyl 4-Hydroxylase. Critical reviews in biochemistry and molecular biology (45) 106-124.
24. Bagley M (2016) Mount Vesuvius & Pompeii: Facts & History. Live Science
25. Giacomelli L, Perrotta A, Scandone R, Scarpati C (2003)The eruption of Vesuvius of 79 AD and its impact on human environment in Pompei. Episodes (26) 234-237.
26. Blong RJ (1984) Volcanic hazards. Academic Press.
27. Sigurdsson H, Carey S, Cornell W, Pescatore T (1985) The eruption of Vesuvius in A.D. 79. Nat. Geo. Res., (1) 332-387.

28. Capasso L, Capasso L, Caramiello S, D'Anastasio R, Di Domenicantonio L, Di Fabrizio A, Di Nardo F, La Verghetta M (2000) Paleobiologia della popolazione di Ercolano (79 d.C.). Recenti Progessi in Medicina. (91) 288-296.
29. Mastrolorenzo G, Petrone P, Pagano P, Incoronato A, Baxter PJ, Canzanella A, Fattore L (2001) Herculaneum victims of Vesuvius in AD 79. Nature, (769).
30. Araki N, Moini M (2011) Age estimation of museum wool textiles from Ovis aries using deamidation rates utilizing matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry (25) 3396–3400.
31. van Doorn NL, Hollund H, Collins MJ, (2011) A novel and non-destructive approach for ZooMS analysis: Ammonium bicarbonate buffer extraction. Archaeological and Anthropological Sciences. (3) 281–289.
32. Hurtado PP, O'Connor PB (2012) Deamidation of collagen. Analytical chemistry (84): 3017–3025.
33. Wilson J, van Doorn NL, Collins MJ (2012). Assessing the extent of bone degradation using glutamine deamidation in collagen. Analytical Chemistry, (84) 9041–9048.
34. Mikšík I (2014). Prince Cangrande's Collagen: Study of Protein Modification on the Mummy of the Lord of Verona, Italy (1291–1329 AD). Chromatographia
35. Geiger T, Clarke S (1987). Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. Journal of Biological Chemistry, (262) 785.
36. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.
37. Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, et al. (2011) Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. Anal. Chem. (83) 2056–2064
38. Wilson J, Van Doorn NL, Collins MJ (2012)Assessing the extent of bone degradation using glutamine deamidation in collagen. Anal. Chem. (84) 9041–9048.
39. Rawlings ND, Barrett AJ(1994) Families of serine peptidases. Meth. Enzymol. Methods in Enzymology (244) 19–61.

6. Proteomic strategies for the characterization of membrane threads.

Nowadays the knowledge of the materials used in a work of art is of the greatest significance to give an insight in the historical context of objects and artists, and to develop appropriate conservation and/or restoration treatments. (1)

The study of the history of clothing and textiles traces the availability and use of textiles and other materials and the development of technology for the making of clothing over human history as well as it can be a signature of the trade history. The wearing of ancient clothing is exclusively a human characteristic and it is a feature of most human societies. Clothing and textiles have been important in human history and they reveal not only the materials available to a civilization, but there is also a social significance of the final product, (2) and materials used for clothes reveal the social status of the person itself . From the ancient times to the present, methods of textile production have continually evolved, and the choices of available textiles have influenced how people carried their possessions, how they clothed themselves, and how they decorated their surroundings. (2)

This chapter deals with the characterization of a textile membrane thread, covered by wire, from a drape found in the tomb of Emperor Henry VII of Luxembourg.

6.1 Henry VII and a unique silk cloth found in his coffin

Resting in Pisa Cathedral, the remains of German king and Holy Roman emperor Henry VII of Luxembourg (1275-1313), were exhumed with the aim of getting more insights into the emperor's physical features and cause of his death, in 2014. (3)

The opening of the sarcophagus revealed the presence of a medieval treasure. A unique silk cloth has been found in the tomb of Henry VII among bones and what remains of his boiled head. The sarcophagus also reveals a crown, a scepter and a unique, 10-foot-long silk cloth. (3)

Celebrated as the "alto Arrigo" (high Henry) in Dante's Divine Comedy, Henry is best remembered for his struggle to reestablish imperial control over the city-states of the 14th--century Italy. He was crowned King of Germany in 1308 and two years later he descended to Italy with the aim of pacifying destructive disputes between Guelf (pro-papal) and Ghibelline (pro-imperial) factions. His goal was to be crowned emperor and restore the glory of the Holy Roman Empire. After meeting strong opposition among anti-imperialist Guelf lords, Henry

entered Rome by force, and was indeed crowned Holy Roman Emperor on the 29 of June 1312. (3)

Henry VII died just a year after his coronation; he failed to defeat opposition by a secular Avignon papacy, city-states and lay kingdoms. Henry died prematurely at Buonconvento, near Siena, on the 24 of August 1313. Rumors about his poisoning began to spread. The emperor's body was hastily buried; two years later, he was reburied in the Cathedral of Pisa.

The emperor's followers burned his body because they had no time to treat the body for transportation; they detached the head and boiled it. His bones were kept in wine to preserve them better. (3)

Researchers found into the Emperor’s coffin ashes and bones showing signs of burning. Anthropological examination has revealed a skeleton that belonged to a 40-years-old man which was 5 feet, 5 inches tall, and who was found kneeling in prayer. Analysis has so far revealed a high concentration of arsenic in the bones, which could support the poisoning theory, although many drugs at that time were arsenic based. (3)

As for the silk cloth, it is unclear how it ended up in the coffin. It's a unique example of the noble production of silk textiles dating back to the beginning of the 14th century.

More than 10 feet long and 4 feet wide, the exquisitely woven cloth features horizontal bands of 4 inches showing alternating colors, a reddish nut-brown (originally red) and blue.

There are blue bands embroidered in gold and silver with pairs of lions facing each other, while an elaborated monochromatic tone-on-tone decoration, currently indecipherable, is visible on the reddish bands. A crimson strip edged with yellow, placed at the top of the piece of fabric, bears traces of an inscription. Other unique features are the finished edge along the length of the fabric -to keep it from unraveling - and the checked bands at the shorter ends marking the beginning and end of the piece. The lions, the most characteristic emblem of sovereignty, as well as other decorations probably symbolizing power, indicate a clear link to the emperor. (3)

What makes this cloth unique is its size, the very high level of craftsmanship and its amazing preservation. According to Gale Owen-Crocker, professor of Anglo-Saxon Culture at the University of Manchester and an expert on medieval clothing and textiles, such silks were treasured possessions of the rich and royal. (3)

The drape found in the sarcophagus is probably a thread composed by animal gut coated with gold.

6.2 Metal threads decorating textiles

The decoration of textiles with metals can be traced thousands of years. The first written record of decoration with gold is probably a quotation from the Bible, which describes the making of Aaron's vestment for service around 12-13th centuries BC. (4)

The value of metals, especially gold, have always attracted men. Besides wearing gold jewels, they haves woven gold into textiles; they used gold for embroidering his clothes; they have decorated his churches and secular textiles with this precious metal. (4)

The metal threads that were used in embroidering and weaving were made in a variety of ways. They consisted of metal strips that were used either directly or, more often, in the production of composite threads, in which the metal strip was wound around a fibrous core of silk, linen, cotton or other yarns. Pure gold, gold alloyed with silver, gilded or gilt-silvered copper and gold-like copper alloys (e.g. pinchbeck alloy or brass) were used as the materials of the metal strips or wires. Wire was directly used or for winding around such cores. (4,5)

Membrane threads form a special group of gold threads: they were once made by gilding leather, parchment, paper or animal gut; and then, they were made cutting the gilded material into strips. These strips were also utilized in weaving and embroidery. For these purposes, they were, either used like solid metal threads, flat or wound around a fibrous core. (4,5)

These membranaceous threads were made with skin or gut of bovines, sheep, goats, pigs or fowls. Current scientific research is focused on metals analysis and only few experiments are about the identification of animal species. (6)

The techniques of productions of these membranaceous threads are not well known. These kind of studies hasn’t been made yet.

The characterization of gut’s animal species has been made considering that some components of connective tissue are amazingly conserved, especially proteins of Collagen.

Despite the effects of aging and deterioration, Collagens are well preserved and highly conserved throughout the evolution of species due to their, essential biological function, as well as, their peculiar sequence as indicate in the chapter 5 .

6.3 Proteins analysis

So far, the analysis of the organic materials in membranaceous threads was limited to its classification in one of three categories: skin, parchment and animal gut, and it is based on its macroscopic appearance. These analyses do not allow to identify the true nature of the sample. (4)

Other methods in cellular and molecular biology (light microscopy, electron microscopy, DNA amplification) have been applied to the study of the proteinaceous substrate of organic metal threads in textiles with the aim to identify the organs and animal species used. The identification of the species involved by the use of molecular biology is still hampered by a number of technical difficulties related to the extent of fragmentation of the DNA molecules.(6)

The application of DNA techniques are limited and DNA survival is restricted whereas proteins are more stable than nucleic acids and can persist much longer. (7)

Mass spectrometry allows for unambiguous identification of peptides and proteins, and even in complex mixtures, species-specific identification becomes feasible with minimal sample consumption. Determination of the peptides is commonly based on theoretical cleavage of known protein sequences and on comparison of the expected peptide fragments, which ones are found in the MS/MS spectra. In this approach, complex computer programs, such as Mascot, perform well identifying known proteins, but they fail when protein sequences are unknown or incomplete. (8)

A main difficulty in identifying proteins is indeed the genomic data. In shotgun proteomics, proteins are routinely identified by matching experimental MS/MS spectra of enzymatic digests of protein samples to simulate spectra from protein databases, usually derived from genomic sequences. Identification in database search procedures is relied on how well experimental spectra fits to theoretical spectra obtained from sequences that are present in databases. Standard database search algorithms fail to identify peptides that are not exactly contained in a protein database. Identification of the specific species will then depend on the detection of two or few more peptides that are conserved among the different species, despite possible differences in other peptides, which will be ignored in a standard database searches. (7, 9, 10)

6.4 Material and Methods

6.4.1 Samples

The sample analyzed was a textile membrane thread from a drape found in the tomb of Emperor Henry VII of Luxembourg. It was provided by the laboratory at the Department of Chemistry and Industrial Chemistry, University of Pisa of Prof. MP Colombini. This thread is composed by animal gut coated with gold.

6.4.2 Proteomic analyses

Reagents. Ammonium hydrogen carbonate (Ambic), urea and TPCK-treated trypsin were from Sigma; Formic acid and Acetonitrile (ACN) were purchased from Baker; deionized water was obtained from Millipore cartridge equipment.

Sample treatment. Few μg of thread were treated in heterogeneous phase with trypsin with the minimally invasive protocol with UREA 6M pretreatment as described in the chapter 4 and the resulting peptide mixtures were analysed by LC-MS/MS.

The acquired MS/MS spectra were transformed in Mascot Generic files (.mgf) format and used to query the SwissProt database 2015_04 (548208 sequences; 195282524 residues), with *Chordata* as taxonomic restriction for protein identification with licensed version of MASCOT software (www.matrixscience.com) version 2.4.0 with the same parameters as reported in the chapter 4. No fixed chemical modification was inserted, but possible oxidation of methionine residues, deamidation at asparagines and glutamines and hydroxylation on Lysine and Proline were considered as variable modifications. Only proteins presenting two or more peptides were considered as positively identified. Individual ion score threshold provided by MASCOT software to evaluate the quality of matches in MS/MS data was 44. Spectra with MASCOT score of <10, having low quality were rejected.

6.5. Results and discussion

As outlined in the introduction, determining the species in the membrane threads can be useful for delineating human uses and habits and for defining economic conditions and social status of the specific person. So far, no investigation about these textile artifacts has been carried out.

To characterise the proteinaceous material in the thread, a fragment of a drape found in the tomb of Emperor Henry VII of Luxembourg was analysed by proteomics procedures.

The samples were subjected to enzymatic digestion with trypsin, and the obtained mixtures of peptides were purified and concentrated using reverse-phase C18 pipette tips and analysed by LC-MS/MS. Table1 reports the proteins identified in the sample:

Table 6.1: Proteins identified in the textile membrane thread in a drape from the tomb of Emperor Henry VII of Luxembourg. Aliquots were treated in heterogeneous phase with trypsin with the minimally invasive protocol with Urea pre-treatment and the resulting peptide mixtures were analyzed by LC-MS/MS. Raw data were used to search with MASCOT MS/MS Ion search. Swiss-Prot/Uniprot database, with Chordata as taxonomic restriction, with Deamidation of Glutamine and Asparagine, methionine oxidation, formation of pyroGlu at the N-terminus Gln and hydroxylation on Lysine and Proline of peptides as variable modifications.

Protein	Score	N° peptides	Sequence coverage (%)
Collagen alpha-2(I) chain	356	11	15%
Collagen alpha-1(III) chain	202	6	11%
Collagen alpha-1(I) chain	196	6	6%

The scores, the sequence coverage, the number of the identified peptides, provide a reliable identification of the material as animal gut as the proteinaceous constituent of the membranaceous thread. Above all, the presence of Collagen type III in the sample is strongly indicative of connective tissue, since type III is specific of reticular fibers, blood vessels and skin.

While few peptide are enough to provide reliable identification of the kind of proteins, for the attribution of the species of origin it's fundamental to detect proteotypic peptides, that are unique to the protein sequence and are specific for a single organism and not in common to other species.

Details of the identifications are given in Tables 6.2, 6.3 and 6.4

Table 6.2: Peptide profile of Collagen alpha-1(I) in the textile membrane thread in a drape from the tomb of Emperor Henry VII of Luxembourg. Peptide mixture was analyzed by LC-MS/MS. Raw data searched by MASCOT MS/MS Ion search. Swiss-Prot/Uniprot database, with Chordata as taxonomic restriction, with Deamidation of Glutamine and Asparagine, methionine oxidation, formation of pyroGlu at the N-terminus Gln and hydroxylation on Lysine and Proline of peptides as variable modifications. Individual ion scores >44 indicate identity or extensive homology. Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html).

Collagen alpha-1(I) chain OS=Bos taurus	
Score: 196 Number of peptides: 6	
Peptide	Score
R.GLPGER.G + Hydroxylation (P)	11
R.GSEGPQGVR.G	26
R.GVVGLPGQR.G + Hydroxylation (P)	30
R.GVQGPPGPAGPR.G + Hydroxylation (P)	34
R.GEPGPAGLPGPPGER.G + 3 Hydroxylation (P)	33
R.GSAGPPGATGFPGAAGR.V + 2 Hydroxylation (P)	45

Table 6.3: Peptide profile of Collagen alpha-1(III) in the textile membrane thread in a drape from the tomb of Emperor Henry VII of Luxembourg. Peptide mixture was analyzed by LC-MS/MS. Raw data searched by MASCOT MS/MS Ion search. Swiss-Prot/Uniprot database, with Chordata as taxonomic restriction, with Deamidation of Glutamine and Asparagine, methionine oxidation, formation of pyroGlu at the N-terminus Gln and hydroxylation on Lysine and Proline of peptides as variable modifications. Individual ion scores >44 indicate identity or extensive homology. Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html). Proteotypic peptides that are unique to the *Bos Taurus* sequence were in bold.

Collagen alpha-1(III) chain OS=Bos taurus	
Score: 202 Number of peptides: 6	
Peptide	Score
R.GPPGAGGPPGPR.G + 2 Hydroxylation (P)	22
R.GLAGPPGMPGAR.G + Oxidation (M); 2 Hydroxylation (P)	36
R.GSPGGPGAAGFPGGR.G + 3 Hydroxylation (P)	40
R.GPPGPPGTNGVPGQR.G + Deamidated (NQ); 3 Hydroxylation (P)	15
R.GVAGEPGRNGLPGGPGLR.G + Deamidated (NQ); 3 Hydroxylation (P)	26
R.GAPGPQGPPGAPGPLGIAGLTGAR.G + 3 Hydroxylation (P)	45

Table 6.4: Peptide profile of Collagen alpha-2(I) in the textile membrane thread in a drape from the tomb of Emperor Henry VII of Luxembourg. Peptide mixture was analyzed by LC-MS/MS. Raw data searched by MASCOT MS/MS Ion search. Swiss-Prot/Uniprot database, with Chordata as taxonomic restriction, with Deamidation of Glutamine and Asparagine, methionine oxidation, formation of pyroGlu at the N-terminus Gln and hydroxylation on Lysine and Proline of peptides as variable modifications. Individual ion scores >44 indicate identity or extensive homology. Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (http://www.matrixscience.com/help/interpretation_help.html). Proteotypic peptides that are unique to the *Bos Taurus* sequence were in bold.

Collagen alpha-2(I) chain OS=Bos taurus	
Score: 356 Number of peptides: 11	
Peptide	Score
R.GLPGER.G + Hydroxylation (P)	11
R.GVVGPQGAR.G	27
R.IGQPGAVGPAGIR.G + Hydroxylation (P)	37
R.GIPGPVGAAGATGAR.G + Hydroxylation (P)	35
R.GIPGEFGLPGPAGAR.G + 2 Hydroxylation (P)	36
R.GDGGPPGATGFPGAAGR.T + 2 Hydroxylation (P)	24
R.GSTGEIGPAGPPGPPGLR.G + 2 Hydroxylation (P)	37
R.GPNGDSGRPGEPGLMGPR.G + Oxidation (M); 2 Hydroxylation (P)	13
R.GERGPPGESGAAGPTGPIGSR.G + Hydroxylation (P)	22
R.GLPGVAGSVGEPGPLGIAGPPGAR.G + 3 Hydroxylation (P)	93
R.GAPGAIGAPGPAGANGDRGEAGPAGPAGPAGPR.G + 2 Hydroxylation (P)	18

A further check of the quality of the experimental fragmentation spectrum of the proteotypic peptides (reported in bold) was also carried out, and it allowed us to ascribe the collagen attribution to bovine species.

Mascot output attributes identified collagens also to different species, but with different score, different number of the identified peptides and different percentages of sequence coverage. No unique peptides were observed for each species except *Bos taurus*.

In most instances, the analysis of the peptides shows that the peptides reported are the same for the same type of collagen in different species. This is expected on the basis of the conservation of the collagen sequences among different species.

As reported in literature, these textile artefacts were usually made of *Bos taurus*, *Ovis aries*, *Capra hircus*, *Sus scrofa* or *Gallus gallus.* (7)

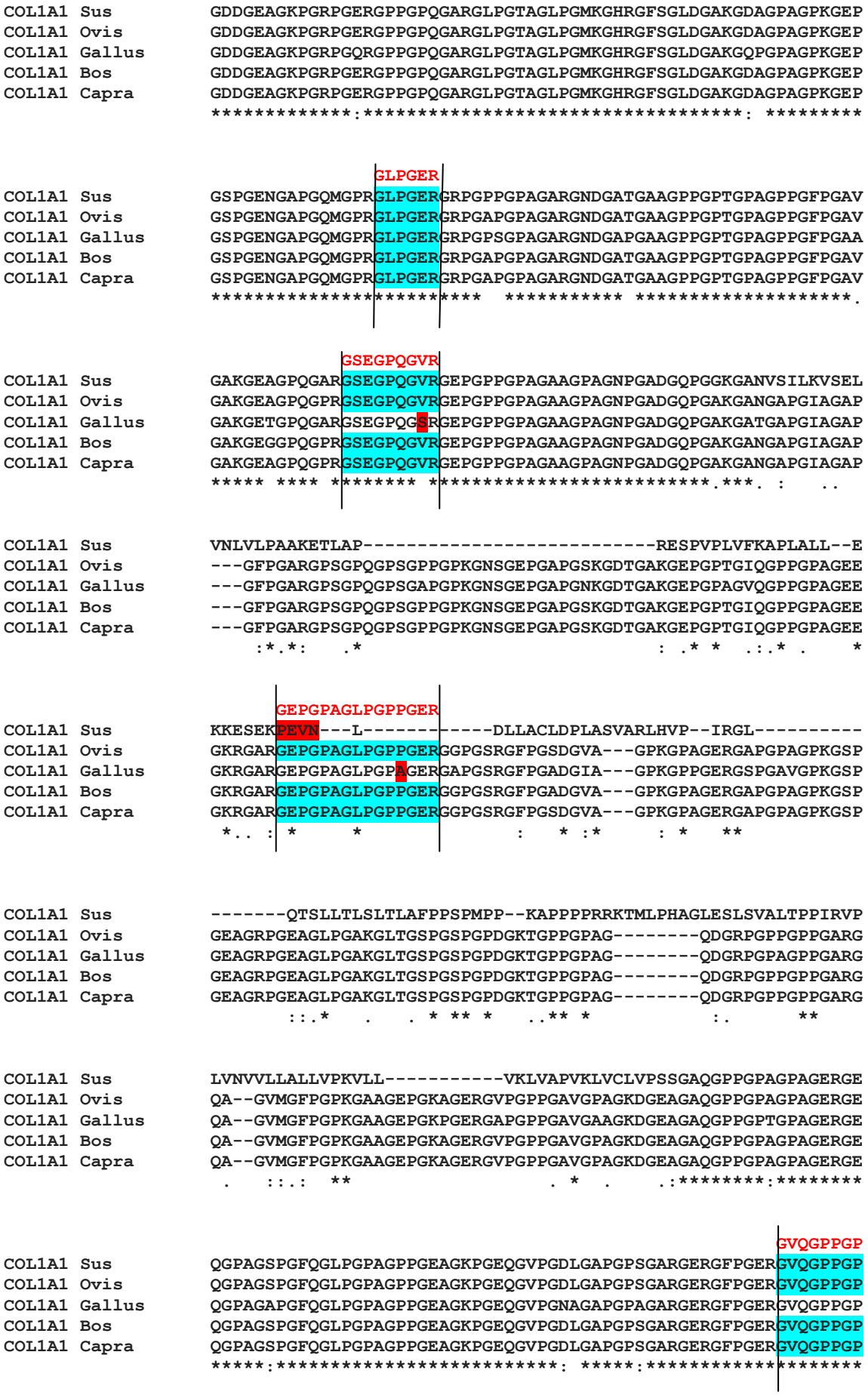
To confirm the identification obtained by Mascot and to exclude the presence of collagen belonging to other species usually used, a multiple alignment with the collagen sequences of these species was therefore created.

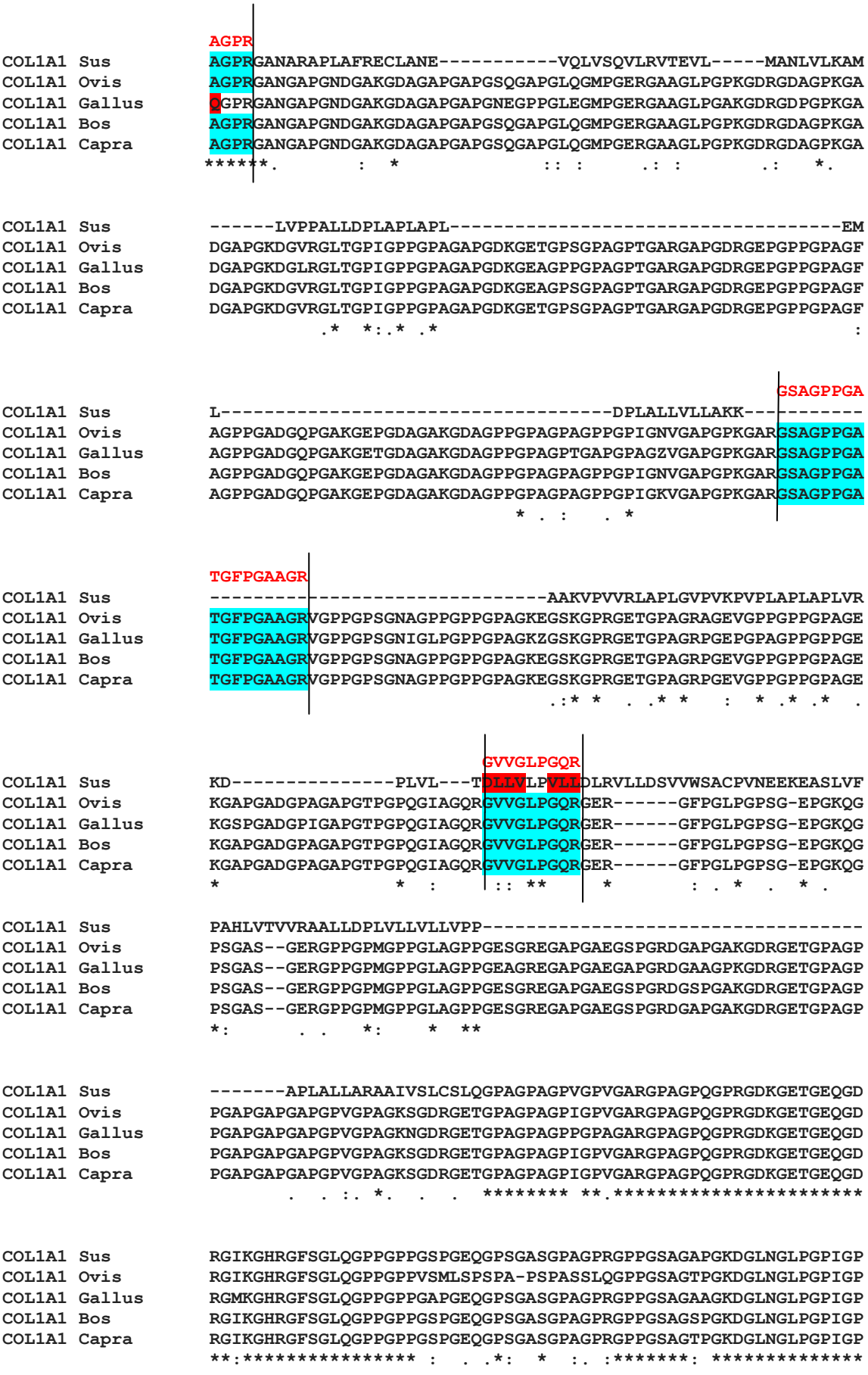
The multiple alignments of collagen alpha-1(I), collagen alpha-2(I) and collagen alpha-1(III), are reported below.

Figure 6.1: Alignment of the sequences of Collagen α-1(I). The aminoacidic sequences from *Sus scrofa, Ovis Aries (W5P481), Gallus gallus (P02457), Bos taurus (P2453)* and *Capra hircus*. The experimentally identified peptides are highlighted in light blue and aminoacids, which are different among species, are highlighted in red.

Protein sequences were aligned using Align tool available online at UniProt which uses Clustal-Omega program. * (asterisk) indicates positions which have a single, fully conserved residue; (colon) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix; . (period) indicates conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

```
COL1A1 Sus          MFSFVDLRLLLLLAATALLTHGQEEGQEEGQQGQEEDIPPVTCVQNGLRYHDRDVWKPVP
COL1A1 Ovis         MFSFVDLRLLLLLAATALLTHGQEEGQEEG---QEEDIPPVTCVQNGLRYHDRDVWKPVP
COL1A1 Gallus       MFSFVDSRLLLLIAATVLLTRGEG----------EEDIQTGSCVQDGLTYNDKDVWKPEP
COL1A1 Bos          MFSFVDLRLLLLLAATALLTHGQEEGQEEG---QEEDIPPVTCVQNGLRYHDRDVWKPVP
COL1A1 Capra        MFSFVDLRLLLLLAATALLTHGQEEGQEEG---QEEDIPPVTCVQNGLRYHDRDVWKPVP
                    ****** *****:***.***:*:           ****   :***:** *.*:***** *

COL1A1 Sus          CQICVCDNGNVLCDDVICDEIKNGPKGD--------------------------------
COL1A1 Ovis         CQICVCDNGNVLCDDVICDELKDCPNAKVPTPPR-PVSPHPPPPPPPPTTTKQRGGRGGP
COL1A1 Gallus       CQICVCDSGNILCDEVICEDTSDCPNAEIPFGECCPICPDVDASPVYPESAGVEGPKGD-
COL1A1 Bos          CQICVCDNGNVLCDDVICDELKDCPNAKVPTDECCPVCPEGQESPTDQETTGVEGPKGD-
COL1A1 Capra        CQICVCDNGNVLCDDVICDELKDCPNAKVPTDECCPVCPEGQESTTDQETTGVEGPKGD-
                    *******.**:***:***:: .: *:..


COL1A1 Sus          --------TGPRGPRGPSGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYD
COL1A1 Ovis         CPRPGAHRPSHSPPQGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYD
COL1A1 Gallus       --------TGPRGDRGLPGPPGRDGIPGQPGLPGPPG---PPGPPGLGGNFAPQMSYGYD
COL1A1 Bos          --------TGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYD
COL1A1 Capra        --------TGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYD
                             .    :*  ******************   **************:*****


COL1A1 Sus          EKSAGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKN
COL1A1 Ovis         EKSTGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKN
COL1A1 Gallus       EKSAGVAVPGPMGPAGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPAGPPGKN
COL1A1 Bos          EKSTGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKN
COL1A1 Capra        EKSTGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKN
                    ***:*::*******:************************************ ******
```


COL1A1 Sus GDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEP
COL1A1 Ovis GDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEP
COL1A1 Gallus GDDGEAGKPGRPGQRGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGQPGPAGPKGEP
COL1A1 Bos GDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEP
COL1A1 Capra GDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEP
GLPGER
COL1A1 Sus GSPGENGAPGQMGPRGLPGERGRPGPPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAV
COL1A1 Ovis GSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAV
COL1A1 Gallus GSPGENGAPGQMGPRGLPGERGRPGPSGPAGARGNDGAPGAAGPPGPTGPAGPPGFPGAA
COL1A1 Bos GSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAV
COL1A1 Capra GSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAV
GSEGPQGVR
COL1A1 Sus GAKGEAGPQGARGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGGKGANVSILKVSEL
COL1A1 Ovis GAKGEAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAP
COL1A1 Gallus GAKGETGPQGARGSEGPQGSRGEPGPPGPAGAAGPAGNPGADGQPGAKGATGAPGIAGAP
COL1A1 Bos GAKGEGGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAP
COL1A1 Capra GAKGEAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAP
COL1A1 Sus VNLVLPAAKETLAP--------------------------RESPVPLVFKAPLALL--E
COL1A1 Ovis ---GFPGARGPSGPQGPSGPPGPKGNSGEPGAPGSKGDTGAKGEPGPTGIQGPPGPAGEE
COL1A1 Gallus ---GFPGARGPSGPQGPSGAPGPKGNSGEPGAPGNKGDTGAKGEPGPAGVQGPPGPAGEE
COL1A1 Bos ---GFPGARGPSGPQGPSGPPGPKGNSGEPGAPGSKGDTGAKGEPGPTGIQGPPGPAGEE
COL1A1 Capra ---GFPGARGPSGPQGPSGPPGPKGNSGEPGAPGSKGDTGAKGEPGPTGIQGPPGPAGEE
GEPGPAGLPGPPGER
COL1A1 Sus KKESEKPEVN---L----------DLLACLDPLASVARLHVP--IRGL----------
COL1A1 Ovis GKRGARGEPGPAGLPGPPGERGGPGSRGFPGSDGVA---GPKGPAGERGAPGPAGPKGSP
COL1A1 Gallus GKRGARGEPGPAGLPGPAGERGAPGSRGFPGADGIA---GPKGPPGERGSPGAVGPKGSP
COL1A1 Bos GKRGARGEPGPAGLPGPPGERGGPGSRGFPGADGVA---GPKGPAGERGAPGPAGPKGSP
COL1A1 Capra GKRGARGEPGPAGLPGPPGERGGPGSRGFPGSDGVA---GPKGPAGERGAPGPAGPKGSP
COL1A1 Sus -------QTSLLTLSLTLAFPPSPMPP--KAPPPPRRKTMLPHAGLESLSVALTPPIRVP
COL1A1 Ovis GEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAG--------QDGRPGPPGPPGARG
COL1A1 Gallus GEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAG--------QDGRPGPAGPPGARG
COL1A1 Bos GEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAG--------QDGRPGPPGPPGARG
COL1A1 Capra GEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAG--------QDGRPGPPGPPGARG
COL1A1 Sus LVNVVLLALLVPKVLL-----------VKLVAPVKLVCLVPSSGAQGPPGPAGPAGERGE
COL1A1 Ovis QA--GVMGFPGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGE
COL1A1 Gallus QA--GVMGFPGPKGAAGEPGKPGERGAPGPPGAVGAAGKDGEAGAQGPPGPTGPAGERGE
COL1A1 Bos QA--GVMGFPGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGE
COL1A1 Capra QA--GVMGFPGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGE
GVQGPPGP
COL1A1 Sus QGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPGP
COL1A1 Ovis QGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPGP
COL1A1 Gallus QGPAGAPGFQGLPGPAGPPGEAGKPGEQGVPGNAGAPGPAGARGERGFPGERGVQGPPGP
COL1A1 Bos QGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPGP
COL1A1 Capra QGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPGP


AGPR
COL1A1 Sus AGPRGANARAPLAFRECLANE----------VQLVSQVLRVTEVL-----MANLVLKAM
COL1A1 Ovis AGPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGA
COL1A1 Gallus QGPRGANGAPGNDGAKGDAGAPGAPGNEGPPGLEGMPGERGAAGLPGAKGDRGDPGPKGA
COL1A1 Bos AGPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGA
COL1A1 Capra AGPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGA
COL1A1 Sus ------LVPPALLDPLAPLAPL----------EM
COL1A1 Ovis DGAPGKDGVRGLTGPIGPPGPAGAPGDKGETGPSGPAGPTGARGAPGDRGEPGPPGPAGF
COL1A1 Gallus DGAPGKDGLRGLTGPIGPPGPAGAPGDKGEAGPPGPAGPTGARGAPGDRGEPGPPGPAGF
COL1A1 Bos DGAPGKDGVRGLTGPIGPPGPAGAPGDKGEAGPSGPAGPTGARGAPGDRGEPGPPGPAGF
COL1A1 Capra DGAPGKDGVRGLTGPIGPPGPAGAPGDKGETGPSGPAGPTGARGAPGDRGEPGPPGPAGF
GSAGPPGA
COL1A1 Sus L----------DPLALLVLLAKK---
COL1A1 Ovis AGPPGADGQPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGAPGPKGARGSAGPPGA
COL1A1 Gallus AGPPGADGQPGAKGETGDAGAKGDAGPPGPAGPTGAPGPAGZVGAPGPKGARGSAGPPGA
COL1A1 Bos AGPPGADGQPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGAPGPKGARGSAGPPGA
COL1A1 Capra AGPPGADGQPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGKVGAPGPKGARGSAGPPGA
TGFPGAAGR
COL1A1 Sus ---------AAKVPVVRLAPLGVPVKPVPLAPLAPLVR
COL1A1 Ovis TGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGPRGETGPAGRAGEVGPPGPPGPAGE
COL1A1 Gallus TGFPGAAGRVGPPGPSGNIGLPGPPGPAGKZGSKGPRGETGPAGRPGEPGPAGPPGPPGE
COL1A1 Bos TGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGPRGETGPAGRPGEVGPPGPPGPAGE
COL1A1 Capra TGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGPRGETGPAGRPGEVGPPGPPGPAGE
GVVGLPGQR
COL1A1 Sus KD---------------PLVL---TDLLVLPVLLDLRVLLDSVVWSACPVNEEKEASLVF
COL1A1 Ovis KGAPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGER------GFPGLPGPSG-EPGKQG
COL1A1 Gallus KGSPGADGPIGAPGTPGPQGIAGQRGVVGLPGQRGER------GFPGLPGPSG-EPGKQG
COL1A1 Bos KGAPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGER------GFPGLPGPSG-EPGKQG
COL1A1 Capra KGAPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGER------GFPGLPGPSG-EPGKQG
COL1A1 Sus PAHLVTVVRAALLDPLVLLVLLVPP----------
COL1A1 Ovis PSGAS--GERGPPGPMGPPGLAGPPGESGREGAPGAEGSPGRDGAPGAKGDRGETGPAGP
COL1A1 Gallus PSGAS--GERGPPGPMGPPGLAGPPGEAGREGAPGAEGAPGRDGAAGPKGDRGETGPAGP
COL1A1 Bos PSGAS--GERGPPGPMGPPGLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGP
COL1A1 Capra PSGAS--GERGPPGPMGPPGLAGPPGESGREGAPGAEGSPGRDGAPGAKGDRGETGPAGP
COL1A1 Sus -------APLALLARAAIVSLCSLQGPAGPAGPVGPVGARGPAGPQGPRGDKGETGEQGD
COL1A1 Ovis PGAPGAPGAPGPVGPAGKSGDRGETGPAGPAGPIGPVGARGPAGPQGPRGDKGETGEQGD
COL1A1 Gallus PGAPGAPGAPGPVGPAGKNGDRGETGPAGPAGPPGPAGARGPAGPQGPRGDKGETGEQGD
COL1A1 Bos PGAPGAPGAPGPVGPAGKSGDRGETGPAGPAGPIGPVGARGPAGPQGPRGDKGETGEQGD
COL1A1 Capra PGAPGAPGAPGPVGPAGKSGDRGETGPAGPAGPIGPVGARGPAGPQGPRGDKGETGEQGD
COL1A1 Sus RGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNGLPGPIGP
COL1A1 Ovis RGIKGHRGFSGLQGPPGPPVSMLSPSPA-PSPASSLQGPPGSAGTPGKDGLNGLPGPIGP
COL1A1 Gallus RGMKGHRGFSGLQGPPGPPGAPGEQGPSGASGPAGPRGPPGSAGAAGKDGLNGLPGPIGP
COL1A1 Bos RGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPRGPPGSAGSPGKDGLNGLPGPIGP
COL1A1 Capra RGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPRGPPGSAGTPGKDGLNGLPGPIGP

```
COL1A1 Sus          PGPRGRTGDAGPVVRSHTFTRSACGPPALEAP----------------NPSNPPSPPSLY
COL1A1 Ovis         PGPRGRTGDAGPAVSPPTLSA--HGPPALKAP--SPRSRYDLSFLPQPPQEKAHDGGRYY
COL1A1 Gallus       PGPRGRTGEVGPVGP--------PGPPGPPGPPGPPSGGFDLSFLPQPPQEKAHDGGRYY
COL1A1 Bos          PGPRGRTGDAGPAGP--------PGPPGPPGPPGPPSGGYDLSFLPQPPQEKAHDGGRYY
COL1A1 Capra        PGPRGRTGDAGPAVSPPTLSA--HGPPGPSNPPSPPSXGYDLSFLPQPPQEKAHDGGRYY
                    ********:.**.           ***.   *                  .:  .    *

COL1A1 Sus          RADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHSDWKSGEYWI
COL1A1 Ovis         RADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHPDWKSGEYWI
COL1A1 Gallus       RADDANVMRDRDLEVDTTLKSLSQQIENIRSPEGTRKNPARTCRDLKMCHGDWKSGEYWI
COL1A1 Bos          RADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHSDWKSGEYWI
COL1A1 Capra        RADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTRRSLSSL---SAHGEYWI
                    *******:***************************** *.*.        *****

COL1A1 Sus          DPNQGCNLDAIKVFCNMETGETCVYPTQPSVPQKNWYISKNPKDKRHVWYGESMTDGFQV
COL1A1 Ovis         DPNQGCNLDAIKVFCNMETGETCVYPTQPSVPQKNWYISKNPKDKRHVWYGESMTGGFQV
COL1A1 Gallus       DPNQGCNLDAIKVYCNMETGETCVYPTQATIAQKNWYLSKNPKEKKHVWFGETMSDGFQF
COL1A1 Bos          DPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYISKNPKEKRHVWYGESMTGGFQF
COL1A1 Capra        DPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYISKNPKDKRHVWYGESMTGGFQF
                    *************:************** :: *****:*****:*:***:**:*: ***.

COL1A1 Sus          RELEPLFQDGLAQGSTGGLILGNPDLCYVLRLMSTEASQNITYHCKNSVAYMDQQTGNLK
COL1A1 Ovis         REGGQG-----------SDPADVAIQLTFLRLMSTEASQNITYHCKNSVAYMDQQTGNLK
COL1A1 Gallus       EYGGEG-----------SNPADVAIQLTFLRLMSTEATQNVTYHCKNSVAYMDHDTGNLK
COL1A1 Bos          EYGGQG-----------SDPADVAIQLTFLRLMSTEASQNITYHCKNSVAYMDQQTGNLK
COL1A1 Capra        EYGGQG-----------SDPADVAIQLTFLRLMSTEASQNITYHCKNSVAYMDQQTGNLK
                    .                .          .********:**:************::*****

COL1A1 Sus          KALLLQGSNEIEIRAEGNSRFTYSVIYDGCTVSPAGPPFIGLWLPRQAQCPPPAHSRWV-
COL1A1 Ovis         KALLLQGSNEIEIRAEGNSRFTYSVTYDGCTSHT------GAWGKTVIEYKTTKTSRLPI
COL1A1 Gallus       KALLLQGANEIEIRAEGNSRFTYGVTEDGCTSHT------GAWGKTVIEYKTTKTSRLPI
COL1A1 Bos          KALLLQGSNEIEIRAEGNSRFTYSVTYDGCTSHT------GAWGKTVIEYKTTKTSRLPI
COL1A1 Capra        KALLLQGSNEIEIRAEGNSRFTYSVTYDGCTSHT------GAWGKTVIEYKTTKTSRLPI
                    *******:***************.*  ****         * *     :      **

COL1A1 Sus          ---TPLS--------------HC--
COL1A1 Ovis         IDVAPLDVGAPDQEFGFDIGSVCFL
COL1A1 Gallus       IDLAPMDVGAPDQEFGIDIGPVCFL
COL1A1 Bos          IDVAPLDVGAPDQEFGFDVGPACFL
COL1A1 Capra        IDVAPLDVGAPDQEFGFDIGSVCFL
                       :*:.              *
```

The multiple alignment of collagen alpha-1(I) hasn't allowed to distinguish among species because we couldn't identify any proteotypic peptides.

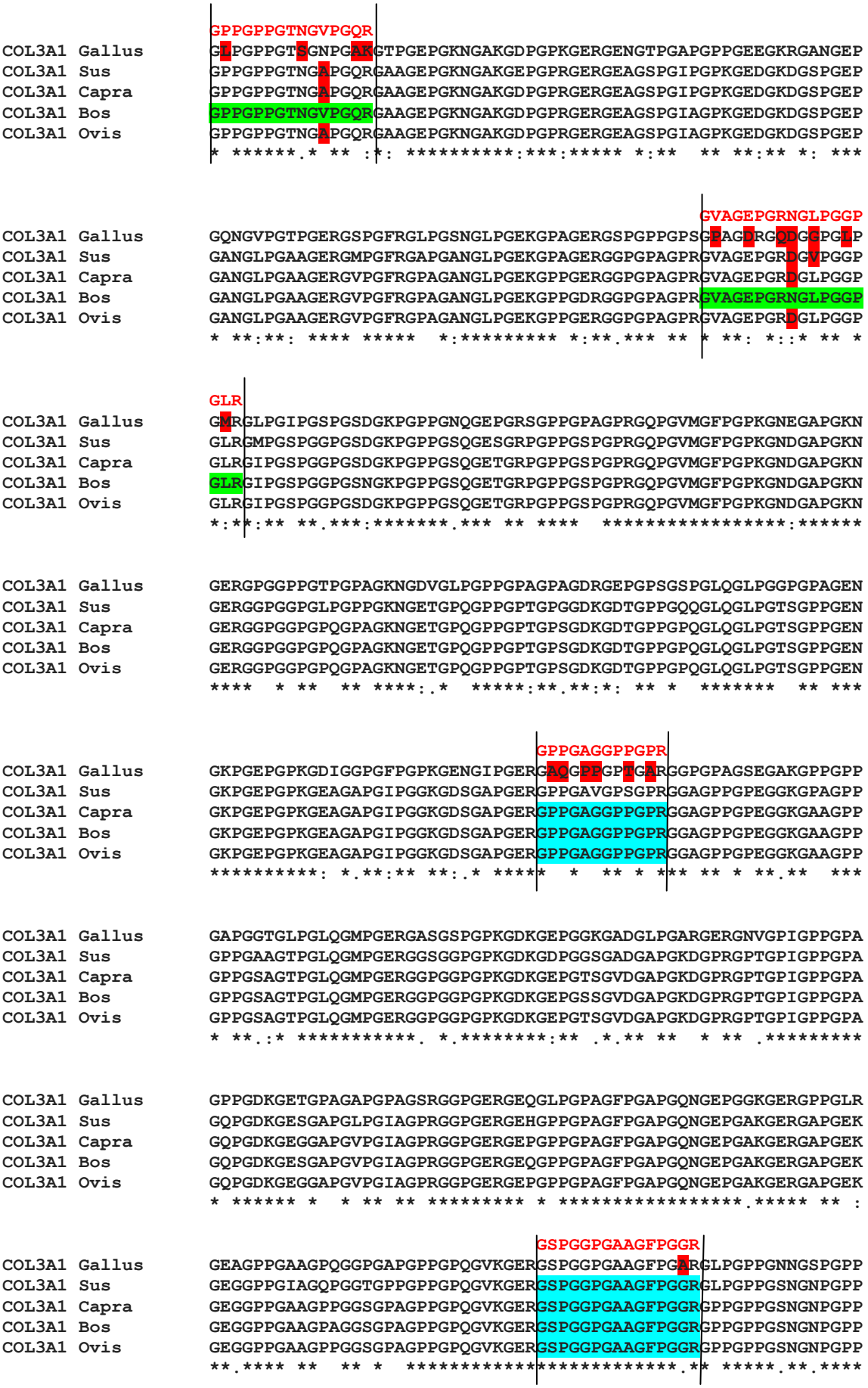
However, by the analysis of data we can exclude *Sus scrofa* and *Gallus gallus* as species used for the manufacturing of the drape. These sequences show low homology with the other species analysed.

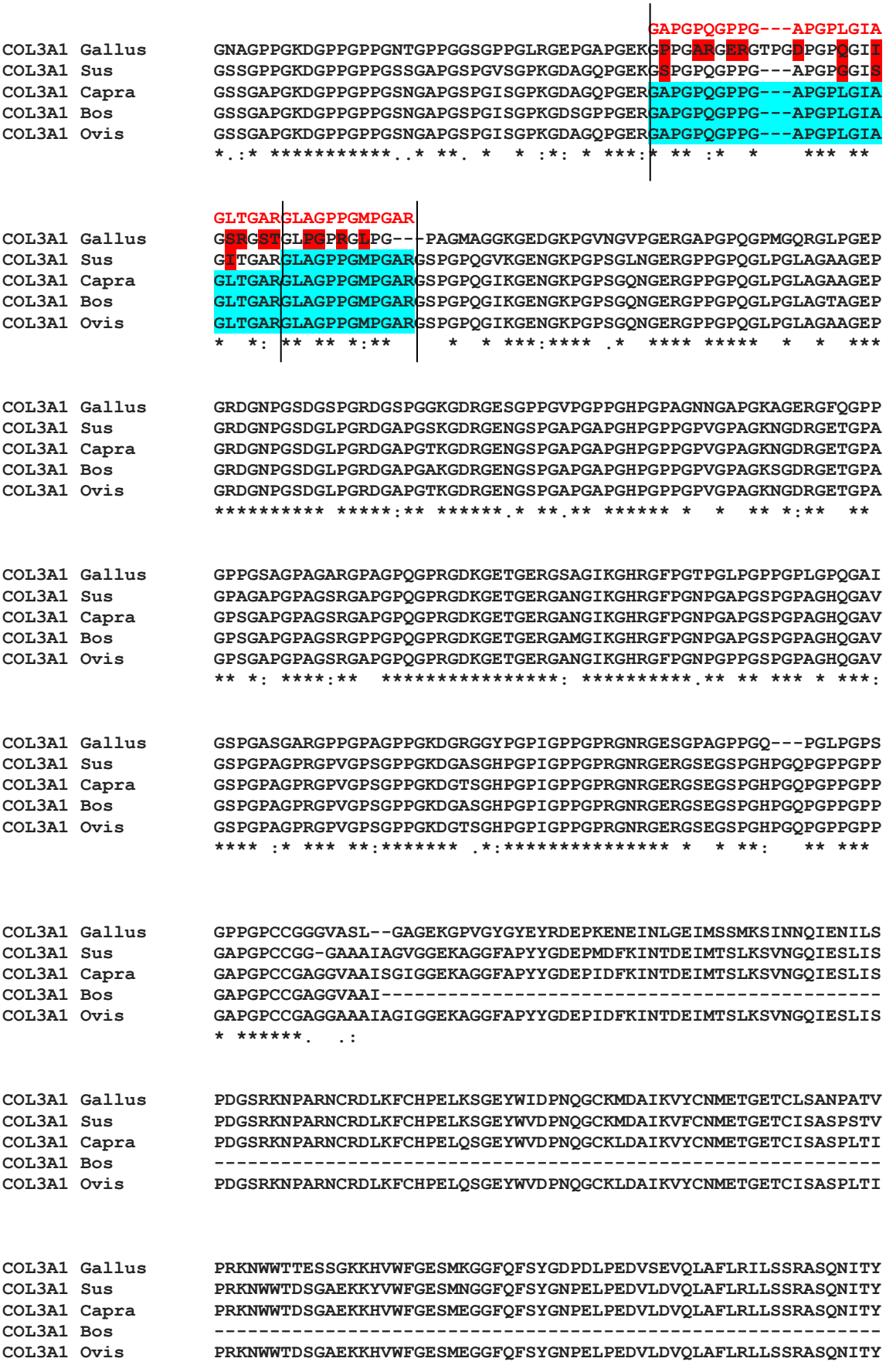
Sus scrofa, Bos taurus, Ovis aries and *Capra hircus* belong to the same order (*Artiodactyla*) phylogenetically, but *Sus scrofa* belongs to another sub-order (*Suiformes*) compared to the other species (*Ruminantia)*. *Gallus gallus* instead belongs to a different class (*Aves*).

Figure 6.2: Alignment of the sequences of Collagen α-1(III). The aminoacidic sequences from *Sus scrofa* (*F1RYI8)*, *Ovis Aries (W5Q4S0)*, *Gallus gallus (CO3A1)*, *Bos taurus (CO3A1)* and *Capra hircus* (*CO3A1)*. The experimentally identified peptides are highlighted in light blue and aminoacids, which are different among species, are highlighted in red.

Protein sequences were aligned using Align tool available online at UniProt which uses Clustal-Omega program. * (asterisk) indicates positions which have a single, fully conserved residue; (colon) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix; . (period) indicates conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

```
COL3A1 Gallus      MMSFVQKVSLFILAVFQPSVILAQQD--ALGGCTHLGQEYADRDVWKPEPCQICVCDSGS
COL3A1 Sus         MTSFVQKGTWLLFALLHPTVILAQQQEAIEGGCSHLGQSYADRDVWKPEPCQICVCDSGS
COL3A1 Capra       MMSFVQKGTWLLFALLHPTVILAQ-QEAVDGGCSHLGQAYADRDVWKPEPCQICVCDSGS
COL3A1 Bos         ------------------------------------------------------------
COL3A1 Ovis        MMSFVQKGTWLLFALLHPTVILAQ-QEAVDGGCSHLGQAYADRDVWKPEPCQICVCDSGS


COL3A1 Gallus      VLCDDIICDDQELDCPNPEIPFGECCPVCPQTTPQPTKLPYT---QGPKGDPGSPGSPGR
COL3A1 Sus         VLCDDIICDDQELDCPNPEIPFGECCAVCPQPPTAPTRPPNGHGPQGPKGDPGPPGIPGR
COL3A1 Capra       VLCDDIICDDQELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGR
COL3A1 Bos         ------------------------------------------------------------
COL3A1 Ovis        VLCDDIICDDQELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGR


COL3A1 Gallus      TGAPGPPGQPGSPGAPGPPGICQSCPSISGGSFSPQYDSYDVKAGSVGMGYPPQPISGFP
COL3A1 Sus         NGDPGLPGQPGSPGSPGPPGICESCPTG-GQNYSPQYESYDVKAGVAGG-----GIGGYP
COL3A1 Capra       NGDPGPPGSPGSPGSPGPPGICESCPTG-SQNYSPQYEAYDVKSGVAGG----VGLAGYP
COL3A1 Bos         ----------------------------------EYEAYDVKSGVAGG-----GIAGYP
COL3A1 Ovis        NGDPGPPGSPGSPGSPGPPGICESCPTG-GQNYSPQYEAYDVKSGVAGG----VGLAGYP
                                                    :*::****:* .*       :.*:*


COL3A1 Gallus      GPP---GPSGPPGPPGHAGPPGSNGYQGPPGEPGQPGPSGPPGPAGMIGPAGPPGKDGEP
COL3A1 Sus         GPAGPPGPPGPPGVSGHPGAPGSPGYQGPPGEPGQAGPAGPPGPPGAIGPSGPAGKDGES
COL3A1 Capra       GPAGPPGPPGPPGTSGHPGAPGAPGYQGPPGEPGQAGPAGPPGPPGAIGPSGPAGKDGES
COL3A1 Bos         GPAGPPGPPGPPGTSGHPGAPGAPGYQGPPGEPGQAGPAGPPGPPGAIGPS---GKDGES
COL3A1 Ovis        GPAGPPGPPGPPGTSGHPGAPGAPGYQGPPGEPGQAGPAGPPGPPGAIGPSGPAGKDGES
                   **    ** ****  ** * **: ********** **:***** * ***:   *****


COL3A1 Gallus      GRPGRNGDRGIPGLPGHKGHPGMPGMPGMKGARGFDGKDGAKGDSGAPGPKGEAGQPGAN
COL3A1 Sus         GRPGRPGERGLPGPPGLKGPAGMPGFPGMKGHRGFDGRNGEKGDTGAPGLKGENGLPGEN
COL3A1 Capra       GRPGRPGERGLPGPPGMKGPAGMPGFPGMKGHRGFDGRNGEKGETGAPGLKGENGIPGEN
COL3A1 Bos         GRPGRPGPRGFPGPPGMKGPAGMPGFPGMKGHRGFDGRNGEKGEPGAPGLKGENGVPGED
COL3A1 Ovis        GRPGRPGERGLPGPPGMKGPAGMPGFPGMKGHRGFDGRNGEKGETGAPGLKGENGIPGEN
                   ***** * **:** ** **  ****:***** *****::* **: **** *** * ** :


COL3A1 Gallus      GSPGQPGPRGPTGERGRPGNPGGPGAHGKDGAPGAAGPPGPPGPPGTAGFPGSPGFKGEA
COL3A1 Sus         GAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEV
COL3A1 Capra       GAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEV
COL3A1 Bos         GAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEV
COL3A1 Ovis        GAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEV
                   *:**  ****  ******* **. **:*:*** *: * ***************** ***.


COL3A1 Gallus      GPPGPAGASGSPGERGEPGPQGQAGPPGPQGPPGRAGSPGNKGEMGPSGIPGAPGLPGGR
COL3A1 Sus         GPAGSPGPSGSPGQRGEPGPQGHAGAAGPPGPPGSNGSPGGKGEMGPAGIPGAPGLMGAR
COL3A1 Capra       GPAGSPGSSGAPGQRGEPGPQGHAGAPGPPGPPGSNGSPGGKGEMGPAGIPGAPGLLGAR
COL3A1 Bos         GPAGSPGSSGAPGQRGEPGPQGHAGAPGPPGPPGSDGSPGGKGEMGPAGIPGAPGLIGAR
COL3A1 Ovis        GPAGSPGSSGAPGQRGEPGPQGHAGAPGPPGPPGSNGSPGGKGEMGPAGIPGAPGLLGAR
                   ** *  * **:**:********:**  ** ****  **** ******:******** *.*
```


GPPGPPGTNGVPGQR
COL3A1 Gallus GLPGPPGTSGNPGAKGTPGEPGKNGAKGDPGPKGERGENGTPGAPGPPGEEGKRGANGEP
COL3A1 Sus GPPGPPGTNGAPGQRGAAGEPGKNGAKGEPGPRGERGEAGSPGIPGPKGEDGKDGSPGEP
COL3A1 Capra GPPGPPGTNGAPGQRGAAGEPGKNGAKGDPGPRGERGEAGSPGIPGPKGEDGKDGSPGEP
COL3A1 Bos GPPGPPGTNGVPGQRGAAGEPGKNGAKGDPGPRGERGEAGSPGIAGPKGEDGKDGSPGEP
COL3A1 Ovis GPPGPPGTNGAPGQRGAAGEPGKNGAKGDPGPRGERGEAGSPGIAGPKGEDGKDGSPGEP
GVAGEPGRNGLPGGP
COL3A1 Gallus GQNGVPGTPGERGSPGFRGLPGSNGLPGEKGPAGERGSPGPPGPSGPAGDRGQDGGPGLP
COL3A1 Sus GANGLPGAAGERGMPGFRGAPGANGLPGEKGPAGERGGPGPAGPRGVAGEPGRDGVPGGP
COL3A1 Capra GANGLPGAAGERGVPGFRGPAGANGLPGEKGPPGERGGPGPAGPRGVAGEPGRDGLPGGP
COL3A1 Bos GANGLPGAAGERGVPGFRGPAGANGLPGEKGPPGDRGGPGPAGPRGVAGEPGRNGLPGGP
COL3A1 Ovis GANGLPGAAGERGVPGFRGPAGANGLPGEKGPPGERGGPGPAGPRGVAGEPGRDGLPGGP
GLR
COL3A1 Gallus GMRGLPGIPGSPGSDGKPGPPGNQGEPGRSGPPGPAGPRGQPGVMGFPGPKGNEGAPGKN
COL3A1 Sus GLRGMPGSPGGPGSDGKPGPPGSQGESGRPGPPGSPGPRGQPGVMGFPGPKGNDGAPGKN
COL3A1 Capra GLRGIPGSPGGPGSDGKPGPPGSQGETGRPGPPGSPGPRGQPGVMGFPGPKGNDGAPGKN
COL3A1 Bos GLRGIPGSPGGPGSNGKPGPPGSQGETGRPGPPGSPGPRGQPGVMGFPGPKGNDGAPGKN
COL3A1 Ovis GLRGIPGSPGGPGSDGKPGPPGSQGETGRPGPPGSPGPRGQPGVMGFPGPKGNDGAPGKN
COL3A1 Gallus GERGPGGPPGTPGPAGKNGDVGLPGPPGPAGPAGDRGEPGPSGSPGLQGLPGGPGPAGEN
COL3A1 Sus GERGGPGGPGLPGPPGKNGETGPQGPPGPTGPGGDKGDTGPPGQQGLQGLPGTSGPPGEN
COL3A1 Capra GERGGPGGPGPQGPAGKNGETGPQGPPGPTGPSGDKGDTGPPGPQGLQGLPGTSGPPGEN
COL3A1 Bos GERGGPGGPGPQGPAGKNGETGPQGPPGPTGPSGDKGDTGPPGPQGLQGLPGTSGPPGEN
COL3A1 Ovis GERGGPGGPGPQGPAGKNGETGPQGPPGPTGPSGDKGDTGPPGPQGLQGLPGTSGPPGEN
GPPGAGGPPGPR
COL3A1 Gallus GKPGEPGPKGDIGGPGFPGPKGENGIPGERGACGPPGPTGARGGPGPAGSEGAKGPPGPP
COL3A1 Sus GKPGEPGPKGEAGAPGIPGGKGDSGAPGERGPPGAVGPSGPRGGAGPPGPEGGKGPAGPP
COL3A1 Capra GKPGEPGPKGEAGAPGIPGGKGDSGAPGERGPPGAGGPPGPRGGAGPPGPEGGKGAAGPP
COL3A1 Bos GKPGEPGPKGEAGAPGIPGGKGDSGAPGERGPPGAGGPPGPRGGAGPPGPEGGKGAAGPP
COL3A1 Ovis GKPGEPGPKGEAGAPGIPGGKGDSGAPGERGPPGAGGPPGPRGGAGPPGPEGGKGAAGPP
COL3A1 Gallus GAPGGTGLPGLQGMPGERGASGSPGPKGDKGEPGGKGADGLPGARGERGNVGPIGPPGPA
COL3A1 Sus GPPGAAGTPGLQGMPGERGGSGGPGPKGDKGDPGGSGADGAPGKDGPRGPTGPIGPPGPA
COL3A1 Capra GPPGSAGTPGLQGMPGERGGPGGPGPKGDKGEPGTSGVDGAPGKDGPRGPTGPIGPPGPA
COL3A1 Bos GPPGSAGTPGLQGMPGERGGPGGPGPKGDKGEPGSSGVDGAPGKDGPRGPTGPIGPPGPA
COL3A1 Ovis GPPGSAGTPGLQGMPGERGGPGGPGPKGDKGEPGTSGVDGAPGKDGPRGPTGPIGPPGPA
COL3A1 Gallus GPPGDKGETGPAGAPGPAGSRGGPGERGEQGLPGPAGFPGAPGQNGEPGGKGERGPPGLR
COL3A1 Sus GQPGDKGESGAPGLPGIAGPRGGPGERGEHGPPGPAGFPGAPGQNGEPGAKGERGAPGEK
COL3A1 Capra GQPGDKGEGGAPGVPGIAGPRGGPGERGEPGPPGPAGFPGAPGQNGEPGAKGERGAPGEK
COL3A1 Bos GQPGDKGESGAPGVPGIAGPRGGPGERGEQGPPGPAGFPGAPGQNGEPGAKGERGAPGEK
COL3A1 Ovis GQPGDKGEGGAPGVPGIAGPRGGPGERGEPGPPGPAGFPGAPGQNGEPGAKGERGAPGEK
GSPGGPGAAGFPGGR
COL3A1 Gallus GEAGPPGAAGPQGGPGAPGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGNNGSPGPP
COL3A1 Sus GEGGPPGIAGQPGGTGPPGPPGPQGVKGERGSPGGPGAAGFPGGRGLPGPPGSNGNPGPP
COL3A1 Capra GEGGPPGAAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGGRGPPGPPGSNGNPGPP
COL3A1 Bos GEGGPPGAAGPAGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGGRGPPGPPGSNGNPGPP
COL3A1 Ovis GEGGPPGAAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGGRGPPGPPGSNGNPGPP

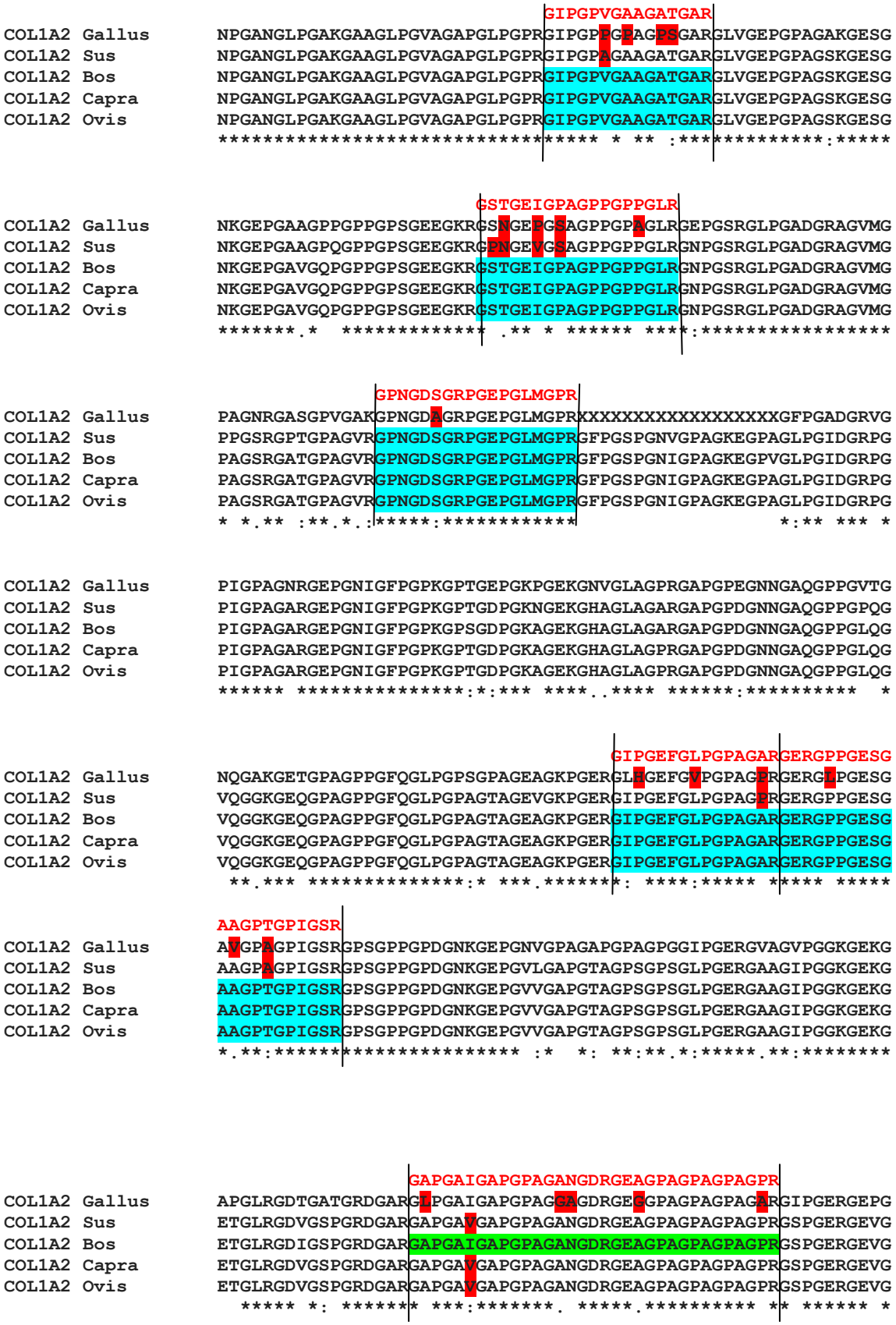

GAPGPQGPPG---APGPLGIA
COL3A1 Gallus GNAGPPGKDGPPGPPGNTGPPGGSGPPGLRGEPGAPGEKGPPGARGERGTPGDPGPQGII
COL3A1 Sus GSSGPPGKDGPPGPPGSSGAPGSPGVSGPKGDAGQPGEKGSPGPQGPPG---APGPGGIS
COL3A1 Capra GSSGAPGKDGPPGPPGSNGAPGSPGISGPKGDAGQPGERGAPGPQGPPG---APGPLGIA
COL3A1 Bos GSSGAPGKDGPPGPPGSNGAPGSPGISGPKGDSGPPGERGAPGPQGPPG---APGPLGIA
COL3A1 Ovis GSSGAPGKDGPPGPPGSNGAPGSPGISGPKGDAGQPGERGAPGPQGPPG---APGPLGIA
GLTGARGLAGPPGMPGAR
COL3A1 Gallus GSRGSTGLPGPRGLPG---PAGMAGGKGEDGKPGVNGVPGERGAPGPQGPMGQRGLPGEP
COL3A1 Sus GITGARGLAGPPGMPGARGSPGPQGVKGENGKPGPSGLNGERGPPGPQGLPGLAGAAGEP
COL3A1 Capra GLTGARGLAGPPGMPGARGSPGPQGIKGENGKPGPSGQNGERGPPGPQGLPGLAGAAGEP
COL3A1 Bos GLTGARGLAGPPGMPGARGSPGPQGIKGENGKPGPSGQNGERGPPGPQGLPGLAGTAGEP
COL3A1 Ovis GLTGARGLAGPPGMPGARGSPGPQGIKGENGKPGPSGQNGERGPPGPQGLPGLAGAAGEP
COL3A1 Gallus GRDGNPGSDGSPGRDGSPGGKGDRGESGPPGVPGPPGHPGPAGNNGAPGKAGERGFQGPP
COL3A1 Sus GRDGNPGSDGLPGRDGAPGSKGDRGENGSPGAPGAPGHPGPPGPVGPAGKNGDRGETGPA
COL3A1 Capra GRDGNPGSDGLPGRDGAPGTKGDRGENGSPGAPGAPGHPGPPGPVGPAGKNGDRGETGPA
COL3A1 Bos GRDGNPGSDGLPGRDGAPGAKGDRGENGSPGAPGAPGHPGPPGPVGPAGKSGDRGETGPA
COL3A1 Ovis GRDGNPGSDGLPGRDGAPGTKGDRGENGSPGAPGAPGHPGPPGPVGPAGKNGDRGETGPA
COL3A1 Gallus GPPGSAGPAGARGPAGPQGPRGDKGETGERGSAGIKGHRGFPGTPGLPGPPGPLGPQGAI
COL3A1 Sus GPAGAPGPAGSRGAPGPQGPRGDKGETGERGANGIKGHRGFPGNPGAPGSPGPAGHQGAV
COL3A1 Capra GPSGAPGPAGSRGAPGPQGPRGDKGETGERGANGIKGHRGFPGNPGAPGSPGPAGHQGAV
COL3A1 Bos GPSGAPGPAGSRGPPGPQGPRGDKGETGERGAMGIKGHRGFPGNPGAPGSPGPAGHQGAV
COL3A1 Ovis GPSGAPGPAGSRGAPGPQGPRGDKGETGERGANGIKGHRGFPGNPGPPGSPGPAGHQGAV
COL3A1 Gallus GSPGASGARGPPGPAGPPGKDGRGGYPGPIGPPGPRGNRGESGPAGPPGQ---PGLPGPS
COL3A1 Sus GSPGPAGPRGPVGPSGPPGKDGASGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
COL3A1 Capra GSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
COL3A1 Bos GSPGPAGPRGPVGPSGPPGKDGASGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
COL3A1 Ovis GSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
COL3A1 Gallus GPPGPCCGGGVASL--GAGEKGPVGYGYEYRDEPKENEINLGEIMSSMKSINNQIENILS
COL3A1 Sus GAPGPCCGG-GAAAIAGVGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLIS
COL3A1 Capra GAPGPCCGAGGVAAISGIGGEKAGGFAPYYGDEPIDFKINTDEIMTSLKSVNGQIESLIS
COL3A1 Bos GAPGPCCGAGGVAAI----------
COL3A1 Ovis GAPGPCCGAGGAAAIAGIGGEKAGGFAPYYGDEPIDFKINTDEIMTSLKSVNGQIESLIS
COL3A1 Gallus PDGSRKNPARNCRDLKFCHPELKSGEYWIDPNQGCKMDAIKVYCNMETGETCLSANPATV
COL3A1 Sus PDGSRKNPARNCRDLKFCHPELKSGEYWVDPNQGCKMDAIKVFCNMETGETCISASPSTV
COL3A1 Capra PDGSRKNPARNCRDLKFCHPELQSGEYWVDPNQGCKLDAIKVYCNMETGETCISASPLTI
COL3A1 Bos ----------
COL3A1 Ovis PDGSRKNPARNCRDLKFCHPELQSGEYWVDPNQGCKLDAIKVYCNMETGETCISASPLTI
COL3A1 Gallus PRKNWWTTESSGKKHVWFGESMKGGFQFSYGDPDLPEDVSEVQLAFLRILSSRASQNITY
COL3A1 Sus PRKNWWTDSGAEKKYVWFGESMNGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITY
COL3A1 Capra PRKNWWTDSGAEKKHVWFGESMEGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITY
COL3A1 Bos ----------
COL3A1 Ovis PRKNWWTDSGAEKKHVWFGESMEGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITY

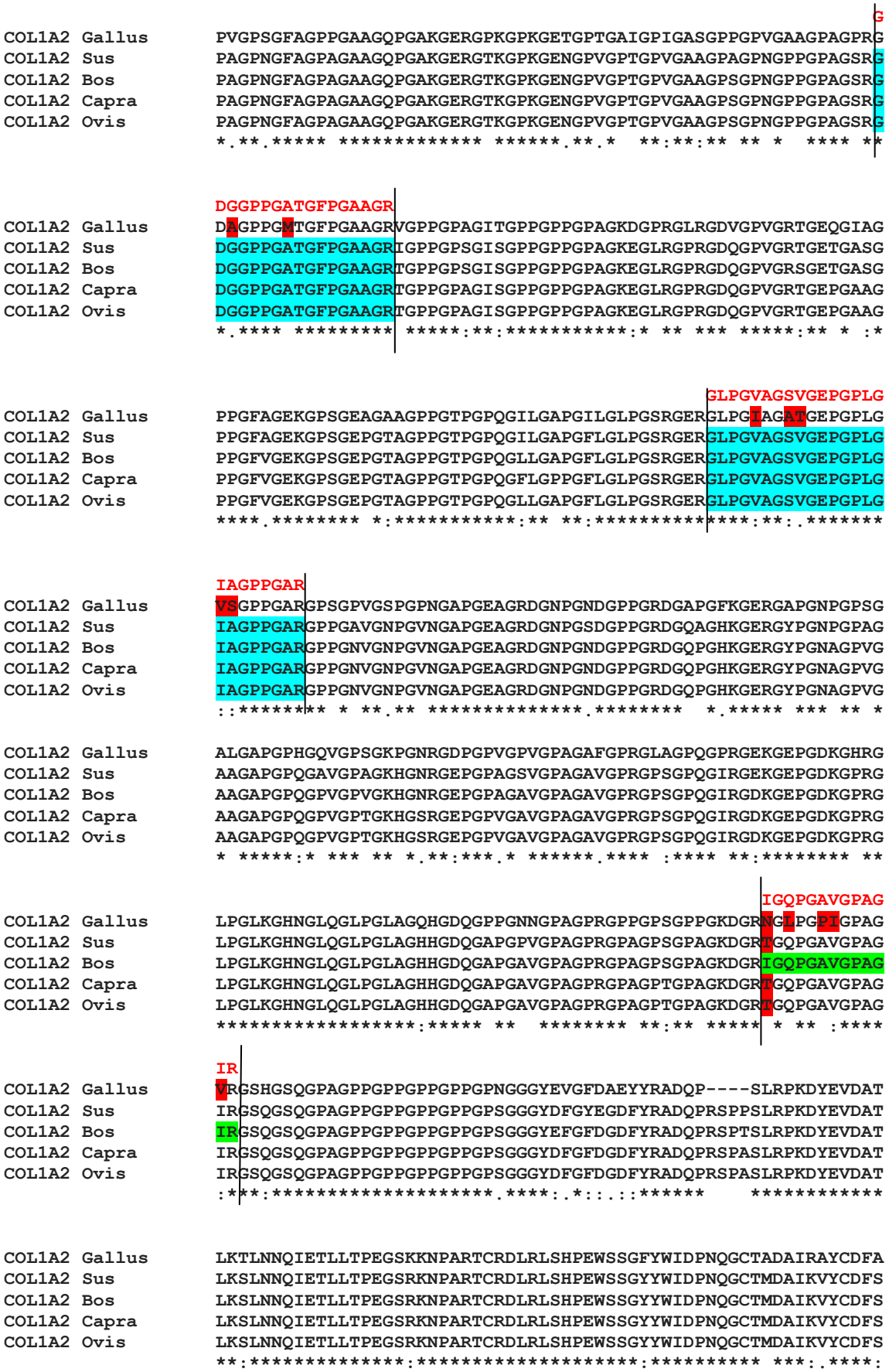
```
COL3A1 Gallus      HCKNSIAYMDQASGNVKKALKLMSSVETEIKAEGNSKYMYAVLEDGCTKHTGEWGKTVFE
COL3A1 Sus         HCKNSIAYMEHASGNVKKALRLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWGKTVFE
COL3A1 Capra       HCKNSIAYMDHASGNVKKALRLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWGKTVFQ
COL3A1 Bos         ------------------------------------------------------------
COL3A1 Ovis        HCKNSIAYMDHASGNVKKALRLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWGKTVFQ


COL3A1 Gallus      YRTRKTMRLPVVDIAPIDIGGPDQEFGVDVGPVCFL
COL3A1 Sus         YRTRKAVRLPIVDIAPYDIGGPDQEFGADIGPVCFL
COL3A1 Capra       YQTRKAVRLPIVDIAPYDIGGPDQEFGADIGPVCFL
COL3A1 Bos         -----------------------------------
COL3A1 Ovis        YQTRKAVRLPIVDIAPYDIGGPDQEFGADIGPVCFL
```

Figure 6.3: Alignment of the sequences of Collagen α-2(I). The aminoacidic sequences from *Sus scrofa* (*343887367*), *Ovis Aries (W5NTT7)*, *Gallus gallus (P02467)*, *Bos taurus (P02465)* and *Capra hircus* (*548462053)*. The experimentally identified peptides are highlighted in light blue and aminoacids, which are different among species, are highlighted in red.

Protein sequences were aligned using Align tool available online at UniProt which uses Clustal-Omega program. * (asterisk) indicates positions which have a single, fully conserved residue; (colon) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix; . (period) indicates conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

```
COL1A2 Gallus      MLSFVDTRILLLLAVTSYLATSQHVSEASAGRKGPRGDKGPQGERGPPGPPGRDGEDGPP
COL1A2 Sus         MLSFVDTRTLLLLAVTSCLATCQSLQEAT-ARKGPTGDRGPRGERGPPGPPGRDGDDGIP
COL1A2 Bos         MLSFVDTRTLLLLAVTSCLATCQSLQEAT-ARKGPSGDRGPRGERGPPGPPGRDGDDGIP
COL1A2 Capra       MLSFVDTRTLLLLAVTSCLATCQSLQEAT-ARKGPSGDRGPRGERGPPGPPGRDGDDGIP
COL1A2 Ovis        MLSFVDTRTLLLLAVTSCLATCQSLQEAT-ARKGPSGDRGPRGERGPPGPPGRDGDDGIP
                   ******** ******** ***.* :.**: .**** **:**:*************:** *

COL1A2 Gallus      ---GPPGPPGPPGLGGNFAAQYDPSKAADFGPGPMGLMGPRGPPGASGPPGPPGFQGVPG
COL1A2 Sus         GPPGPPGPPGPPGLGGNFAAQYDGK-GVGAGPGPMGLMGPRGPPGAVGAPGPQGFQGPAG
COL1A2 Bos         GPPGPPGPPGPPGLGGNFAAQFDAK-G--GGPGPMGLMGPRGPPGASGAPGPQGFQGPPG
COL1A2 Capra       GPPGPPGPPGPPGLGGNFAAQFDGK-G--GGPGPMGLMGPRGPPGASGAPGPQGFQGPPG
COL1A2 Ovis        GPPGPPGPPGPPGLGGNFAAQFDGK-G--GGPGPMGLMGPRGPPGASGAPGPQGFQGPPG
                      ******************:* . .   **************** * *** ****  *

                                                          GVVGPQGAR
COL1A2 Gallus      EPGEPGQTGPQGPRGPPGPPGKAGEDGHPGKPGRPGERGVAGPQGARGFPGTPGPPGFKG
COL1A2 Sus         EPGEPGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKG
COL1A2 Bos         EPGEPGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKG
COL1A2 Capra       EPGEPGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKG
COL1A2 Ovis        EPGEPGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKG
                   ********** * **************************.*******:******* *****

                                                          GLPGER
COL1A2 Gallus      IRGHNGLDGLTGQPGAPGTKGEPGAPGENGTPGQPGARGLPGERGRIGAPGPAGARGSDG
COL1A2 Sus         IRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGNDG
COL1A2 Bos         IRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDG
COL1A2 Capra       IRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDG
COL1A2 Ovis        IRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDG
                   **********.*******.*************** **********:**********.**

COL1A2 Gallus      SAGPTGPAXXXXXXXX-XXXXXXXXXGEIGPAGNEGPTGPAGPRGEIGLPGSSGPVGPPG
COL1A2 Sus         SVGPVDPAGPIGSAGPPGFPGAPGPKGELGPVGNPGPAGPAGPRGEVGLPGVSGPVGPPG
COL1A2 Bos         SVGPVGPAGPIGSAGPPGFPGAPGPKGELGPVGNPGPAGPAGPRGEVGLPGLSGPVGPPG
COL1A2 Capra       SVGPVGPAGPIGSAGPPGFPGAPGPKGELGPVGNPGPAGPAGPRGEVGLPGLSGPVGPPG
COL1A2 Ovis        SVGPVGPAGPIGSAGPPGFPGAPGPKGELGPVGNPGPAGPAGPRGEVGLPGLSGPVGPPG
                   *.**. **                  **:**.** **:********:**** ********
```


GIPGPVGAAGATGAR
COL1A2 Gallus NPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPPGPAGPSGARGLVGEPGPAGAKGESG
COL1A2 Sus NPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPAGAAGATGARGLVGEPGPAGSKGESG
COL1A2 Bos NPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGESG
COL1A2 Capra NPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGESG
COL1A2 Ovis NPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGESG
GSTGEIGPAGPPGPPGLR
COL1A2 Gallus NKGEPGAAGPPGPPGPSGEEGKRGSNGEPGSAGPPGPAGLRGEPGSRGLPGADGRAGVMG
COL1A2 Sus NKGEPGAAGPQGPPGPSGEEGKRGPNGEVGSAGPPGPPGLRGNPGSRGLPGADGRAGVMG
COL1A2 Bos NKGEPGAVGQPGPPGPSGEEGKRGSTGEIGPAGPPGPPGLRGNPGSRGLPGADGRAGVMG
COL1A2 Capra NKGEPGAVGQPGPPGPSGEEGKRGSTGEIGPAGPPGPPGLRGNPGSRGLPGADGRAGVMG
COL1A2 Ovis NKGEPGAVGQPGPPGPSGEEGKRGSTGEIGPAGPPGPPGLRGNPGSRGLPGADGRAGVMG
GPNGDSGRPGEPGLMGPR
COL1A2 Gallus PAGNRGASGPVGAKGPNGDAGRPGEPGLMGPRXXXXXXXXXXXXXXXXXGFPGADGRVG
COL1A2 Sus PPGSRGPTGPAGVRGPNGDSGRPGEPGLMGPRGFPGSPGNVGPAGKEGPAGLPGIDGRPG
COL1A2 Bos PAGSRGATGPAGVRGPNGDSGRPGEPGLMGPRGFPGSPGNIGPAGKEGPVGLPGIDGRPG
COL1A2 Capra PAGSRGATGPAGVRGPNGDSGRPGEPGLMGPRGFPGSPGNIGPAGKEGPAGLPGIDGRPG
COL1A2 Ovis PAGSRGATGPAGVRGPNGDSGRPGEPGLMGPRGFPGSPGNIGPAGKEGPAGLPGIDGRPG
COL1A2 Gallus PIGPAGNRGEPGNIGFPGPKGPTGEPGKPGEKGNVGLAGPRGAPGPEGNNGAQGPPGVTG
COL1A2 Sus PIGPAGARGEPGNIGFPGPKGPTGDPGKNGEKGHAGLAGARGAPGPDGNNGAQGPPGPQG
COL1A2 Bos PIGPAGARGEPGNIGFPGPKGPSGDPGKAGEKGHAGLAGARGAPGPDGNNGAQGPPGLQG
COL1A2 Capra PIGPAGARGEPGNIGFPGPKGPTGDPGKAGEKGHAGLAGPRGAPGPDGNNGAQGPPGLQG
COL1A2 Ovis PIGPAGARGEPGNIGFPGPKGPTGDPGKAGEKGHAGLAGPRGAPGPDGNNGAQGPPGLQG
GIPGEFGLPGPAGARGERGPPGESG
COL1A2 Gallus NQGAKGETGPAGPPGFQGLPGPSGPAGEAGKPGERGLHGEFGVPGPAGPRGERGLPGESG
COL1A2 Sus VQGGKGEQGPAGPPGFQGLPGPAGTAGEVGKPGERGIPGEFGLPGPAGPRGERGPPGESG
COL1A2 Bos VQGGKGEQGPAGPPGFQGLPGPAGTAGEAGKPGERGIPGEFGLPGPAGARGERGPPGESG
COL1A2 Capra VQGGKGEQGPAGPPGFQGLPGPAGTAGEAGKPGERGIPGEFGLPGPAGARGERGPPGESG
COL1A2 Ovis VQGGKGEQGPAGPPGFQGLPGPAGTAGEAGKPGERGIPGEFGLPGPAGARGERGPPGESG
AAGPTGPIGSR
COL1A2 Gallus AVGPAGPIGSRGPSGPPGPDGNKGEPGNVGPAGAPGPAGPGGIPGERGVAGVPGGKGEKG
COL1A2 Sus AAGPAGPIGSRGPSGPPGPDGNKGEPGVLGAPGTAGPSGPSGLPGERGAAGIPGGKGEKG
COL1A2 Bos AAGPTGPIGSRGPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERGAAGIPGGKGEKG
COL1A2 Capra AAGPTGPIGSRGPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERGAAGIPGGKGEKG
COL1A2 Ovis AAGPTGPIGSRGPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERGAAGIPGGKGEKG
GAPGAIGAPGPAGANGDRGEAGPAGPAGPAGPR
COL1A2 Gallus APGLRGDTGATGRDGARGLPGAIGAPGPAGGAGDRGEGGPAGPAGPAGARGIPGERGEPG
COL1A2 Sus ETGLRGDVGSPGRDGARGAPGAVGAPGPAGANGDRGEAGPAGPAGPAGPRGSPGERGEVG
COL1A2 Bos ETGLRGDIGSPGRDGARGAPGAIGAPGPAGANGDRGEAGPAGPAGPAGPRGSPGERGEVG
COL1A2 Capra ETGLRGDVGSPGRDGARGAPGAVGAPGPAGANGDRGEAGPAGPAGPAGPRGSPGERGEVG
COL1A2 Ovis ETGLRGDVGSPGRDGARGAPGAVGAPGPAGANGDRGEAGPAGPAGPAGPRGSPGERGEVG


G
COL1A2 Gallus PVGPSGFAGPPGAAGQPGAKGERGPKGPKGETGPTGAIGPIGASGPPGPVGAAGPAGPRG
COL1A2 Sus PAGPNGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPAGPNGPPGPAGSRG
COL1A2 Bos PAGPNGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG
COL1A2 Capra PAGPNGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG
COL1A2 Ovis PAGPNGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG
DGGPPGATGFPGAAGR
COL1A2 Gallus DAGPPGMTGFPGAAGRVGPPGPAGITGPPGPPGPAGKDGPRGLRGDVGPVGRTGEQGIAG
COL1A2 Sus DGGPPGATGFPGAAGRIGPPGPSGISGPPGPPGPAGKEGLRGPRGDQGPVGRTGETGASG
COL1A2 Bos DGGPPGATGFPGAAGRTGPPGPSGISGPPGPPGPAGKEGLRGPRGDQGPVGRSGETGASG
COL1A2 Capra DGGPPGATGFPGAAGRTGPPGPAGISGPPGPPGPAGKEGLRGPRGDQGPVGRTGEPGAAG
COL1A2 Ovis DGGPPGATGFPGAAGRTGPPGPAGISGPPGPPGPAGKEGLRGPRGDQGPVGRTGEPGAAG
GLPGVAGSVGEPGPLG
COL1A2 Gallus PPGFAGEKGPSGEAGAAGPPGTPGPQGILGAPGILGLPGSRGERGLPGTAGATGEPGPLG
COL1A2 Sus PPGFAGEKGPSGEPGTAGPPGTPGPQGILGAPGFLGLPGSRGERGLPGVAGSVGEPGPLG
COL1A2 Bos PPGFVGEKGPSGEPGTAGPPGTPGPQGLLGAPGFLGLPGSRGERGLPGVAGSVGEPGPLG
COL1A2 Capra PPGFVGEKGPSGEPGTAGPPGTPGPQGFLGPPGFLGLPGSRGERGLPGVAGSVGEPGPLG
COL1A2 Ovis PPGFVGEKGPSGEPGTAGPPGTPGPQGLLGAPGFLGLPGSRGERGLPGVAGSVGEPGPLG
IAGPPGAR
COL1A2 Gallus VSGPPGARGPSGPVGSPGPNGAPGEAGRDGNPGNDGPPGRDGAPGFKGERGAPGNPGPSG
COL1A2 Sus IAGPPGARGPPGAVGNPGVNGAPGEAGRDGNPGSDGPPGRDGQAGHKGERGYPGNPGPAG
COL1A2 Bos IAGPPGARGPPGNVGNPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNAGPVG
COL1A2 Capra IAGPPGARGPPGNVGNPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNAGPVG
COL1A2 Ovis IAGPPGARGPPGNVGNPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNAGPVG
COL1A2 Gallus ALGAPGPHGQVGPSGKPGNRGDPGPVGPVGPAGAFGPRGLAGPQGPRGEKGEPGDKGHRG
COL1A2 Sus AAGAPGPQGAVGPAGKHGNRGEPGPAGSVGPAGAVGPRGPSGPQGIRGEKGEPGDKGPRG
COL1A2 Bos AAGAPGPQGPVGPVGKHGNRGEPGPAGAVGPAGAVGPRGPSGPQGIRGDKGEPGDKGPRG
COL1A2 Capra AAGAPGPQGPVGPTGKHGSRGEPGPVGAVGPAGAVGPRGPSGPQGIRGDKGEPGDKGPRG
COL1A2 Ovis AAGAPGPQGPVGPTGKHGSRGEPGPVGAVGPAGAVGPRGPSGPQGIRGDKGEPGDKGPRG
IGQPGAVGPAG
COL1A2 Gallus LPGLKGHNGLQGLPGLAGQHGDQGPPGNNGPAGPRGPPGPSGPPGKDGRNGLPGPIGPAG
COL1A2 Sus LPGLKGHNGLQGLPGLAGHHGDQGAPGPVGPAGPRGPAGPSGPAGKDGRTGQPGAVGPAG
COL1A2 Bos LPGLKGHNGLQGLPGLAGHHGDQGAPGAVGPAGPRGPAGPSGPAGKDGRIGQPGAVGPAG
COL1A2 Capra LPGLKGHNGLQGLPGLAGHHGDQGAPGAVGPAGPRGPAGPTGPAGKDGRTGQPGAVGPAG
COL1A2 Ovis LPGLKGHNGLQGLPGLAGHHGDQGAPGAVGPAGPRGPAGPTGPAGKDGRTGQPGAVGPAG
IR
COL1A2 Gallus VRGSHGSQGPAGPPGPPGPPGPPGPNGGGYEVGFDAEYYRADQP----SLRPKDYEVDAT
COL1A2 Sus IRGSQGSQGPAGPPGPPGPPGPPGPSGGGYDFGYEGDFYRADQPRSPPSLRPKDYEVDAT
COL1A2 Bos IRGSQGSQGPAGPPGPPGPPGPPGPSGGGYEFGFDGDFYRADQPRSPTSLRPKDYEVDAT
COL1A2 Capra IRGSQGSQGPAGPPGPPGPPGPPGPSGGGYDFGFDGDFYRADQPRSPASLRPKDYEVDAT
COL1A2 Ovis IRGSQGSQGPAGPPGPPGPPGPPGPSGGGYDFGFDGDFYRADQPRSPASLRPKDYEVDAT
COL1A2 Gallus LKTLNNQIETLLTPEGSKKNPARTCRDLRLSHPEWSSGFYWIDPNQGCTADAIRAYCDFA
COL1A2 Sus LKSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFS
COL1A2 Bos LKSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFS
COL1A2 Capra LKSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFS
COL1A2 Ovis LKSLNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFS

```
COL1A2 Gallus      TGETCIHASLEDIPTKTWYVSKNPKDKKHIWFGETINGGTQFEYNGEGVTTKDMATQLAF
COL1A2 Sus         TGETCIRAQPENIPAKNWYRN--SKVKKHVWLGETINGGTQFEYNMEGVTTKEMATQLAF
COL1A2 Bos         TGETCIRAQPEDIPVKNWYRN--SKAKKHVWVGETINGGTQFEYNVEGVTTKEMATQLAF
COL1A2 Capra       TGETCIRAQPEDIPVKNWYRN--SKAKKHVWVGETINGGTQFEYNVEGVTTKEMATQLAF
COL1A2 Ovis        TGETCIRAQPEDIPVKNWYRN--SKAKKHVWVGETINGGTQFEYNVEGVTTKEMATQLAF
                   ******:*. *:**.*.** .   * ***:*.************* ******:*******

COL1A2 Gallus      MRLLANHASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELRAEGNSRFTFSVLVDG
COL1A2 Sus         MRLLANHASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDG
COL1A2 Bos         MRLLANHASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDG
COL1A2 Capra       MRLLANHASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDG
COL1A2 Ovis        MRLLANHASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDG
                   **************************************** ********::*****

COL1A2 Gallus      CSKKNNKWGKTIIEYRTNKPSRLPILDIAPLDIGGADQEFGLHIGPVCFK
COL1A2 Sus         CSKKTNEWRKTIIEYKTNKPSRLPILDIAPLDIGDADQEVSVDVGPVCFK
COL1A2 Bos         CSKKTNEWQKTIIEYKTNKPSRLPILDIAPLDIGGADQEIRLNIGPVCFK
COL1A2 Capra       CSKKTNEWKKTIIEYKTNKPSRLPILDIAPLDIGGADQEIRLNIGPVCFK
COL1A2 Ovis        CSKKTNEWKKTIIEYKTNKPSRLPILDIAPLDIGGADQEIRLNIGPVCFK
                   ****.*:* ******:****************** ****. :.:******
```

The analysis of the multiple alignments of the collagen alpha-1(III) and collagen alpha-2(I) reveal some peptides that allow to unambiguously assign a *bovine* origin to the proteins (bold in tables 6.3 and 6.4).It is worth mentioning that no unique peptides was observed for any species but *Bos taurus*.

Two unique peptides of bovine collagen α-1 (III) were identified

- R.GPPGPPGTNG**V**PGQR.G + Deamidated (NQ); 3 Hydroxylation (P) (P)
- R.GVAGEPGR**N**GLPGGPGLR.G + Deamidated (NQ); 3 Hydroxylation (P) (P)

Two unique peptides of collagen α-2(I) from *Bos taurus* were identified:

- R.GAPGA**I**GAPGPAGANGDRGEAGPAGPAGPAGPR.G + 2 Hydroxylation (P)
- R.**I**GQPGAVGPAGIR.G + Hydroxylation (P)

A further check of the quality of the fragmentation spectra that allowed this assignment was necessarily carried out, confidently confirming the bovine collagen attribution.

The results herein reported allowed us to identify bovine gut as the constituent material of the membranaceous textile used in the manufacture of the drape found in the tomb of Emperor Henry VII of Luxembourg.

6.6 Conclusions

Animal or vegetable strips covered by wire were the most commonly reported membranaceous threads used in decoration of textiles. (4,5)

As reported in literature, these threads were made with skin or gut of bovines, sheep, goats, pigs or fowls. (6)

The aim of these studies was to characterize the proteinaceous material in the drape found in Pisa cathedral in the tomb of Emperor Henry VII of Luxemburg (3). The material used as strip was identified by the proteomic approach, and confirmed by a clever use of sequence alignement tools. A close inspection of the experimental data, contributed to confidently identify proteotypic peptides that are unique to bovine collagenic sequences.

These studies allowed us to reliably identify bovine gut as the constituent material of the membranaceous textile used in the manufacture of the drape found in the tomb of Emperor Henry VII of Luxembourg.

6.7 References

1. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.
2. Bellis M (2016). "The History of Clothing – How Did Specific Items of Clothing Develop?".The About Group.
3. Lorenzi R (2014) Unique Silk Cloth Found in Emperor Henry VII's Coffin. Hstory.
4. Jarò M (1990) Gold Embroidery and Fabrics in Europe: XI -XIV Centuries. Gold Bulletin (23) 40 -57.
5. Cauzzi D, Chiavari C, Martini C, Prandstraller D (2006) Fili in rame argentato dall'abito di gala della duchessa Maria Luigia di Parma (1791 -1847): indagini sulla tecnologia di produzione e sullo stato di conservazione. Archeometallurgia (9) 25-31.
6. De Reyer D, Jeantet AY, Pilbout S, Anglo A, Monnerot M(2002) Les lamelles del fils métalliques organiques dans les textiles médiévaux: approche méthodologique de leur origine biologique. Studies in Conservation (47)122-133.
7. Welker F, Collins MJ, Thomas JA, Wadsley M, Brace S, Cappellini E, et al. (2015) Ancient proteins resolve the evolutionary history of Darwin/'s south American ungulates. Nature (522)81–84.
8. Fremout W, Dhaenensc M, Saverwynsa S, Sanyovaa J, Vandenabeeled P, Deforcec D, Moensb L (2012) Development of a dedicated peptide tandem mass spectral library for conservation science. Analytica Chimica ActaVolume (728) 39–48.
9. Waters MR et al. (2011) Pre-Clovis Mastodon Hunting 13,800 Years Ago at the Manis Site, Washington. Sci. (334) 351–353.
10. Welker F. et al. (2015) Ancient proteins resolve the evolutionary history of Darwin/'s South American ungulates. Nature (522) 81–84.

7. Identification of remains in archaeological potsherds by mass spectrometry

For ages, human activity left material traces of its creative intelligence. Art or archaeological remains can be found, preserved, and studied revealing us, as real witnesses of a period, invaluable historical, socioeconomic, and cultural information. However, the detailed study of organic material can be a real analytical challenge, especially in archaeological remains, that have been exposed to harsh conditions during burial and excavations. Protein study and identification are very challenging in archaeological potsherds. The difficulties result principally from sample burials and leaching of hydrophilic molecules. As a consequence, most of the studies deal with hydrophobic, mainly lipidic fractions. (1)

Ceramics, amphorae and other tools, are among the most abundant objects found during archaeological excavations and can be considered as real witnesses of the uses and habits of the old civilizations. Their typology and their production sites are now quite well understood, although several questions still remain unanswered about their content (like examples oil, wine, fish sauce, etc.), and therefore their uses. (2)

The current chapter present a characterization of a collection of amphorae and tools from two archaeological sites: Eastern Sudan and Cuma (Italy).

Particularly, a collection of 8 potsherds of amphorae from Sudan and, 4 amphora's fragments recovered in Cuma were analysed to establish their contents. Moreover, 7 pieces of tools made in stone from the same archaeological site in Sudan were analysed too. The aim of the analysis of these tools, were to understand the ancient uses of their.

The traditional approach, consisting in studying the shape, the economic context of the production area, or the workshop situation, is not always successful in solving this problem and the use of analytical chemistry techniques can be crucial. However, because of the small quantities of organic traces preserved and their degradation state, analysis of archaeological objects represents a real challenge. (2)

Therefore, all the samples were analysed with the aim of establishing a relationship between the vessel and its use, specific experimental procedures, based on mass spectrometric

technologies have been developed to analyze the traces of different organic compounds that may be present in these archaeological potsherds. (2)

Literature data were reported below, with the aim of understanding the typical contents and the uses of the amphorae and tools during the centuries.

7.1 Amphoraes

Pottery is one of the most copiously available items encountered in an excavation because of the natural wear of ceramic containers during their lifetime. Pots break and are replaced with others. Some of them were manufactured over a long period of time, some ameliorated in style and form by the urge of fashion trends, or some were corrected because they did not live up to the specific functions they were designed for. (3)

Obvious functions are those represented by cooking pots that pertain to cooking a stew or boiling water. Storage jars were used to carry liquids as oil, wine and water as well as recipients for preserving wheat, salted fish and meat and still more. Furthermore, pouring utensils, e.g. decanters that can be used at the table and lamps to light up one table, a room or a cave. Other pottery as bowls, cups, jugs and the like were used for a variety of functions such as mixing food, serving it, pouring liquids and eating or drinking from them. (3)

Through the changes in the various styles of ceramic vessels, it is possible to follow how man saw the changing situations around him. Each material in nature has its own characteristics in terms of strength, workability, colour and texture that may be used in the manufacture of every possible utensil for daily use. If the ancients changed their utensils in form or size or the style was abandoned, we have to search for how well the previous tool or vessel performed in order to be able to speculate on its change into some other shape. (3)

7.2 Tools

Quite a lot is known about ancient tools thanks to the importance the Egyptians attributed to their use in the next world. The graves of craftsmen often contained tools or models of tools, and tomb walls were at times decorated with scenes of artisans at work demonstrating their techniques. In addition, just to make sure that one would not be left without the necessary implements some had lists of tools carved into the walls. (4)

Tools were made especially in wood, iron, copper, bronze or stone. Our attention was focused on some stone tools. (4)

7.2.1 Stone tools

Stone has basically three functions for which it is suited: pounding, grinding and cutting. A relatively high specific weight and hardness give it much more impact power than a similarly sized piece of timber would have. This hardness is also what restricts its use as a hammer, because it splinters easily. The stone was either chipped or ground into the desired shape depending on the kind of stone: Fine sandstone, limestone and the like were ground serving as grinding stones and the like, while flint was generally chipped and used for cutting. Some materials like granite could only be worked with spherical hammer stones made of diorite, a stone of even greater hardness. Thanks to their roundness and composition these hammer stones rarely splintered. Applied with measured force they were used to slowly pulverize and shape the workpiece. The possibility of creating cutting edges is due to the hardness and crystallinity of the stone. It would be wrong to think that one could just bend down, pick up any stone and make a blade of it. Few kinds of stone are suitable for knapping, the most widely used being flint and chert. Opaque flint was preferred by the toolmakers to the clearer varieties. It seems that the resulting flakes were longer and straighter, their edges tougher and they did not splinter as easily. (4)

Because of the flint's brittleness little pressure should be exerted on thin blades. Stone knives had therefore often no real handle. At one end the blade was left blunt for a few centimetres and wrapped with some fabric and string to give a minimal amount of grip, just enough for cutting, but not enough for breaking the blade. But the knowledge of how to make proper handles and how to fasten them to blades dates back to prehistoric days. (4)

The amount of work a knapper invested in making a tool was dependent on the length of time it could be expected to be used. Axes which received harsh treatment and therefore broke quickly were generally fashioned with a few well-placed strokes [5]. The heads were fastened to the handles by cutting a socket into the wood, inserting the blade and tying it with a cord two or three feet in length. No cement was used. (4)

Broken tools were often reshaped and dull edges resharpened. Axe heads were sometimes ground down to such an extent that little stone was left protruding from the socket, before they were discarded. (4)

7.3 Potsheds content

Because of their hydrophobicity, lipids have a strong capacity to resist natural leaching and therefore these organic compounds have been widely investigated, indicating the presence of

oils, fats or waxes in ceramics. Studies looking for other classes of organic compounds such as the degraded products from tannins and polyphenols were carried out to reveal the presence of residues of wine products in the amphorae. Archaeobotanical methods based on pollen analysis were also considered to highlight the presence of plant-based beverages such as wine and med in archaeological containers. Regarding marine products such as fish and shells, traces of polyunsaturated fatty acids in archaeological pottery have been described in literature as specific marker. (2)

Additionally, encouraging data were obtained from analysis of ancient proteins. For archaeological vessels, the proteomic approach was applied for the first time on potsherd from Alaska. The analysis was performed on samples without visible residues and, nevertheless succeeded in identifying proteins from muscles of various fish species. (2)

7.4 Materials and Methods

7.4.1 Samples

Several samples from excavations Eastern Sudan, dating back to around IV millennium B.C. were provided by Prof. Andrea Manzo, Università "L'Orientale", Napoli. Two types of samples were available: amphorae and tools. Moreover, 4 samples from Cuma excavation were analysed in collaboration with Prof. Giovanna Greco, Dip. Scienze Umanistiche, Università degli studi di Napoli "Federico II".

7.4.2 Proteomic analyses

Aliquots of each sample were treated in heterogeneous phase with trypsin with the minimally invasive protocol as indicated in the chapter 4, and the resulting peptide mixtures were analysed by LC-MS/MS.

In particular, the peptide mixture (1µl) was analyzed by nanoLC-MS/MS, on a CHIP MS 6520 QTOF equipped with a capillary 1200 HPLC system and a chip cube (Agilent Technologies, Palo Alto, CA) as indicated for proteomics analysis in the previously chapters.

The acquired MS/MS spectra were transformed in Mascot Generic files (.mgf) format and used to query the NCBIr database, without taxonomic restriction for protein identification with licensed version of MASCOT software (www.matrixscience.com) version 2.4.0 with the same parameters as reported in the chapter 4 No fixed chemical modification was inserted, but possible oxidation of methionine residues, and deamidation at asparagines and glutamines were considered as variable modifications. Only proteins presenting two or more peptides

were considered as positively identified. Spectra with MASCOT score of <10, having low quality were rejected.

7.4.3 Chemical analysis of organic molecules

Reagents.

Chloroform and methanol were from Sigma; hexane was purchased from Romil; deionized water was obtained from Millipore cartridge equipment.

Samples treatment

The samples were extracted with 1 mL of chloroform-methanol solution (2:1 v/v; 3x 60 min sonication). After centrifugation (20 min, 5000 rpm) the supernatant solvent was removed from the residue and placed in a pirex. The three extracts were combined and the solvent removed by evaporation under a gentle stream of nitrogen and then extracted with hexane (5ml; 2x 3 hours sonication). The extracts were combined and the solvent removed by evaporation under a gentle stream of nitrogen. Afterwards, the samples were redissolved in hexane (500 µL) and analyzed by GC/MS.

GC-MS analysis. Analyses were performed with an Agilent 6890 Series GC, coupled to a detector MSD quadrupole mass spectrometer 5973, equipped with a gas chromatograph by using a Zebron ZB-5HT Inferno (5%-Phenyl-95%-Dimethylpolysiloxane) fused silica capillary column (Column 30 m x 0.32 mm x 0.10 µm) from Phenomenex. The stationary phase used consists of polydimethylsiloxane containing 5% of a phenyl derivative.

The injection temperature was 250 °C. The oven temperature was increased from 90 °C to 140 °C at 10 °Cmin^{-1} held for 1 min, increasing to 280 °C at 20 °Cmin^{-1} and held for 3 min, and finally to 300 °C at 10 °C min^{-1} held for 5 min and remaining at this temperature for 4 min; carrier gas (He) at a constant flow rate of 1 mL min-1. MS interface was kept at 250 °C; electron ionization source was set at290 °C, +70eV (30-450 *m/z*) and electron multiplier was at 1.7 kV. Each species was interpreted on the basis of electron impact spectra (NIST library 2008).

7.5 Results e discussions

Chemical analyses were performed on a collection of Amphorae and tools from Sudan and one amphora from Cuma.

The investigation was carried out in two analytical directions. The samples were submitted to proteomics analysis and, in parallel, preliminary analyses were conduced on the same samples by GC-MS to determine free fatty acids. Indeed, it is possible that fats are highly hidrolysed in archaeological samples.

7.5.1 Tools from Sudan

In a first instance, the standard minimally invasive treatment based on the enzymatic digestion and subsequent analysis by LC-MS/MS and database search with MASCOT software was carried out in the search of proteins. Proteins were identified in NCBI database, with methionine oxidation and deamidation at asparagine and glutammine as variable modifications. Since no hypothesys of the possible use of the tools had been made in advance, that meant we had no idea of the material that has come into contact with, the whole database, with no restriction, was used in a first database search.

Interestingly, it was possible to identify proteins from cotton (*Gossypium*) in one sample (4Ua), suggesting a possible use of the tool in the cotton manufacture. The sample 4U is a sliver of stone that measures approximately 2/3 cm and having sharp sides.

To refine the protein identification, once we had established that proteins were from Cotton, a further database search was carried out in a database restricted to viridiplantae as taxonomic restriction, improving the match scores and therefore, the reliability of the identifications.

The identification of proteins shows these overall results:

Table 7.1: Sample 4Ua. Identification of proteins after trypsin digestion. Peptide mixtures were analysed by LC-MS/MS and raw data searched by MASCOT software. Proteins were identified in NCBI database, with *Viridieplantae* as taxonomic restriction and with methionine oxidation and deamidation at asparagine and glutammine as variable modifications.

Protein	Score	Peptide	Peptide Score
Vicilin A (*Gossypium hirsutum*)	398	K.IYVVTNGR.G + Deamidated (NQ)	34
		R.QQGQGMFR.K	22
		K.ELAFGVESR.L	41
		R.EILEAVFNTR.S	14
		R.SEQLDELFGGR.Q	68
		K.ESYNVVPGVVVR.I	44
		R.EEEEQEEQEVER.R	75
		R.EEEEEESDEGEQQQR.N	37
Vicilin C72 (*Gossypium hirsutum*)	302	K.IYLVTNGR.G	37
		R.QQGQGMFR.K	22
		K.ELAFGVSSR.L	23
		R.ALSQEATSPR.E + Deamidated (NQ)	39
		R.EILEPAFNTR.S	42
		R.SEQLDELFGGR.Q	68
		K.ESYNVVPGVVVR.V	44
Vicilin A (*Gossypium herbaceum*)	285	K.IYVVTNGR.G + Deamidated (NQ)	34
		R.QQGQGMFR.K	22
		K.ELAFGVESR.L	41
		R.ALSQEATSPR.E + Deamidated (NQ)	39
		R.EILEAVFNTR.S	14
		R.SEQLDELFGGR.Q	68
		K.ESYNVAPGVVVR.I	29
		R.EEEEEESDEGEQQQR.N	37
Legumin B (*Gossypium hirsutum*)	147	R.TPASSADVFNPR.G	75
		R.EEEESQESGGNNVLSGFR.D	72
Legumin precursor (*Gossypium hirsutum*)	116	R.LELSAER.G	46
		R.LQVVTPPR.M	31
		R.ADIFNPQAGR.I	28
Oleosin (*Gossypium hirsutum*)	84	R.TFDQVVR.G	39
		R.QATGTEQLDADR.A	44

Literature data reports that cotton was used in Africa, India and China until 5000 AD (5). As above mentioned, samples analyzed traced back to 4 millennia AD and come from Sudan, so it is possible that this tool were used for harvesting or manufacturing cotton.

No proteins were detected in any of the other samples, nor any protein was confidently identified when digestion with trypsin was carried out after Urea 6M sample pre-treatment as indicated in the chapter 4.

Non proteinoceous molecules were then searched on other aliquots of the samples. Briefly, powder on the ancient tools fragment was extracted with chloroform-methanol solution and then with hexane. The preparation of samples was followed by GC/MS analysis.

The results obtained are summarized in Tables 7.2 to 7.7

Table 7.2 : 1U sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
20.80	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
22.12	10,13-octadecadiynoic acid, methylester
22.26	5,8,11-heptadecatriynoic acid, methylester
22.36	pregn-4-ene-3,20-dione-16-hydroxy-(16α)
22.51	cedran-diol,8S,13
22.95	androst-5-ene-3,17,19-triol
23.15	13alpha-delta(8)-dihydroebietic acid
23.34	10,18-bisnorabieta-8,11,13-triene
23.56	octadecanoic acid-4-hydroxy-methylester

Table 7.3 : 2U sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
20.61	benzestrol
20.75	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
22.12	10,13-octadecadiynoic acid, methylester
22.26	5,8,11-heptadecatriynoic acid, methylester
22.38	hexadecanoic acid, methylester
22.51	cedran-diol,8S,13
22.95	primaric acid

Table 7.4 : 3U sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
20.743	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
21.036	7-methyl-Z-tetradecen-1-olacetate
22.110	10,13-octadecadiynoic acid, methylester
22.359	pregn-4-ene-3,20-dione-16-hydroxy-(16α)
22.952	androst-5-ene-3,17,19-triol
23.339	10,18-bisnorabieta-8,11,13-triene
23.399	tetrahydroaraucaolone
23.563	10-octadecanoic acid-methylester
24.276	5,8,11,14-eicosatetraynoic acid
25.643	palustric acid

Table 7.5 : 4Ua sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
18.509	2,5-di-tert-butyl-1,4-benzoquinone
20.597	benzestrol
20.743	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
22.110	10,13-octadecadiynoic acid, methylester
22.205	20-methylpregn-5-en-3-beta-16alpha-20-triol
22.256	ergosta-5,22-dien-3-ol,acetate(3β,22E)
22.359	pregn-4-ene-3,20-dione-16-hydroxy-(16α)
22.523	cedran-diol,8S,13
22.944	pimaric acid
23.150	13alpha-delta(8)-dihydroabietic acid
23.339	10,18-bisnorabieta-8,11,13-triene
23.399	lycoramine-6-O-dimethyl-3-desoxy
23.563	10-octadecanoic acid-methylester
24.534	bufa-14,20,22-trienolide-3-hydroxy(3β,5β)
24.697	ethyliso-albcholate
24.869	retinal,9-cis
25.626	palustric acid
27.938	5α-cholest-8-ene-3β,6α-diol-14-methyl, diacetate

Table 7.6 : 4Ub sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
20.288	5-(4,5,6,7-tetrahydrobenzofuran-2-yl)pentanoic acid
20.606	benzestrol
20.743	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
22.119	10,13-octadecadiynoic acid, methylester
22.248	pyrrolidine-1-|4-(4-chlorophenyl)-3-phenyl-2-butenyl
22.368	pentadecanoic acid-14-methyl,methylester
22.506	cedran-diol,8S,13
22.944	pimaric acid
23.339	10,18-bisnorabieta-8,11,13-triene
25.609	palustric acid

Table 7.7 : 5U sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention time	Molecules
20.606	benzestrol
20.743	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane)-1',4',2,6,6-pentamethyl
20.855	phenol-2-methyl-4-(1,1,3,3-tetramethylbutyl)
20.924	benzestrol
22.119	10,13-octadecadiynoic acid, methylester
22.256	1-methyl-10,18-bisnorabieta-8,11,13-triene
22.514	cedran-diol,8S,13
22.875	1,4-hexadien-3-one-5-methyl-1-|2,6,6-trimethyl-2,4-cyclohexadien-1-yl|
22.953	androst-5-ene-3,17,19-triol
23.150	13alphadelta(8)-dihydroabietic acid
23.279	1-heptatriacotanol
23.339	10,18-bisnorabieta-8,11,13-triene
23.408	lycoramine-6-0-demethyl-3-desoxy
23.563	10-octadecanoic acid,methylester
23.709	antracene-9-dodeciltetradecahydro
24.276	10,18-bisnorabieta-5,7,9(10),11,13-pentaene
24.345	17-methoxy-d-omo-18noradrosta,4,8,13,15,17-pentaen-3-one
25.643	palustric acid

In the samples analysed, hormones and fatty acid were found. The presence of these ones could be indicative of the use of the tools for hunting, fishing, manufacturing leather and food.

The molecules identified in the samples 4Ua, that containing also cotton proteins, suggest that these tools were simultaneously used for different uses. Another hypothesis is that this tool had a handle made in cotton. Furter experiments will be necessary to exclude that hormones identification is due to contamination processes.

In some samples, the presence of resinous materials, could be indicative of the use of these substances as binder to tie different parts of the tools, for example join a tip to a cane obtaining a lance for fishing or hunting.

A careful inspection of the results by the archaeologists, considering historical and geographical information is now urgently required to postulate an hypothesis of the actual use of the tools.

7.5.2 Amphorae from Sudan

8 samples from potsherds of amphorae were analysed in order to detect any chemical compound that could help in defying their use

Proteomics analysis, that were carried out as first approach, produced no conclusive results. Then, the described methodology developed for the identification of non proteinaceous molecules was applied. Briefly, the ancient amphorae fragments were crushed into fine powder, that were subsequently extracted with chloroform-methanol solution and then with hexane. The extracts were analyzed by GC/MS.

The results obtained are summarized in Tables 7.8 to 7.16.

Table 7.8 : A1 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
19.40	tetradecane
20.92	hexadecane
21.05	7-methyl-Z-tetradecen-1-olacetate
21.26	eicosane, 2-methyl
22.62	octadecane, 3-ethyl-5-(2-ethylbutyl)
22.85	heptadecane, 2, 6, 10, 15-tetramethyl
23.34	10, 18-bis norabieta-8, 11, 13-triene
24.24	heptacosane

Table 7.9 : sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20.93	9-octadecen-12-ynoic acid, methylester
21.05	Z-8-methyl-9-tetradecenoic acid
21.26	octadecane, 6-methyl
22.37	pentadecanoic acid, 14-methyl-, methylester
22.58	octadecane, 3-ethyl-5-(2-ethylbutyl)
22.84	2-1-(2-1,3-dithian-2-yl-ethyl)-pent-4-enyloxy-tetrahydropyran
23.22	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene

Table 7.10 : A3 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20.69	6-(1-adamantylamino)-2, 4, 5-trichloronicotinonitrile
20.75	spiro-1-(ciclohex-2-ene)-2-(oxabicyclopentane), 1, 4, 2, 6, 6-pentamethyl
20.81	1-cyclopropene-1-pentanol, α, ε, ε, 2-tetramethyl-3-(1-methylethenyl)
21.04	1, 6-dioxacyclododecane-7, 12-dione
22.11	10, 13-octadecadiynoic acid, methylester
22.36	cyclopropane butanoicacid,pentyl, methylcyclopropyl-methylester
22.62	cholesta-8, 24-dien-3-ol, 4-methyl-(3β, 4α)
23.21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.79	bicyclo heptan-2-one, 1, 7, 7-trimethyl-3-(phenylmethylene)
24.20	5, 8, 11, 14-eicosatetraynoic acid
24.84	ethyliso-allocholato
26.00	5α-cholest-8-en-3-one, 6α-hydroxy-14-methyl, acetate

Table 7.11: A4 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20.74	spiro-1-(ciclohex-2-ene)-2-(oxabicyclopentane), 1, 4, 2, 6, 6-pentamethyl
20.81	cyclohexanone, 2-(2-butynyl)
20.92	1-cyclopropene-1-pentanol, α, ε, ε, 2-tetramethyl-3-(1-methylethenyl)
21.03	1, 6-dioxacyclododecane-7, 12-dione
21.14	4α-dichloromethyl-4, 4α, 5, 6, 7, 8-hexahydro-3H-naphthalen-2-one
22.11	10, 13-octadecadiynoic acid, methylester
22.35	hexadecanoic acid, methylester
22.51	10-heptadecen-8-ynoic acid, methylester
22.56	octadecane, 3-ethyl-5-(2-ethylbutyl)
22.62	androstan-17-one, 3-ethyl-3-hydroxy-(5α)
22.74	1-heptatriacotanol
22.86	cholesta-8, 24-dien-3-ol, 4-methyl-(3β, 4α)
23.21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.71	cyclopropane butanoicacid,pentyl, methylcyclopropyl-methylester
23.79	5, 8, 11, heptadecatriynoic acid, methylester
24.19	5-(2-(5-oxopenta-1, 3-dienyl)phenyl)penta-2, 4-dienol
25.29	ethyliso-allocholato
26.00	5α-cholest-8-en-3-one, 6α-hydroxy-14-methyl, acetate

Table 7.12: sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20.81	4-(3, 3-dimethyl-but-1-ynyl)-4-hydroxy-2, 6, 6-trimethylcyclohex-2-enone
21.04	1, 6-dioxacyclododecane-7, 12-dione
22.15	10, 13-octadecadiynoic acid, methylester
22.62	cholesta-8, 24-dien-3-ol, 4-methyl-(3β, 4α)
23.21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.78	5, 8, 11-heptadecatriynoic acid, methylester
26.00	5α-cholest-8-en-3-one, 6α-hydroxy-14-methyl, acetate

Table 7.13: A6 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
21.03	1, 6-dioxocyclododecane-7, 12-dione
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.78	acetic acid,17-(1-acetyloxy-ethyl)-10, 13-dimethyl-3-oxododecahydro…

Table 7.14: A7 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20.69	6-(1-adamantylamino)-2, 4, 5-trichloronicotinonitrile
21.03	1, 6-dioxocyclododecane-7, 12-dione
22.11	10, 13-octadecadiynoic acid, methylester
22.86	1, 4-hexadien-3-one, 5-methyl-1-(trimethyl-2, 4-cyclohexadien-1-yl)
22.94	1-heptatriocotanol
23.21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.79	cyclohexane, hexaethyldene
24.2	5, 8, 11, 14-eicosatetraynoic acid
26.00	5α-cholest-8-en-3-one, 6α-hydroxy-14-methyl, acetate

Table 7.15: A8 sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
21.03	1, 6-dioxocyclododecane-7, 12-dione
22.11	10, 13-octadecadiynoic acid, methylester
22.94	1-heptatriocotanol
23.21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23.33	10, 18-bis norabieta-8, 11, 13-triene
23.79	5, 8, 11-heptadecatriynoic acid, methylester

On the basis of the molecules identified in each sample, we propose the following hypotheses:

- Sample 1A: The main product detected were hydrocarbons. Literature data reports that terpenes, glicerides, phenols, but also hydrocarbons are components of essential oils. (6) On the basis of these information, we suggest that this amphora, could be used to store essential oils but additional experiments will be necessary to search for the other molecules typical of essential oils.

- Samples 2A, 4A, 5A, 8A: Steroids and fatty acids were detected. As reported in literature, animal lipids have a strong capacity to resist to natural leaching, because of their hydrophobicity, and they are best preserved in ancient amphorae. They are stably trapped inside the porous of the amphorae and they are protected from degradation. (7) Indeed, saturated fatty acids like pentadecanoic acid ($C_{15:0}$), palmitic acid ($C_{16:0}$), margaric acid ($C_{17:0}$) and oleic acid ($C_{18:0}$) suggest the presence of milk and its derivatives. However, these analyses are not complete enough to highlight the presence of lipids and further experiments will be essential to validate our hypotheses.

- Samples 3A, 7A: Steroids and two fatty acids, Linoleic and Arachidonic acid, were detected as major component. Literature data reports that their simultaneous presence is suggestive of the presence of fishing products (9), leading to the hypothesis that these amphorae was used to store fish.
- Sample 6A: In this sample a derivative of acetic acid was found. This compound could be a degradative product of ethanol, possible residue of alcoholic drinks, such as wine and beer. This amphorae could therefore be used to store wine or beer.

The hyphotesis reported are summarized in the table 7.16

Table 7.16: Amphorae from Sudan: Possible content of amphorae as suggested by GC/MS analysis.

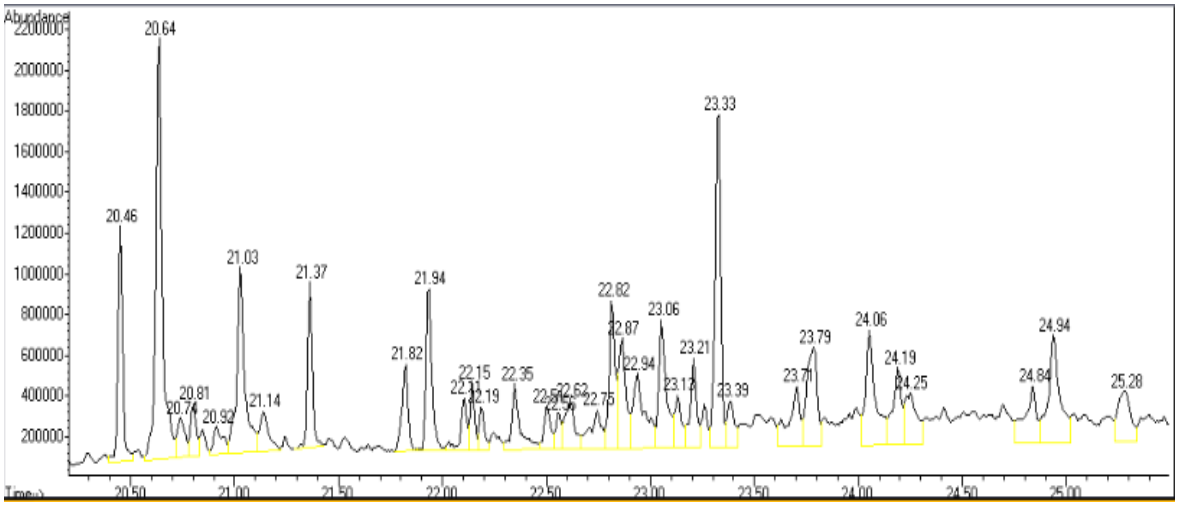
Sample	Content
1A	Essential Oil
2A	Milk or derivative
3A	Fishing products
4A	Milk or derivative
5A	Milk or derivative
6A	Wine or beer
7A	Fishing products
8A	Milk or derivative

Moreover, in all the samples, compounds that are derived from Abietic acid were found. Abietic acid was found in all resins produced by trees of *Pinaceae* family.

These resins is protective and wood preservatives, and is produced by parenchymatous epithelial cells that surround the resin ducts in trees of coniferous forests. It is used as a component of adhesives, rubbers, inks, and as emulsifier. Moreover, Colofony was used to seal amphorae in ancient times.(10-12)

Chromatogram of sample 4A was reported as example.

Figure 7.1: A4 sample: GC chromatogram of the sample from chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.



The preocedure used was focused on the research of free fatty acids and does not allow to identify lipids because, even though these ones were extracted with this procedure, they would not be detected by GC/MS because glycerides are too heavy. Indeed, further experiments using procedures like hydrolysis or derivatization, will be indispensable to validate our hypotheses.

Next step for a confident and reliable attribution of the use of the amphorae will be the discussiom with the archaeologists, that will contribute with historical and geographical information.

7.5.3 Amphorae from Cuma

A similar approach as described above was carried out on a collection of potsherds of one amphora (4 samples) from Cuma.

Also in this case, proteomics analysis produced no conclusive results, and therefore, a GC-MS analysis of the extract in organic solvent was carried out.

The main compounds identified in the GC-MS analysis are summarized in Tables 7.17 to 7.20.

Table 7.17: Bottom sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Bottom	
Retention time	Compound
20.75	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane),1',4,2,6,6-pentamethyl
22.12	10,13-octadecadynoic acid, methylester
22.37	pentadecanoic acid, 14-methyl,methylester
22.51	cedran-diol 8S,13
22.95	androst-5-ene-3,17,19-triol
23.22	pregna-5,16-dien-20-one,3-(acethyloxy)-16-methyl-(3β)
23.34	10,18-bisnorabieta-8,11,13-triene
24.55	bufa-20,22-dienoide,3,14-dihydroxy-(3β,5β)
25.05	pregna-4,6-diene-3-20-dione,17-(acetyloxy)6,16-dimethyl-(16α)
25.38	spirost-8-en-11-one,3-hydroxy-(3β,5α,14β,20β,22β,25R)
25.65	palustric acid
26.07	hexadecanoic acid,1(1-methylethyl)-1,2-ethanediylester

Table 7.18: Potsheds from kitchen sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Potsheds from kitchen	
Retention time	Compound
22.13	10,13-octadecadynoic acid, methylester
22.26	7,9-di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione
22.40	hexadecadionoic acid, methylester
22.53	spiro|androst-5-ene-17,1'-cyclobutan|-2'-one,3-hydroxy-(3β,17β)
23.29	1-heptatriacotanol
23.35	10,18-bisnorabieta-8,11,13-triene
23.41	d-alb-dec-2-enoic acid,5,8-anhydro-2,3,4,9-tetradeoxy…
25.66	palustric acid

Table 7.19: Edge from kitchen sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Edge	
Retention time	Compound
20.74	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane),1',4,2,6,6-pentamethyl
22.13	10,13-octadecadynoic acid, methylester
22.36	pentadecanoic acid,1,4-methyl,methylester
22.61	hexadecadionoic acid 1-(hydroxymethyl)-1,2-ethanediylester
23.22	pregna-5,6-dien-20-one,3-(acetiloxy)-16-methyl…
23.34	10,18-bis norabieta-8,11,13-triene

Table 7.20: Lid from kitchen sample: Molecules identified by GC-MS, analysis of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Lid	
Retention time	Compound
18,53	2-butanone,4-(2,6,6-trymethyl-2-cyclohexen-1-ylidene
18,87	phenol-2,4-bis(1,1-dimethylethyl)
19,82	phenol-2,4,6-tris-(1,1-dimethylethyl)
20,31	benzene1-(1,1-dimethylethyl)4-(2-ethoxyethoxy)
20,46	benzene,1,2,4,5-tetrakis(1-methylethyl)
20,61	benzestrol
20,70	phenol,4-(1,1,3,3-tetramethylbutyl)
20,75	spiro-1-(cyclohex-2-ene)-2'-(5'-oxabicyclo|2.1.0|pentane),1',4,2,6,6-pentamethyl
20,92	benzestrol
22,13	10,13-octadecadynoic acid, methylester
22,26	1-heptatriacotanol
22,37	pregn-4-ene-3,20-dione,16-hydroxy-(16α)
22,52	cedran diol,8S,13
22,95	pimaric acid
23,15	1-phenantrenecarboxylic acid,7-ethenyl-1,2,3,4,4a,4b…
23,23	2-hydroxy-3-methylanthraquinone,o-trymethylsilyl
23,28	1-heptatriacotanol
23,34	10,18-bisnorabieta-8,11,13-triene
23,41	lycoramine,6-o-dimethyl-3-desoxy
23,81	3-methoxy-17aβ-methyl-17aα-acetoxy-d-homoestra-1,3,5(10),8-tetraen
24,28	9-α-hydroxyarteethen
24,35	acetic acid, 17-(1-acetoxyethyl)10,13-dimethyl-3-oxo…
24,53	ethyiso-albcholate
25,05	oxazole,4-carboxylic acid,5-(2-fluoro-4,5-dimethoxy-ethylester
25,17	naphathelene-.6-dicarboxylic acid,pentylester-4-pentyl-pentylester
25,53	bufa-20,22-dienolide,3,14-dihydroxy-(3β,5β)

In three samples (all the samples except Lid), steroids and fatty acids were detected. We hypotized that milk or its derivative were contained in the amphora from Cuma. Similarly as what observed for the amphorae from Sudan, further experiments will be necessary to validate our hypotesis.

In all the samples, beside Abietic acid derivatives, also Palustric and Pimaric acid, other typical compounds in resin acid, were also found. These resins were used to seal and waterproof the amphorae, to preserve and protect the material inside.

Also in this case, data obtained will be discussed with the archaeologists, to circumstantiate with historical and geographical information the chemical data.

7.6 Conclusions

Archaeological remains can be considered as real witnesses of a period, providing invaluable historical, socioeconomic, and cultural information. Ceramics, amphorae and other tools, that are among the most abundant objects found during archaeological excavations can reveal uses and habits of the oldest civilities. However, the precise study of organic material in cultural heritage pieces can be a real analytical challenge, especially because exposure to environmental factors during burial and the natural leaching of hydrophilic molecules. (1)

With the aim of establishing a relationship between the crockery pieces and their use, (2) proteomic and GC-MS analyses were used to identify organic materials eventually left over the potsherd of Amphorae and tools from Sudan and one amphora from Cuma.

The analyses conduced, suggest that the tools could have been used for hunting, fishing and for manufacturing leather and food, while amphorae could have been used to store a relatively wide range of materials, including milk, fish and essential oils.However, additional experiments will be essential to validate our hypotheses.

7.7 References

1. Solazzo C, Fitzhugh WW, Rolando C, Tokarski C (2008) Identification of protein remains in archaeological potsherds by proteomics. Anal Chem(80)4590-4597.
2. Dallongeville S, Garnier N, Casasola DB, Bonifay M, Rolando C, Tokarski C (2011) Dealing with the identification of protein species in ancient amphorae. Anal Bioanal Chem. (99) 3053-3063
3. Gunneweg J. (2016) Clay, Technology and Qumran Pottery: What to verify and publish ? Researchgate.
4. Cagle AJ (2001) The Spatial Structure of Kom el-Hisn: An Old Kingdom Town in the Western Nile Delta, Egypt. Artifact Analysis, dissertation, University of Washington.
5. Schlingloff D (1974) Cotton-Manufacture in Ancient India. *Journal of Economic and Social History of the Orient*, (17).
6. Al-Qudah MA, Abu Zarga MH (2010) Chemical composition of Essential oils from Aerial Parts of SisymbriumIrio from Jordan. E-Journal of Chemistry.
7. Correa-Ascencio M, Evershed RP (2014)High throughput screening of organic residues in archaeological potsherds using direct acidified methanol extraction. The Royal Society of Chemistry.
8. Dudd SN, Evershed RP (1998) Direct Demonstration of Milk as an Element of Archaeological Economies. Science (282) 1478-1481.
9. Gunstone FD (1996) Fatty acid and Lipid Chemistry. Chapman & Hall.
10. Stenius (2000). Forest Products Chemistry. Papermaing Science and Technology. (3). 73-76.
11. Norlin LH (2002) "Tall Oil" in Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH, Weinheim
12. Ansaldi M (2002) Il verde e il nero Myristica.

8. Characterization of Embalming Materials in Animal Mummies

Despite the term ‘Egyptian mummification’ is often seen as a very narrow and specific definition, even within the study of ancient Egypt itself, the truth is that it evolved, and indeed regressed, over time. Within the mummification practice there were inevitable variations, despite some general consistencies, and it was eventually confined to the privileged elite for much of its history—through whose eyes much of the received wisdom continues to be seen—with the particular status of individuals inevitably affecting the treatment they ultimately received. (1)

8.1 History

The ancient Egyptians believed that the preservation of the physical body was essential for the continuation of life after death. The body of the deceased had to be treated to render it incorruptible. Ancient Egyptian written sources do not report how mummification was carried out; and reliefs and paintings contain no narrative descriptions of embalming techniques. The only technical data on mummification come from Herodotus (II, 85-88), Diodorus Siculus (I, 91) and a few other classical authors. Illustrations relating to mummification are extremely rare. Some of the scenes from the wooden coffin of Djedbastiufankh (Late Period) in the Hildesheim Museum could be interpreted as being part of the embalming process. Other information about embalming comes from the text of the Ritual of Embalming, which survives in two copies in the Papyrus Boulaq 3 and the Papyrus Louvre 5158, both from the Roman period.(2,3) The examination of mummies from different periods has yielded further knowledge about the technique of mummification throughout the history of ancient Egypt. The data obtained show that the method described by Herodotus and Diodorus Siculus agrees more with that used during the New Kingdom and the Late Egyptian periods. (4)

Where scholarly and scientific studies on Egyptian mummies have been carried out they have historically tended to have an anatomical and medical perspective(5,6) which while unquestionably providing valuable insights has necessarily only limited answers on the embalming materials themselves and the procedures. (1)

8.2 Animal mummies

Throughout the history of ancient Egypt, animals were highly respected. In no other culture animals have been so influential in so many aspects of life, nor has any culture depicted animals as often in their artwork or writing.It is estimated that 2 in every 4 or 5 Egyptian hieroglyphs relates to animals. Egyptians believed that animals were crucial to both physical and spiritual survival-vital to physical survival because they were a major source of food and to spiritual survival based on how well a person treated animals during their life on earth.[2] Some animals were considered to be incarnations of the gods, and therefore, it is understandable why Egyptians would have wanted to hold such animals in the highest consideration, giving them a proper burial through mummification. The Egyptian religion taught of life after death. In order to determine a person's admittance or denial to the afterlife, the gods would ask a series of judgment questions. One of these crucial questions would be whether they had mistreated any animals during their life on earth. Because of this religious belief, the killing of an animal was considered a serious crime punishable by death.(7)

Animal mummification was an enormous part of Egyptian culture. Animals were typically mummified for four main purposes—to allow beloved pets to go on to the afterlife, to provide food in the afterlife, to act as offerings to a particular god, and because some were seen as physical manifestations of specific gods.(7)

For example, cats were seen as the incarnation of Bastet, goddess of music and joy and protector of women. The Apis bull, a sacred animal to the Egyptians, came to be known as the incarnation of Osiris, god of embalming and cemeteries. Likewise, ancient Egyptians associated hawks with Horus (the god of light), ibises with Thoth (the god of wisdom and learning), and so on. (7)

In an extended interview for National geographic, John Taylor, an Egyptian-antiquities expert at the British Museum in London, says that animal mummies represent a largely untapped resource for scientists and historians. Some animal cemeteries contain literally millions of examples of one particular species. "You can't image that all of them were given the most elaborate treatment," Taylor said.

However, Taylor agrees there is no doubt that some animals were preserved to an extremely high standard. "In the case of the Apis bull, for example, we know that mummification was a

very ritualized process that was taken extremely seriously. It was certainly no less elaborate than the mummification of an important human," he said. "But to what extent does that elaborate care for animals extend further down the scale?"

"Specimens from museum collections have often been selected because they are the finest specimens," Taylor added. "They don't necessarily represent a cross section of what was actually found."

Lizard, fish, and even beetle mummies from ancient Egypt have been unearthed. Pets were also mummified and buried in tombs with their owners. Such pets weren't limited to dogs and cats but included baboons, monkeys, and gazelles.

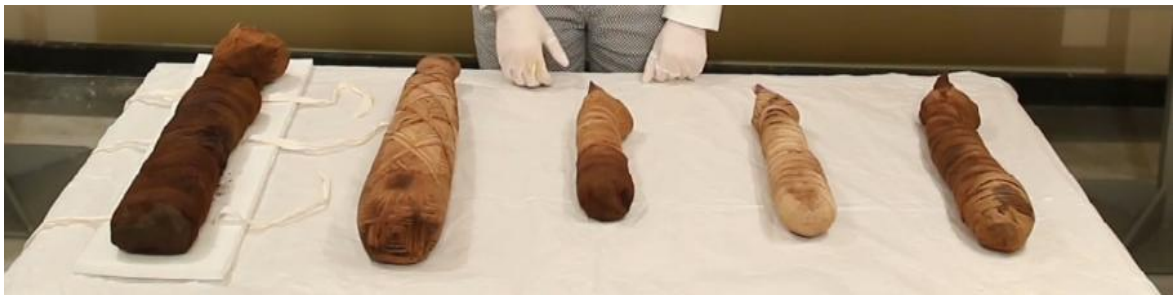
As with other animals, the precise embalming procedures involved in pet mummification remains largely unknown.

Evershed, the study coauthor, said, "The mummification process just isn't documented by the ancient Egyptians, which is why we are doing the chemical analysis. It was a secretive process."

He said further mummy studies should help to unravel other mysteries about ancient Egyptian society. For instance, knowledge of the precise embalming agents used should shed new light on the extent of trade between Egypt and its neighbors. "Coniferous [tree] resin was widely found around the Mediterranean, and Romans traded it very widely," Evershed said. Just like those famously bandaged Pharaohs, it seems the vast menagerie of animal mummies left behind by the ancient Egyptians have many more secrets wrapped within.

Today, to characterize animal embalming materials is still a challenge and this chapter aims to characterize the materials used in the mummification processes of 4 cat mummies (These mummies are among those reported in figure X), that are undergoing at restoration intervention at the Museo “Umberto Scerrato” in Napoli, under the supervision of Dr.ssa Maria Diletta Pubblico and Prof. Andrea Manzo, from University “L’Orientale” di Napoli,

Figure 1: Cat mummies



All the mummies show the same features. Their bodies were covered with linen bandages and their face was painted in black and red. Somatic characteristics were well reproduced. The ears were covered with linen bandages too, to shape a cone that was painted in red. "

8.3 Embalming procedure

The first real attempts of mummification started in the Archaic period; the earliest mummies were wrapped in many layers of linen bandaging and sometimes the external features of the faces were fashioned out of the linen to give the mummy a more 'lifelike' appearance. The next stage in the development towards mummification of the body involved the evisceration of the body to inhibit the process of decomposition; an incision was made in the abdomen, usually on the left side, and the abdominal organs (intestines followed by the liver and stomach) were extracted, except for the kidneys; next the diaphragm was cut out and the lungs removed. The heart, seat of the mind, was left in place. After the internal organs were removed, the thorax and abdomen were cleansed with palm wine and the body cavity was packed with temporary stuffing materials. The brain was also removed, but it was not kept since the Egyptians did not think it was important. The body was completely covered with dry natron, a naturally occurring mixture of salts containing sodium carbonate, sodium bicarbonate, sodium chloride and sodium sulphate in varying proportions. The drying-out process lasted 40 days. After complete dehydration of the body, the temporary packing material was removed and the body cavities were stuffed with linen cloth and bags of several materials including natron, resin-soaked bandages, sawdust and earth, in such a way as to give the body a more natural shape. The brain cavity was also filled with resinous materials or linen; the whole body was then coated in viscous material obtained by mixing a variety of natural substances. (4,8)

8.3.1 Embalming materials and chemical characterization

The materials used for embalming have always been a matter of controversy. (9-11) Natural resins, gums, waxes, oils and other natural substances (12,13), together with various embalming practices, have been utilized over the centuries, but very little historical information about them is available and very little analysis of the embalming materials has been carried out. The few chemical characterizations undertaken explain the black colour of mummies primarily by the presence of bitumen from the Dead Sea or from Iraq, plant resinous extracts and tannins(14-20). Although it is a challenge for a chemist to distinguish between plant resins, pitches, bitumen, oils, waxes and other natural compounds and their degradation products, which may all be present in a balm, some analytical methods are proposed in the literature, and are generally based on high performance liquid chromatography (HPLC) (19) and gas chromatography-mass spectrometry (GC-MS). (14-18, 20). However, most of these methods are directed to the identification of a few of the substances that are simultaneously present in the sample, and thus only partial information can be obtained. This Chapter describes an analytical procedure for the extraction and identification of the main natural organic compounds simultaneously present in a animal mummy balms. (4)

Millions of votive mummies of mammals, birds and reptiles were produced throughout ancient Egypt, with their popularity increasing during the reign of Amenhotep III (1400 BC) and there after. The scale of production has been taken to indicate that relatively little care and expense was involved in their preparation compared with human mummies. (21-23) The accepted view is that animals were merely wrapped in coarse linen bandages and/or dipped in ‘resin’ before death.(22-24) However, similarly to human mummification there was a range of qualities of treatments, and visual inspection of animal mummies suggests that the procedures used were often as complex as those used in humans (for example, evisceration and elaborate bandaging). Moreover, the ancient Egyptians treated animals with great respect, regarding them both as domestic pets and representatives of the gods; for example, the cat symbolized the goddess Bastet; the hawk, Horus; the ibis, Thoth, and so on. (24)

We report here the results of chemical investigations of tissues and wrappings from cats mummies using gas chromatography–mass spectrometry.

8.3.2 Bitumen in ancient Mummies

When the Arabs initially encountered mummified bodies, the consistency of the embalming materials and their dark colour led them to believe that the substance was bitumen known to them as mum. Hence, the Arabs called the bodies thus prepared ‘mumia’, after the black material they thought had been used to make them. Both the name for Egyptian preserved bodies (mumia, mummy, momie, mumie, mummia, etc.) and the idea that their preservation was due to bitumen were widely adopted, including by Egyptologists from the nineteenth century onwards (7,25-29). Critically, the modern term ‘bitumen’ refers to a specific naturally occurring petroleum product, also known as asphalt that has lost its volatile hydrocarbon components via biodegradation and/or evaporation, leaving a black, semi-viscous, or even solid, material. However, as bitumen was black, early Egyptologists and more recent researchers, fell into the habit of describing all black mummy balms as ‘bitumen’, when its ubiquity certainly should not be assumed when its presence has only been confirmed in a handful of individuals (25-27,28). Significantly, there is no consensus among the classical authors as to bitumen use in mummification. Diodorus Siculus, in the first century BC (30), and Pliny (31), in the first century AD, mention bitumen when writing about the Dead Sea. The former mentions that bitumen was sold to the Egyptians for embalming but it is not mentioned when he discusses mummification. In the fifth century BC, Herodotus (32) provides the most thorough descriptions of mummification, makes no mention of the use of bitumen during the process, but does describe its use in other contexts completely unrelated to preserving the dead. Strabo (33) reports the common sources of bitumen in the first century BC, and also refers to the substance when writing about the Dead Sea, and states: The Egyptians use the asphalt for embalming the dead. Until recently, it was thought that the trade route for the Egyptians to the Dead Sea was only available in Ptolemaic and later times (34), but archaeological discoveries and chemical analyses have revealed molecular evidence for trade during the earlier Chalcolithic and Early Bronze Age periods (3900–2200 BC; (18)). Clearly, therefore, bitumen was available to the Egyptians, but its role in mummification is poorly documented. The dark or black colour is an unsatisfactory diagnostic for bitumen as many organic materials when heated, aged and/or decayed naturally darken. Thus, it is possible that the dark colour of many of the balms may be ascribed to a variety of non-bituminous sources. (35) It was only in the early twentieth century that some chemists questioned the fact whether all black balms were indeed bitumen (27), and from that time on, with subsequent tests of mummification material it became clear that a variety of materials,

including petroleum bitumen, were employed in mummification throughout its over 3300 year history in Egypt. (16, 36)

8.4 Materials and Methods

8.4.1 Samples

The samples were a collection of 7 tissues and bandaging samples from 4 animal mummies. It was provided by professor A.Manzo, Università degli studi di Napoli L'Orientale.

8.4.2 Chemical analysis of organic molecules

Aliquots of each sample were extracted with chloroform-methanol solution (2:1 v/v) and then with hexane as reported in the chapter 7 and the extracts were analyzed by GC/MS.

In particular, the analyses were performed with an Agilent 6890 Series GC, coupled to a detector MSD quadrupole mass spectrometer 5973, equipped with a gas chromatograph as indicated for the analysis of non proteinaceous molecules in the previously chapters.

For SIM, the mass spectrometer was set to monitor m/z 191 (hopanes and other triterpanes) and 217 (steranes). (36)

8.5 Results and discussion

The study of animal embalming procedures were an important challenge to increase the knowledge of the uses and habits of ancient people like Egyptians that used mummification as funerary tradition.

To characterize embalming materials used in animal mummies and to define if these substances are comparable to those used to human ones, a chemical investigation of samples from the collection of animal mummy balms was carried out.

The investigation proceeded in two stages. Initially, all the balms were screened by GC-MS to determine the major components of the balms. Since, from the literature, (36) the biomarkers for petroleum bitumen are known to be generally present in trace, in a second stage, extracts were analysed by GC-MS in Selected Ion Monitoring (SIM) mode. The extracts were screened for the presence of diagnostic hopane (andothertriterpane) and sterane petroleum biomarkers, according to the procedure reported in (36).

The results obtained in the first step are summarized in Tables 8.1 to8.7.

Table8. 1: 66GU sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20,92	octadecane, 6-methyl
21,04	Z-8-methyl-9-tetradecenoic acid
21,26	octadecane, 2-methyl
22,57	tetradecane, 2, 6, 10-trimethyl
22,85	heptadecane, 2, 6, 10, 15-tetramethyl
23,33	10, 18-bis norabieta-8, 11, 13-triene
24	heptacosane
25,27	octadecane, 3-ethyl-5-(2-ethylbutyl)

Table 8.2: 66 black/red sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20,74	spiro-1-(ciclohex-2-ene)-2-(oxabicyclopentane), 1, 4, 2, 6, 6-pentamethyl
20,92	1-cyclopropene-1-pentanol, α, ε, ε, 2-tetramethyl-3-(1-methylethenyl)
21,03	1, 6-dioxacyclododecane-7, 12-dione
21,14	estra-1, 3, 5-trien-17β-ol
22,15	10, 13-octadecadiynoic acid, methylester
22,63	n-hexadecanoic acid
22,94	1-heptatriocotanol
23,21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23,33	10, 18-bis norabieta-8, 11, 13-triene
24,19	5, 8, 11, 14-eicosatetraynoic acid
26	5α, 8α-cholestane-3β, 6α-diol, 8, 9-epoxy-1, 4-methyl-acetate

Table 8. 3: 67 black sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
21,03	1, 6-dioxacyclododecane-7, 12-dione
22,14	10, 13-octadecadiynoic acid, methylester
22,63	1-heptatriocotanol
23,21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23,33	10, 18-bis norabieta-8, 11, 13-triene
23,78	acetic acid,17-(1-acetyloxy-ethyl)-10, 13-dimethyl-3-oxododecahydro…
24,84	ethyliso-allocholato
26	5α-cholest-8-en-3-one, 6α-hydroxy-14-methyl, acetate

Table 8.4: 67 ear sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20,81	spiro-1-(ciclohex-2-ene)-2-(oxabicyclopentane), 1, 4, 2, 6, 6-pentamethyl
20,91	9-octadecen-12-ynoic acid, methylester
21,03	1, 6-dioxacyclododecane-7, 12-dione
22,15	10, 13-octadecadiynoic acid, methylester
22,94	2-[4-methyl-6-(2, 6, 6-trimethylcyclohex-1-enyl)hexa-1, 3, 5-trienyl]cyclohex-1-en-1-carboxaldehyde
23,21	pregna-5,16-dien-20-one, 3-(acetyloxy)-16-methyl-(3β)
23,33	10, 18-bis norabieta-8, 11, 13-triene

Table 8.5: 67 GU sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
20,92	decane, 2, 3, 5, 8-tetramethyl
21,04	7-methyl-Z-tetradecen-1-olacetate
21,26	heptadecane, 2, 6, 10, 15-tetramethyl
23,33	10, 18-bis norabieta-8, 11, 13-triene
24	heptacosane
26	5α, 8α-cholestane-3β, 6α-diol, 8, 9-epoxy-1, 4-methyl-acetate

Table 8.6: 68sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
16,66	tri(1, 2-propyleneglycol), monomethylether
21,03	1, 6-dioxacyclododecane-7, 12-dione
22,87	1, 4-hexadien-3-one, 5-methyl-1-(2, 6, 6-trimethyl-2, 4-cyclohexadien-1-yl)
23,33	10, 18-bis norabieta-8, 11, 13-triene

Table 8.7: 69sample: retention time of the organic fraction of the samples obtained after chloroform-methanol extraction, followed by hexane extraction. The resulting molecules were analysed by GC/MS. Identification was done by comparison of the retention time and mass spectra to those of standards in the instrument manufacture database NIST MS Search 2.0.

Retention Time	Molecules
21,03	1, 6-dioxacyclododecane-7, 12-dione
21,37	phenol, 2-(1, 1-dimethylethyl)-4-(1, 1, 3, 3-tetramethylbutyl)
23,33	10, 18-bis norabieta-8, 11, 13-triene

The analyses reveal the presence of highly complex mixtures of hydrocarbons (66GU and 67GU), terpenes (66black/red and 67GU) and phytosterols (66black/red, 67GU and 67 black). These mixtures are comparable to those found in a natural petroleum hydrocarbon source. (36) In addition, antiseptic substances were found; among these ethyl isoallocolate (67 black). that has antiseptic and antioxidant properties.(37)

Coniferous resin was also present in all the samples analyzed. Derivative of abietic acid, a typical molecule of aged archaeological conifer resins (4,38,39-41) was identified. The use of coniferous resins is most notable since their terpenoids were known to inhibit microbial degradation via mechanisms (physico-chemical barriers and antimicrobial action) analogous to their protective roles in the plants from which they were derived. (42,43)

8.5.1 Screening mummy balms for bitumen biomarker

In the second stage, GC-MS in selected ion monitoring (SIM) operating mode was used to screen all the 7 samples for the presence of bitumen biomarkers, principally hopanes and steranes, considered as marker of the use of bitumen in balms. Literature data report that steranes concentrations were always lower than hopanes (and other triterpanes). (36)

In all the samples of mummy balms analysed, chromatograms show evidence for the presence of hopanes (retention time 21.81, 22.64, 23.35 min). It is worth mentioning, all the samples analysed did not contain sterane in detectable amount. However, several moecules with the same *m/z* values were obtained from the analysis. For this reason, further experiments, using derivatization protocols, will be necessary to confirm that in the samples there is truly hopane and not other compounds.

The mixtures we have characterized are of comparable complexity to those used to mummify humans (1,4,36). Moreover, the use of other protocols, like derivatization or hydrolysis, could

be provided additional data. Particularly, with these experiments, we could be identify other molecules like lipids, other di-tri terpenoids or beewax that were typically used in embalming materials.

8.6 Conclusions

Mummification was practised in ancient Egypt for more than 3000 years, emerging from initial observations of buried bodies preserved by natural desiccation. The use of organic balms (and other funerary practices) was a later introduction necessitated by more humid burial environments, especially tombs. The dark colour of many mummies led to the assumption that petroleum bitumen (or natural asphalt) was ubiquitous in mummification. (36)

Animal mummification was an enormous part of Egyptian culture. Animals were typically mummified for several purposes, especially religious. Egyptians treated animals with great respect, regarding them both as domestic pets and representatives of the gods.(7)

In this Chapter, we have described an analytical procedure for the extraction and identification of the natural organic compounds (not only bitumen) present in animal mummy balms with the aim to characterize embalming materials used in animal mummies and to define if these substances are comparable to the techniques used to human ones. We have tested 7 samples from 4 cat mummies provided by the museum of "L'Orientale" and we have showed that a variety of natural products possessing a range of preservative properties were applied during animal mummification.

Preliminary analyses show that the mixtures identified in the cats mummies, are comparable with those ones used to mummify humans. (1,4,36)

However, further experiments, based on hydrolyzation and derivatization processes, will be indispensable to identify other molecules like lipids, other di-tri terpenoids or bee waxes. Indeed, these molecules were tipically used during embalming processesand their identification can confirm that the mixtures used in these cats mummies are comparable with those ones used to mummify humans.

8.7 References

1. Bianucci R,. Habicht ME, Buckley S, Fletcher5 J, Seiler R, Öhrström LM, Vassilika E, Böni T,. Rühli FJ (2015) Shedding New Light on the 18th Dynasty Mummies of the Royal Architect Kha and His Spouse Merit. PLOS ONE 1-21.
2. Sauneor S (1952) Rituel de l'embaumement. Imprimerie Nationale, Le Caire.
3. Goyon JC (1972) Rituels fundraires de l'Ancienne Egypte. Le rituel de l'embaumement . Les editions du Cerf, Paris.
4. Colombini MP, Modugno F, Silvano F, Onor M (2000) Characterization of the Balm of an Egyptian Mummy from the Seventh Century B.C. Studies in Conservation (45) 19-29
5. Cockburn A, Cockburn E, Reyman TA. (1998) Mummies, Disease and Ancient Cultures. CambridgeUniversity Press.
6. Aufderheide AC. (2003) The Scientific Study of Mummies. Cambridge University Press.
7. Velde H (1980)"A Few Remarks upon the Religious Significance of Animals in Ancient Egypt." Numen (27) 76-82.
8. Ikram S, Dodson A (1998) The Mummy in Ancient Egypt. Thames and Hudson Ltd, London.
9. Lucas A (1962) Ancient Egyptian Materials and Industries. Edward Arnold (ed), London.
10. Rullkotter J, Nissenbaum A (1988) Dead Sea asphalt in Egyptian mummies: molecular evidence. Naturwissenschaften (75) 618-621.
11. Nissenbaum A (1992) Molecular archaeology: organic geochemistry of Egyptian mummies. Journal of Archeological Science (19) 1-6.
12. Zohary M (1982) Plants of the Bible. Cambridge University Press.
13. Mills JS, White R. (1987) The Organic Chemistry of Museum Objects. Butterworths, London.
14. Drenkhan R, Germer R. (1991) 'Identifierung von pflanzlichen Inhaltsstoffen einer Mumi- fizierungsmasse aus zeit um 600 v.Chr' in Mumie und Computer. Th. Schaefer Druckerei GmbH, Hannover .
15. Mejanelle P, Bleton J, Goursard S, Tchapla A (1997) 'Identification of phenolic acids and inositols in balms and tissues from an Egyptian mummy'. Journal of Chromatography 767.

16. Connan J, Dessort D (1991) 'Du bitume dans des baumes de momies Egyptiennes (1295 av.J.C-300 ap.J.C.): determination de son origine et evaluation de sa quantita'. Acadimie des Sciences de Paris (312) 1445-1452.
17. Connan J, Dessort D (1989) 'Du bitume de la mer Morte dans les baumes d'une momie Egyptienne: identification par criteres moleculaires'. Acadimie des Sciences de Paris (309) 1665-1689.
18. Connan J, Nissenbaum A, Dessort D (1992) 'Molecular archeology: export of Dead Sea asphalt to Canaan and Egypt in the Chalcolithic-Early Bronze Age (4th-3rd millennium BC)'. Geochimica et Cosmochimica Acta (56) 2743-2759.
19. Vieillescazes C, Coen S (1993) 'Caracterisation de quelques resines utilises en Egypte ancienne'. Studies in Conservation (38) 255-264.
20. Proefke ML, Rinehart KL, Mastura R, Ambrose SH, Wisseman S (1992) 'Probing the mysteries of ancient Egypt. Chemical analysis of a Roman period Egyptian mummy'. Analytical Chemistry (64 106) 111.
21. Lucas, A. (1989)Ancient Egyptian Materials and Industries. Histories and Mysteries of Man, London.
22. Malek, J. (1997)The Cat in Ancient Egypt. British Museum, London.
23. Brier B, (1996) Egyptian Mummies. Michael O' Mara Books, London.
24. Buckley SA, Clark KA, Evershed RP (2004) Complex organic chemical balms of Pharaonic animal mummies. Nature (431) 294-299.
25. Granville AB (1825) An essay on Egyptian mummies: with observations on the art ofembalming among the ancient Egyptians. Phil. Trans. R. Soc. (115) 269–316.
26. Budge EAW (1883) The mummy: a handbook of Egyptian funerary archaeology. New York, NY: Dover Publication Inc.
27. Lucas A. (1914) The question of the use of bitumen or pitch by the ancient Egyptians inmummification. J. Egypt. Archaeol. (1) 241–245.
28. Brier B (1996) Egyptian mummies. Brockhampton Press.
29. Aufderheide A (2003) The scientific study of mummies. Cambridge University Press.
30. Diodorus Siculus (1935) The library of history. Transl. Oldfather CH, London.
31. Pliny the Elder (1989) Natural history III-VII. Loeb Classical Library.
32. Herodotus (1996) The histories. Penguin.
33. Strabo (1969) Geography. Leob Classic Libraries.
34. Forbes RJ. (1955) Studies in ancient technology. The Netherlands: Brill, Leiden.

35. Serpico M, White R (2000) Resin, pitch and bitumen. In Ancient Egyptian materials and technologies, Cambridge University Press.
36. Clark KA, Ikram S, Evershed RP (2016) The significance of petroleum bitumen in ancient Egyptian mummies. The royal society.
37. Ahmad I, Aqil F (2009) New strategies combating bacterial infection. Wiley-Blackwell.
38. Barakat AO. (2005) Organic geochemistry indicates Gebel El Zeit, Gulf of Suez, is a source of bitumen used in some Egyptian mummies. Geoarchaeology (20) 211–228.
39. Proefke ML, Rinehart KL, Raheel M, Ambrose SH, Wisseman SU (1992) Probing the mysteries of ancient Egypt: chemical analysis of a Roman Period Egyptian mummy. Anal Chem (64) 105–111.
40. Koller J, Baumer U, Kaup Y, Etspu¨ ler H, Weser U (1998) Embalming was usedin Old Kingdom. Nature (391) 343–344.
41. Pollard AM, Heron C (1996) Archaeological chemistry. Royal Society of Chemistry, Cambridge
42. Briggs DEG, Eglinton G (1994) Chemical traces of ancient life. Chemistry in Britain. (30): 907–912.
43. Jones J, Higham TFG, Oldfield R, O'Connor TP, Buckley SA (2014) Evidence for Prehistoric Origins of Egyptian Mummification in Late Neolithic Burials. PLoS ONE (9) 1-13.

9.Dedicated protein-based binders database applied to targeted MS method in Cultural Heritage

Proteomics methods rely on mass spectrometry technologies and can be broadly divided in two main categories: discovery and targeted proteomics. The most common MS-based proteomics pipelines aimed at identifying proteins, follow the classical bottom up approach, where protein samples are enzimatically digested to peptides, which are then separated by liquid chromatography and eventually analyzed by tandem mass spectrometry. In discovery proteomics (or shotgun proteomics, global profiling proteomics), peptides and thereof proteins are identified by correlating experimental spectra to virtual ones in a protein sequence database. So far, all the applications to the cultural heritage field are discovery proteomics based experiments. (1,2,3)

The potential of this operating mode lies in the possibility of uncovering novel, or at least unexpected proteins, in what can be called a “blind” analysis, without any “a priori” knowledge or hypothesis this approach is therefore compatible with both well and poorly characterized systems in cultural heritage, the only requirements being the actual presence of proteins in the sample and the presence of the protein sequences in a properly selected database.

Despite the tremendous success of these discovery, standard based proteomics approaches so far in the field of cultural heritage objects, (4–14) the complexity of these unusual samples, in terms of chemical composition, presence of contaminating protein-based materials, as well as the high levels of molecular damage found in ancient samples, can often overwhelm even the most modern mass spectrometer. Moreover, a “blind” analysis is not always absolutely required in this kind of samples, where some hypothesys can in principle be worked out and the results are expected in a limited range of possibilities. This is for instance the case of proteinaceous binders in paintings, where generally the choices are among milk, egg or animal glue based binders. In such cases, the basic analytical challenge is the unequivocal discrimination, with the highest sensitivity, among a discrete number of possibilities, overcoming the unavoidable problems strictly connected to the peculiar nature of the samples: the unusual physical state, the complex chemical composition (i.e. proteins are only

minor, although essential, component in the actual preparation of the color, which is full of inorganic pigments), the unavoidable presence of contaminating proteins coming from the environmental dust, and the chemical modifications induced by aging processes.

In such cases, an alternative mass spectrometric strategy, that selectively and nonredundantly detects the ions of one or, more desirably, a few peptides that unequivocally identify a protein among a restricted range of possibility would thus represent a significant advance in the solution of the analytical challenge.

As an alternative to a discovery, data-dependent operation mode of the mass spectrometer, it is possible to operate in a targeted mode, by programming the instrument to continuously monitor selected ions that upon fragmentation generate specific species, and these ions transitions can be diagnostic of the presence of a specific peptide, and, thereof, can constitute a signature of the presence of a protein. Targeted proteomics is recognized as the most sensitive and selective way to detect pre-selected components in a complex matrix. Adoption of targeted mass spectrometry approaches such as Multiple Reaction Monitoring (MRM) is well underway in the proteomics community, and is gradually complementing large shotgun studies with hypothesis-driven experiments, (15–19) prompted by the growing demand for reliable quantification methods. In this context, the targeted MS methods (Multiple Reaction Monitoring, MRM, and Selected Reaction Monitoring, SRM, are terms that can be used interchangeably) can be developed very rapidly, although developing a robust and precise assay can be quite demanding, the main challenge being the proper selection of the target peptides that would ensure reliable, high quality data. There is no report so far of any application of targeted proteomics in the field of cultural heritage, that would be desirable to increase sensitivity of protein identification in highly complex and rather contaminated samples, beside paving the way to the fascinating perspective of quantitate proteins and/or, even more attractive, modifications on proteins in artworks.

In this chapter we report, to the best of our knowledge, the first example of application of targeted proteomics to samples in the field of cultural heritage. Proteinaceous binders in painting are mainly derived from egg, milk casein and animal glue, and constitute a perfect benchmark for testing the application of targeted proteomics to cultural heritage samples. A targeted proteomics approach could rapidly answer with a sensitivity comparable to immune-based strategies to the analytical question of identifying the type of binders in paintings, a critical aspect for art conservationists and restorers that can plan a proper restoration

campaign on the basis of a scientific diagnosis. We therefore developed an MRM assay to detect and distinguish proteins in painting binders and tested it on samples that had already produced positive identifications and, later, on samples for which no protein identification was achieved in discovery proteomic analyses.

Selection of the proteotypic peptides, a critical step in the assay set up, was preliminarily carried out on an experimental basis, from the statistical analysis of the most recurrent peptides in a large dataset of samples that had been analyzed by classical discovery strategies.

Results herein presented demonstrate the improvement in selectivity and sensitivity achieved with the targeted approach in respect to classical profiling experiments and pave the way to possible application of this most advanced analytical strategy in the field of archaeology and artworks.

9.1 Materials and Methods

9.1.1 Reagents.

Ammonium hydrogen carbonate (AMBIC), Ethylenediaminetetraacetic acid (EDTA); Tri(hydroxymethyl)aminomethane (TRIS), TPCK-treated trypsin were from Sigma; recombinant Peptide N-Glycosidase F (PNGaseF) was from Roche. Formic acid and Acetonitrile (ACN) were purchased from Baker, respectively. Deionised water was obtained from Millipore cartridge equipment.

9.1.2 Painting samples

Replicas and historical samples analysed and used for database building consist of pigments and protein-based binders as reported in Table B1. (Appendix B)

Aliquots of each sample were treated in heterogeneous phase with trypsin with the minimally invasive protocol either without any sample pre-treatment or with a particular pre-treatment as indicated in the chapter 4 or with a deglycosylating step. (20) A deglycosylation pre-treatment step was performed as described by Vinciguerra *et al* (21). Briefly, 50 µL of AMBIC 50 mM containing 60mU/µL of PNGaseF solution was added to microsamples (ca 600-800 µg), and incubated at 37°C for two hours. The reaction was stopped by boiling the sample for 2 minutes. Subsequently, enzymatic digestion was carried out with the urea 6 M pretreatment as described in the chapter. 4. (22)

The resulting peptide mixtures were analysed by LC-MS/MS and/or MALDI-TOF. (10,21,22,23,24)

In particular, the peptide mixture (1µl) was analyzed by nanoLC-MS/MS, on a CHIP MS 6520 QTOF equipped with a capillary 1200 HPLC system and a chip cube (Agilent Technologies, Palo Alto, CA) as indicated for proteomics analysis in the chapter 4. For some samples, however, the peptide mixtures were analyzed by Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) on a 4800 Plus MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA) as indicated in the chapter 4.

9.2 MS database building and statistical analysis

A collection of 116 mass spectrometric analyses, carried out on 107 independent samples, were screened to identify the proteins most frequently detected in standard global profiling experiments. The collected samples were grouped on the basis of the binders.

The statistical analysis was performed by screening the samples proteins and peptides most frequently detected in standard global profiling experiments in a large dataset of samples. A relational database was designed on PivotTables (https://support.office.com) to simplify the data processing. Moreover, data evaluation was carried out regardless the treatment procedures they have undergone and the mass spectrometers that was used in the specific analysis.

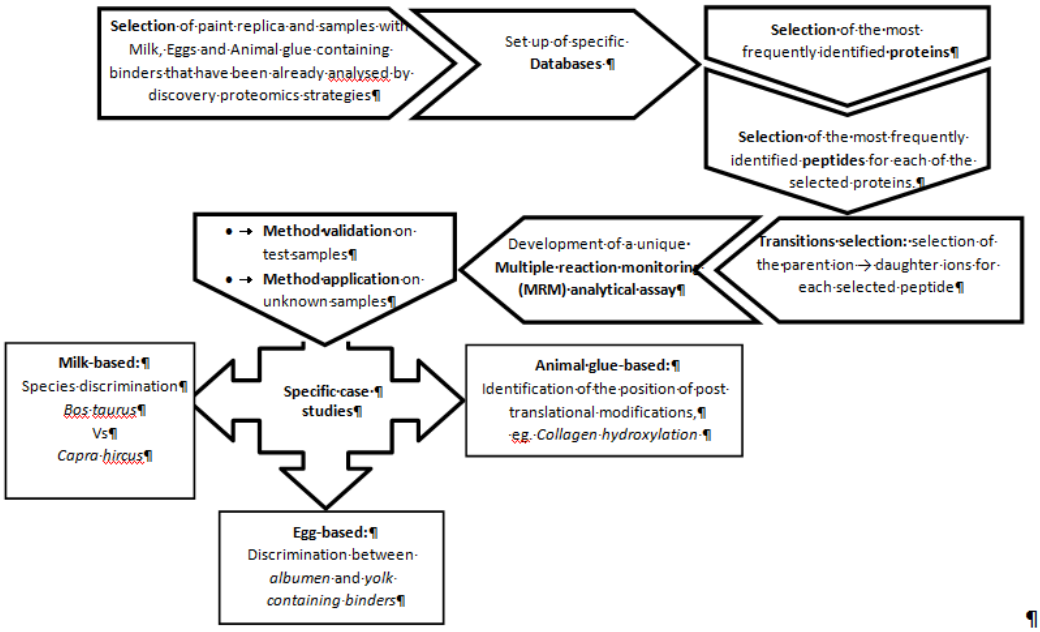
9.3 MRM assay design

The MRM assay requires the setup of specific mass spectrometry acquisition parameters that was carried out by Prof. Angela Amoresano and Dr. Anna Illiano, Dep. of Scienze Chimiche, Università "Federico II", Napoli.

Skyline Software drew the choice of specific transitions and parameters related to ionization and fragmentation for each analyte. Peptide mixtures were analyzed by LC-MS/MS analysis using a Xevo TQ-S (Waters) with an UPLC Microflow Source coupled to an UPLC Acquity System (Waters) using an IonKey (Advion) device. For each run, 1 µL peptide mixture was injected and separated on a TS3 1.0 mm × 150 mm analytical RP column (Waters, Milford, MA, USA) at 60 ◦C with flow rate of 3 µl/min. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B (0.1% Formic acid in 100% ACN) in A (0.1% Formic acid in water (LC–MS grade)) from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min with eluent A. MRM mass spectrometry analyses were performed in positive ion mode. Skyline Software drew the choice for the selection of virtual best transitions and collision energy calculated to generate maximal fragmention intensities for each analyte. The developed MRM method was used with MRM detection

window set to 0.5–1.6 min per peptide, the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35V.

Figure9. 1: Workflow of the development and the assessment of the MRM assay for protein-based binder identification.



9.4 Results and discussion

The basic scheme of the development of a targeted proteomic assay starts with the selection of a target list of proteins, on the basis of previous experiments, literature or knowledge. Then, target peptides that optimally represent the target proteins are selected, and this step is followed by the selection of a set of suitable transitions for each selected peptide. In our assay development, schematically described in figure 9.1 we opted for an empirical selection of proteins and then of peptides based on our previous experience with samples from cultural heritage, and the designed assay was then validated on samples that had already been characterized in profiling experiments and tested on unknown ones. Moreover, we validated the use of an MRM assay for some specific requests that could be relevant in specific analyses of proteinaceous binders.

9.4.1 Database construction and proteins selection

Selection of diagnostic peptides that can constitute the signature of proteins and, in our case, thereof of a material, is the critical step in the experimental design of the assay because it will affect both the specificity and the sensitivity of the targeted analysis. Predictive models based on experimental data (25) have proven to be extremely powerful in the selection of

proteotypic peptides in targeted proteomics, and several software packages exist that facilitate this step of the assay development (26). However, among the factors that could affect the effective detection of a peptide, for sure we need to consider the physical state of the sample. The peculiar, solid and aged nature of the samples we aim to develop the assay for, prompted us to approach the selection in a definitively empirical mode. We therefore screened the results obtained on the collection of painting samples that we have been analysing throughout years in our lab, including both paint replicas and historical samples, regardless the treatment procedures they have undergone and the mass spectrometers that have been used in the specific analysis, in the perspective of picking the most experimental procedure independent markers that could be used in any laboratory and in the aim of minimizing the influence of sample pretreatment procedures in the selection of the most representative and discriminative peptides.

Figure B1 reports the summary of the results of the screening of the samples that contained animal glue, data clearly indicates that COL1α1 (collagen alpha-1(I)) from *Bos taurus* (P02453) was detected in 29 samples out of 29 samples and COL1α2 (collagen alpha-2(I)) (P02465) in 28 samples. When milk was used (B2) alpha S1 casein from *Bos taurus* (P02662) was identified (43 samples), as could be expected considering its relative abundance. Moreover, beta casein (P02666) (77%) and alpha S2 casein (P02663) (72%) could be detected in 33 samples and in 31 samples out of 43 total milk containing samples, respectively.

The case of egg containing samples is more complex, since we have to distinguish samples where only albumen was used from samples with only yolk and samples containing both yolk and albumen. It's worth to underline that while is technically quite easy to have "pure" albumen from an egg, the yolk is likely contaminated by the egg white. Therefore, in the sake of simplicity, on the basis of experiments carried out on eggs, we grouped the results obtained with the different egg containing samples, (43) as reported in Figure B3, according to the presence of albumen only (27 samples), and albumen plus yolk (16 samples), without considering yolk only containing samples separately, for which contamination from albumen proteins is very likely. As expected chicken ovalbumin (P01012), the main protein found in egg white, making up 60-65% of the total protein, is identified in almost all the samples: 38 out of 43 total samples and 27 out of 27 albumen only samples, accounting for 90% and 100%, respectively. Similarly, ovotransferrin (P02789) is frequently identified (29 out of 43 total samples, 67%, that increases up to 78% if we consider the albumen only samples, since

it was identified in 21 out of 27 albumen only samples). Vitellogenin-2 (P02845) and apovitellenin-1 (P02659) are the most frequently identified proteins from yolk (both of them have been identified in 14 samples out of 16 yolk and albumen containing ones, counting for 87%). Moreover, Vitellogenin-1 (P87498) can be identified in 12 out of 16 samples (75%). It is worth noting that neither vitellogenin-2 nor apovitellenin-1 have been detected in any albumen only containing sample.

From the results presented, we selected collagen alpha-1(I) and collagen alpha-2(I) as animal glue representative proteins, alpha S1 casein, alpha S2 casein and beta casein as the representative proteins for milk based binder, and ovalbumin and ovotranferrin as generic representative of egg containing binder, and, more specifically, presence of vitellogenin-2, apovitellenin-1 and vitellogenin-1 can be used as indicator of yolk presence.

9.4.2 Selection of the proteotypic peptides

Once we have identified the target proteins as the most frequently detected in the different binders, we deepened our analysis to identify the peptides to be used in the MRM assay. Our strategy of selection was essentially similar to that carried out in the selection of the proteins, i.e., we listed the peptides of the selected proteins that allowed identification, and screened the list for the most frequently observed ones.

The frequency of observation was calculated from the detailed analysis reported in table B2 for collagen-based samples. The selection of the peptides was carried out with the same rules also for eggs and milk-containing samples, which is part of the research project of Dr. R. Vinciguerra. We opted to select 3/4 target peptides per target protein.

It is worth mentioning that in most cases, we carried out both the MALDI-TOF and the LC-MS/MS analyses on the same peptide mixtures, in order to evaluate the independence of the observation from the specific mass spectrometric analysis. However, frequency of observation has been calculated considering the detection of the specific peptide in any of the analyses only once per sample. Moreover, in the perspective of selecting specific ions to develop an appropriate MRM assay, modified peptides were considered separately.

9.4.3 Animal glue proteotypic peptides

In the Table 9.1 the peptides that have been selected from the analysis of the animal glue containing samples for collagen alpha-1(I) and alpha-2(I) are reported.

Table 9.1: Selected peptides**.** Peptides selected upon the analysis of the occurrence in animal glue containing samples. The times the peptide has been observed and the corresponding percentage is reported. Details of the occurrence for all the peptide and their related graphical representations are reported in Tables B2 and Figures B4 and B5. Position in the protein sequence is given as flanking apexes.

Sample group (number of samples)	**Protein**	**Peptide**	**MH^+**	**N° of observations**	**%**
Animal glue containing samples (29)	**collagen alpha-1(I)**	^{1062}SGDRGETGPAGPAGPIGPVGAR^{1083}	1975.99	24	82
		^{1084}GPAGPQGPR^{1092}	836.44	24	82
		^{958}GVVGLPGQR^{966} + Hydroxy (P)	898.51	23	79
	collagen alpha-2(I)	^{326}GIPGPVGAAGATGAR^{340} + Hydroxy (P)	1267.67	22	75
		^{572}GIPGEFGLPGPAGAR^{586} + 2 Hydroxy (P)	1427.73	21	72
		^{1066}IGQPGAVGPAGIR^{1078}	1192.68	19	65

The detailed analysis is shown by the related graphical representations that are reported in Figure B4 and B5 for collagen alpha-1(I) and alpha-2(I), respectively. In the selection of collagenic reporter peptides, it has also been taken into account the occurrence of hydroxylation and the possible isomers for the relative position in the sequence of this modification, that would not alter the mass of the peptide, but rather the ions formed upon fragmentation and therefore the transition to be selected. For instance, in the case of collagen alpha-2(I), we selected peptide ^{326}GIPGPVGAAGATGAR^{340}, with hydroxylation on proline in position 328, that was detected in 22 samples (75%), as later confirmed also in the MRM analysis.

9.4.4 Transition selection

Once selected, target peptides for each target protein were analysed by using Skyline software in order to determine the virtual best transitions and collision energy calculated to generate maximal fragmention intensities. The transitions generated by Skyline were then compared with fragmentation spectra obtained by LC-MS/MS experimental data, when available. Peptides chosen by in silico analysis and showing the best signal to noise in the experimental fragmentation spectra were used for MRM method setting up. The full list of the selected peptides is reported in Table B3 together with the corresponding transitions and mass spectral parameters that have been selected for the MRM assays.

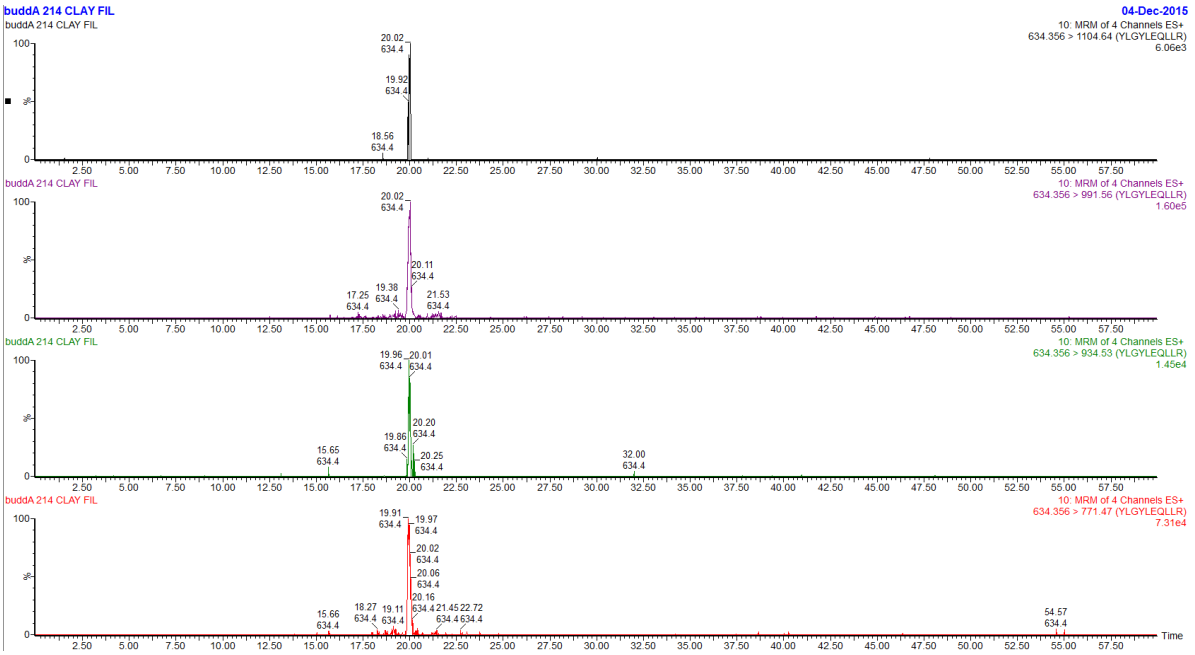
9.4.5 MRM Analysis

The development of the MRM assay required the setup of acquisition parameters that was carried out by Prof. Angela Amoresano and Dr. Anna Illiano. As considering LC-MRM/MS analyses a unique MRM method was built. This method contains all precursor ion-daughter ions transitions associated with target peptides. As a preliminary test of the developed assay, three selected samples, containing either milk, animal glue or egg as single proteinaceous component, were analysed by using the set up MRM unique method

9.4.6 MRM Analysis

As an example, the TIC chromatograms for the four transitions selected for peptide ^{106}YLGYLEQLLR^{115} of alpha S1 casein from *Bos taurus* are shown in figure 9.2. All the signals have been recorded at the same retention time, thus confirming that the transitions belong to the same peptide. As expected, all the selected proteins were identified with at least one peptide, each with at least three transitions. In no case any false positive were observed.

Figure 9.2: MRM traces chromatogram of four transitions (644.36>771.47; 644.36>943.53; 644.36>991.56; 644.36>1104.64) related to YLGYLEQLLR peptide of alpha s1 casein that co-eluted at 20 min. The analysis was carried out in one of the 14 Buddha's fragments for which no protein identification was achieved in discovery proteomic analyses.



Similar analyses were carried out on mixed test samples, and the results confirmed the validity of the method.

As next step we analyzed historical samples that have been already analysed by standard LC-MSMS.

The detection of several transitions that are ascribable to peptides from collagens and caseins were detected in a sample from the paintings from the Monumental Cemetry of Pisa, thus confirming the data acquired by standard proteomic approach. (23)

9.5 Specific applications

The same MRM procedure as described above can also be used to answer to specific requests in addition to the identification of the kind of binder.

For instance, a specific request in case of milk binders can be the discrimination of milk originating from different animal sources. Among the peptides that we selected for the identification of milk, some can be used to distinguish the animal origin at different levels. In most cases, the peptides will not be unique to a single animal source.

While in discovery proteomics identification of more than one peptides of the single proteins helps a lot in assigning the matches to a specie by the so-called principal of parsimony, in MRM operation mode peptides, and thereof transitions, should be selected properly in the assay development, if we want to answer questions like that. The principle of parsimony and the use of sequence alignment tools allowed us to assess that in the decoration of the giant Buddha statua from the Baymian valley in Afghanistan a mixture of cattle and goat milks was used as binder. Detection of peptides that were unique for the single sequences allowed to confidently establish the co-presence of the alpha S1 and alpha S2 caseins from *Bos taurus* and *Capra hircus*. In a development of the MRM assay to be specifically used to identify the origin of the milk, we selected FFVAPFPEVFGK and FVVAPFPEVFR peptides of alpha S1 casein from *Bos taurus* and *Capra hircus*, respectively, and the some four specific transitions, and tested the assay on the sample of Buddha's decoration that was negative in standard discovery analysis and positive for milk detection in the MRM assay as described above. Identification of the pre-selected four perfectly co-eluted transitions for both of peptides, clearly demonstrates the presence of *Bos taurus* and *Capra hircus* milk in mixture.

Although still in the process of optimization, the MRM method will also be used to obtain various supporting information in the case of eggs and animal glue. In the case of animal glue-based binders, it could be used in the identification of post translational modifications e.g. selective manner to identify the hydroxylation sites, very useful to evaluate the conservation state of the cultural heritage objects; and in the case of egg-based binders could

be used to discriminate between albumen and yolk eventually used in different ancient painting techniques.

This demonstrates that MRM analysis can also be used for answering to questions more specific than simply identificative purposes.

Moreover, the improvement in selectivity and sensitivity achieved with the targeted approach in respect to classical profiling experiments could pave the way to possible application of this most advanced analytical strategy in the field of archaeology and artworks.

9.6 Conclusions

The field of diagnosis in cultural heritage is continuously in search for more and more sensible methods that could reliably answer the important questions of the materials used in a work of art while preserving as much as possible the artistic object itself. This means that while completely non destructive techniques will always be the first choice approach, the development and the improvement of more and more sensible techniques, that can provide more and more details in a single analysis, will reduce the impact of invasive methods.

Herein, we first applied, to the best of our knowledge, a targeted proteomics approach to the identification of the proteinaceous binders in paintings. This specific issue has been successfully approached by standard proteomics, i.e. MALDI-TOF and LC-MSMS untargeted approaches, in recent years (see references therein for the state of the art). In this chapter we have demonstrated that a targeted proteomic approach can reliably improve the sensibility of the analysis, providing positive results in some cases were standard LC-MSMS analysis had failed, or when it provided partial results, such as in the case of the paintings decorating the walls of the Monumental Cemetery in Pisa. In that specific case, the overabundant milk and animal glue that were used in restorations and in the "strappo" procedure hid the presence of egg as additional and possibly original binder in the standard LC-MSMS analysis, while egg peptides were unambiguosly revealed in the targeted search.

Should MRM analysis overtake the classical approach? Not definitively, since the "limit" of targeted approaches is that the targets should be hypothesized in order to set up the specific assay: the study case of the Giant Buddha samples is the perfect example to illustrate this point of view. By analyzing the sample directly by MRM, we would have definitively and reliably assessed that milk was used in the decoration of the Statua, but, possibly, we would have not suspected the presence of a mixture of bovine and goat milk, since the peptides that have been chosen are either non specie specific or are specific for bovine milk. In that specific case, our satisfactory conclusion after the MRM analysis would have been that

bovine milk was used, while the close inspection of classical proteomics analysis results clearly identified the compresence of goat milk proteins. However, once the suspect was instilled, as we herein demonstrate, a specifically tailored MRM approach unambiguously assessed, with its unsurpassed sensitivity, that both milks were present. This result, obviously, paves the way to the exploitation of the MRM assay in the challenging field of species identification, that is made extremely challenging by the paucity of material and even more by the degradation state.

9.7 References

1. en Have ST, Ahmad Y, Lamond AI (2011) Proteomics — current novelties and future directions. J. Anal. Bioanal. Tech. (1) 1–6
2. Ostrom PH, Schall M, Gandhi H, Shen TL, Hauschka PV,. Strahler JR, et al. (2000) New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. Geochim. Cosmochim. Acta (64) 1043–1050.
3. Dallongeville S, Garnier N, Rolando C,Tokarski C (2015) Proteins in Art, Archaeology, and Paleontology: From Detection to Identification. Chem. Rev. (116) 2-79.
4. Tokarski C, Martin E, Rolando C (2006) Post-traductionnelles, M., and Cedex, A. Identification of Proteins in Renaissance Paintings by Proteomics. Anal. Chem. (78) 1494–1502.
5. Dallongeville S, Richter M, Schafer S, Kuhlenthal M, Garnier N, Rolando C, Tokarski C (2013) Proteomics applied to the authentication of fish glue: application to a 17th century artwork sample. Analyst (138) 5357–5364.
6. Dallongeville S, Garnier N, Casasola DB, Bonifay M, Rolando C, Tokarski C (2011) Dealing with the identification of protein species in ancient amphorae. Anal. Bioanal. Chem. (399) 3053–3063.
7. Solazzo C, Fitzhugh WW, Rolando C, Tokarski C (2008) Identification of protein remains in archaeological potsherds by proteomics. Anal Chem(80)4590-4597.
8. Dallongeville S, Koperska M, Garnier N, Reille-Taillefert G, Rolando C,Tokarski C (2011) Identification of animal glue species in artworks using proteomics: Application to a 18th century gilt sample. Anal. Chem. (83) 9431–9437.
9. Chambery A, Di Maro A, Sanges C, Severino V, Tarantino M, Lamberti A, Parente A, Arcari P (2009) Improved procedure for protein binder analysis in mural painting by LC-ESI/Q-q-TOF mass spectrometry: detection of different milk species by casein proteotypic peptides. Anal. Bioanal. Chem. (395) 2281–91.
10. Leo G, Cartechini L, Pucci P, Sgamellotti A, Marino G, Birolo L (2009) Proteomic strategies for the identification of proteinaceous binders in paintings. Anal. Bioanal. Chem. (395) 2269–2280.
11. Fremout W, Dhaenens M, Saverwyns S, Sanyova J, Vandenabeele P, Deforc, D, Moens L (2012) Development of a dedicated peptide tandem mass spectral library for

conservation science. Anal. Chim. Acta (728) 39–48.

12. Fremout W, Dhaenens M, Saverwyns S, Sanyova J, Vandenabeele P, Deforce D, and Moens L (2010) Tryptic peptide analysis of protein binders in works of art by liquid chromatography–tandem mass spectrometry. Anal. Chim. Acta (658) 156–162.
13. Rao H, Yang Y, Abuduresule I, Li W, Hu X, Wang C (2015) Proteomic identification of adhesive on a bone sculpture-inlaid wooden artifact from the Xiaohe Cemetery, Xinjiang, China. J. Archaeol. Sci. (53) 148–155.
14. Villa P, Pollarolo L, Degano I, Birolo L, Pasero M, Biagioni C, et al. (2015) A milk and ochre paint mixture used 49,000 years ago at sibudu, South Africa. PLoS One (10)
15. Picotti P, Bodenmiller B, Mueller L N, Domon B, Aebersold R (2009) Full Dynamic Range Proteome Analysis of S. cerevisiae by Targeted Proteomics. Cell (138) 795–806.
16. Pan S, Aebersold R, Chen,R, Rush J, Goodlett DR, McIntosh MW, Zhang J, Brentnall TA (2009) Mass spectrometry based targeted protein quantification: methods and applications. J Proteome Res (8) 787–797.
17. 27. Kiyonami R, Schoen A, Prakash A, Peterman S, Zabrouskov V, Picotti P, Aebersold R, Huhmer A, Domon B (2011) Increased Selectivity, Analytical Precision, and Throughput in Targeted Proteomics. Mol. Cell. Proteomics (10) 110.002931–M110.002931
18. Carr SA, Abbatiello SE, Ackermann BL, Borchers C, Domon B, Deutsch E W, et al. (2014) Targeted Peptide Measurements in Biology and Medicine: Best Practices for Mass Spectrometry-based Assay Development Using a Fit-for-Purpose Approach. Mol. Cell. Proteomics (13) 907–917.
19. Picotti P, Aebersold R (2012) Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods (9) 555–566
20. Bailey UM, SchulzBL (2013).Deglycosylation systematically improves N-glycoprotein identification in liquid chromatography–tandem mass spectrometry proteomics for analysis of cell wall stress responses in Saccharomyces cerevisiae lacking Alg3p. J. Chromatogr. (923)16–21 .
21. Vinciguerra R, Galano E, Vallone F, Greco G, Vergara A, Bonaduce I, et al. (2015) Degly-cosylation step to improve the identification of egg proteins in art samples. Anal. Chem.

22. Vinciguerra R, De Chiaro A, Pucci P, Marino G, Birolo L (2015) Proteomic strategies for cultural heritage: form bones to paintings. Microchem. J. (126) 341–348.
23. Leo G, Bonaduce I, Andreotti A, Marino G, Pucci P, Colombini MP, et al. (2011) Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. Anal. Chem. (83) 2056–2064
24. Bonaduce I, Colombini MP. Diring S (2006) Identification of garlic in old gildings by gas chromatography-mass spectrometry. J. Chromatogr. (1107) 226–232.
25. .MallickP, Schirle M, Chen SS, Flory MR, Lee H, MartinD, Ranish J, Raught B, Schmitt R, Werner T, Kuster B, Aebersold R (2007) Computational prediction of proteotypic peptides for quantitative proteomics. Nat. Biotechnol. (25) 125–131.
26. Colangelo CM, Chung, L, Bruce C, Cheung KH (2013) Review of software tools for design and analysis of large scale MRM proteomic datasets. Methods (61) 287–298
27. Pereira PC et al. (2014)Milk nutritional composition and its role in human health. NUT (30) 619–627.

Appendix A

Table A1: Deamidation sites of *human* Collagen alpha-1(I).In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from Herculaneum X indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

ErcolanoCollagene alpha 1 (I)																					
Position AA	CC104	FPC106	CC107	FMC107	CC108	FMC108	CC109	FMC109	FMC110	CC111	FMC111	CC112	FMC112	CC113	FPC113	CC114	FPC114	CC 115	FMC115	CC116	FMC116
162 Q																					
199 Q				X	X	X	X	X		X-D	X	X	X-D	X-D	X	X	X-D		X		X
202 Q				X	X-D	X-D	X	X-D		X-D	X	X-D	X-D	X-D	X-D	X-D	X-D		X-D		X
229 N																					
250 Q																					
295 N			D	X	X	X	X-D			X-D	X-D	X-D	X-D		X-D	D		X			
300 Q			X	X	X	X	X			X-D	X	X-D	X		X	X		X			
324 N			X-D	X-D	X-D	X-D	X-D	X	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D		X	X	X
358 Q				X-D	X	X-D	X-D	X	X		X	X	X-D	X-D		X	X		X-D		
367 Q				X		X	X	X			X	X-D	X				X				
387 N			X	X		X	X		X	X-D	X	X-D	X-D		X	X-D	X-D				
393 Q			X	X		X	X		X-D	X-D	X	X-D	X-D		X	X-D	X-D				
400 N		D	X	X	X	X-D	X-D	D	X-D	X-D	X-D	X-D	X-D	D	X	X					
421 Q		X-D							X	X-D	X-D		X-D		D			X			
432 N												X	X-D			X					
457 Q	D		X-D		X-D	X	X	X		X-D	X	X-D	X-D		X	X					
561 Q					X				X	X		X-D	X			X					X
576 Q	D	D			X					X		X-D	X			X				X-D	X
619 Q			X	X	X-D	X	X	X	D	X-D	X	X-D	X	D	X-D	X					
634 Q			X		X			X		X		X		X			X				X
643 Q			X		X			X		X		X		X			X				X
661 Q			X		X			X		X-D		X		X			D				X
688 Q	X	D							X	X					D			X			
700 N				X-D	X-D	X-D	X-D			X	X-D	X-D	X-D				X-D				
705 N				X-D	D	X	X-D			X	X-D	X-D	X-D				X-D				
721 Q			X	X-D	X-D	X	X	X-D	X	X	X-D	X-D	X-D	X	X	X	X-D				
727 Q			X-D	X-D	X-D	D	X	X-D	X	X	X	X-D	X-D	D	X	X	X				
822 Q			X-D	X	X-D	X	X	X-D		X-D	X	X-D	X-D	X-D		X	X-D				

855N			X-D	X	X-D	X	X-D	X-D		X-D	X	X-D	X-D	X-D		X	X-D		X-D		
891N			X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D		X-D	X	X	X-D			
952 Q		X	X	X-D	X-D	X-D	X-D	X-D	X	X-D	X	X-D	X-D	X-D	X	X-D	X-D		X-D		X
957 Q		X	X-D	X-D	X-D	X-D	X-D	X-D	X	X-D	X	X-D	X-D	X-D	X-D	X	X-D		X-D		X
966 Q	X			X-D	X	X	X		X	X				X	X-D			X			
985 Q				X	X	X	X	X		D	X	D	X		X		X		X	X	
1090 Q		D	X		D						D										
1102 Q																					
1117 Q			X	X	X-D	X	X	X	X	X-D	X	X-D	X-D	X-D	X	X	X		X		X
1129 Q			X-D	X-D	X-D	X-D	X-D	X-D	X	X-D	X	X-D	X-D	X-D	X-D	X	X-D		X-D		D
1156 N		D	X	X-D	X-D	X	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X	X		X	X	X
1203 Q												X		D						D	
1206 Q												X		X-D						D	

Table A2: Deamidation sites of human Collagen alpha-2(I). In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from Herculaneum X indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

ErcolanoCollagene alpha 2 (I)																					
Position AA	CC104	FPC106	CC107	FMC107	CC108	FMC108	CC109	FMC109	FMC110	CC111	FMC111	CC112	FMC112	CC113	FPC113	CC114	FPC114	CC 115	FMC115	CC116	FMC116
80 Q																					
111 Q			X	X	X	X	X	X	X	X-D	X	X-D	X-D	X		X	X		X-D	X	
114 Q			X	X	X	X	X	X	X	X-D	X	X	X	X-D		X	X		X	X	
125 Q			X	X	X	X	X	X	X	X	X	X-D	X-D	X		X	X		X	X	
162 Q		X									X		X								
183 N																					
191 Q																					
207 N			D							D			D								
212 Q			D							D			X								
272 N			D	D	X-D	D	D		X	X-D	X-D	X-D	X-D	X	X-D	X-D	X-D		X-D	X-D	D
299 N		X	X-D	X	X-D	X-D	X	X		X-D	D	X-D	X-D	X-D		X-D	X-D		X-D	X-D	X
303 N		X	X-D	X	X-D	X-D	X-D	X-D		X-D	D	X-D	X-D	X-D		X-D	X-D		X-D	X-D	X-D
359 N						X				X-D		X				X-D	D				
369 Q						X				X-D		X				X	X		X		
384 N			X	X-D	X-D	X-D	X-D		D	X-D		X	X-D		X-D	X	D		X		
435 N									X			D	X-D		X					X	X
458 N												D	X		X-D			X	X		
491 N	X	D									X		X-D	X	X				X		
507 N													X								
527 N										X		X				X					
528 N										X		X				X					
531 Q										X		X-D				D					
537 Q										X		X				D					
540 Q										X		X				X					
546 Q			X	X						X						X	X		X		X
555 Q			X-D	X						X-D						X	X		X		X

620 N			X-D	X-D	X-D	X-D	X-D	X-D		X-D	X-D	X-D	X-D	X-D	X-D	X	X-D		X-D	X	X
668 N												X									
723 N			X-D	X-D	X-D	X-D	X-D	X-D		X-D		X-D	X-D	X-D		X-D	X-D				
734 Q			X-D	X-D	X-D	X-D	X-D	X-D		X-D		X-D	X-D	X-D		X-D	X-D				
750 N			D	X-D	X-D	X-D	X-D	X-D	X	X-D	X	X-D	X-D	X-D	X-D	X-D	X-D		X-D	X	X-D
768 N			X-D	X-D	X-D	X	X-D	X-D	X	X-D	X-D	X-D	X-D	X-D	X-D	X-D	X-D		X-D	X	X
825 Q																					
864 Q			X-D	X	X-D	X-D	X	X-D		X-D	X-D	X-D	X-D	X-D	X	X-D	X-D		X-D	D	X-D
918 N		X	X-D	X-D	X-D	X-D	D	D	X-D	X-D	X-D	X-D	X-D		X		X-D		X		
929 N																	X-D				
932 N																	X-D				
941 Q																					
953 N									X			X	X	X	X		X		X		
977 N											X		X								
1002 Q		X											D								
1026 N												D		D					X	X	
1029 Q												D		X					X	X	
1041 Q												D		X-D					X	X	
1083 Q																					
1086 Q																					

Table A3. Deamidation sites of human Collagen alpha-1(I).In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from Pompeii X indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

Collagene alpha 1 (I)		
Position AA	Pompei 3A	Pompei 3D
162 Q		
199 Q	X	
202 Q	X	
229 N		
250 Q	X	
295 N	X	
300 Q	X	
324 N	X	X
358 Q	X-D	X
367 Q	X-D	
387 N	X-D	
393 Q	X	
400 N	X	X-D
421 Q	X-D	
432 N	X	X
457 Q	XD	X
561 Q		
576 Q	X-D	X-D
619 Q	X	
634 Q		
643 Q		
661 Q		
688 Q	X-D	X
700 N	X	
705 N	X	
721 Q	X	
727 Q	X-D	
822 Q		
855N	X	
891N	X	X
952 Q	X	
957 Q	X	
966 Q	X-D	X
985 Q		
1090 Q	X-D	
1102 Q		
1117 Q	X	

1129 Q	X	
1156 N	X-D	X
1203 Q	X	
1206 Q	X-D	

Table A4. Deamidation sites of human Collagen alpha-2(I).In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from Pompeii X indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

Collagene alpha 2 (I)		
Position AA	Pompei 3A	Pompei 3D
80 Q		
111 Q	X	
114 Q	X-D	
125 Q	X	
162 Q	X-D	X
183 N		
191 Q		
207 N		
212 Q		
272 N	X-D	X
299 N	X-D	X
303 N	X-D	X
359 N	X	
369 Q	X	
384 N	X	
435 N	X-D	
458 N	X-D	X
491 N	X-D	X
507 N		
527 N		
528 N		
531 Q		
537 Q		
540 Q		
546 Q		
555 Q		
620 N	X-D	X
668 N	X	
723 N		
734 Q		
750 N	X-D	X
768 N	X	
825 Q	X-D	

864 Q	X-D	X-D
918 N	X-D	
929 N		
932 N		
941 Q		
953 N	X	X
977 N		
1002 Q	X	
1026 N		
1029 Q		
1041 Q		
1083 Q		
1086 Q		

Table A5: Deamidation sites of human Collagen alpha-1(I). In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from ScalandronebayX indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

ScalandroneCollagene alpha 1 (I)														
Position AA	CC 180	FMC 180	CC181	FMC181	CC182	RC182	CC183	FMC183	CC 184	FMC184	CC 186	FMC186	CC187	FMC187
162 Q														
199 Q	X	X-D			X	X	X	X	X-D	X	X	X	X-D	X
202 Q	X-D	X			D	X	X-D	X	X	X	X-D	X	X-D	X-D
229 N		X												
250 Q	X-D	X-D	D		X-D	X	X	X	X					
295 N	X	X			X	X-D	X-D	X	X-D	X-D	X-D	X		
300 Q	X	D			X	X-D	D	X	X-D	X	X	X		
324 N	X				X	X	X-D	X	X-D	X-D	X			X
358 Q			X			X		X	X		D			X
367 Q	X	X	X-D		X-D	X-D	D	X-D						
387 N		D			X-D	X-D	X-D	X	X	X		X		
393 Q		D			X	X	X	X	X-D	X-D		D		
400 N	X-D	X	D	X-D	X	X-D	X-D	X-D	X	X-D	X		X	
421 Q	D	X-D	X	D	X-D	X-D	X-D	X-D	X-D					
432 N								X						
457 Q	X-D	X-D			X-D	X-D	X	X	X-D	X				
561 Q		X			X		X	X				X		X
576 Q	X	X-D		X-D	X-D	X-D	X-D	X-D	X-D	X		X		X
619 Q	D			X	X-D	X-D	X-D		X-D	X-D	D	X		X
634 Q							X		X		X-D			X
643 Q							X		X		X			X
661 Q							X		X		X-D			X
688 Q	X-D	X-D		X-D	X-D	X-D	X	X	D	D				
700 N	D				X			X	X		X-D			
705 N	D				X-D			X-D	X-D		D			
721 Q	X	X-D			X	X-D	X-D	X	X-D	X-D	X-D	X		
727 Q	X-D	X			X-D	X-D	X-D	X	X-D	X	X	X		
822 Q	X	X-D			X	X-D	X-D	X	X	X	X-D			X-D
855N	X	X			X	X	X-D	X	X-D		D		X	
891N	X-D	X		D	X-D	X-D	X-D	X-D	X-D	X-D	X	X		
952 Q	X-D	X-D			X	X-D	X	X	X	X	X-D			X
957 Q	X	D			X	X-D	X-D	X	X-D	X	X-D			X
966 Q	X-D	X	X	X	X	X-D	X-D	X-D			X		X	
985 Q	X-D				X	X-D	X	X			X			X
1090 Q	X-D	X-D			X	X-D	X	X-D	D	X-D				

1102 Q														
1117 Q	X	X			X	X-D	X-D	X	X	X	X	X		X
1129 Q	X-D	X-D			X-D	X-D	X-D	X-D	X-D	X-D	X-D	D		X-D
1156 N	X-D	X-D	X-D	D	X-D	X-D	X-D	X	X-D	X-D	X-D	X	X	X
1203 Q														
1206 Q														

Table A6: Deamidation sites of human Collagen alpha-2(I).In the table the putative deamidation sites of collagen alpha 1(I) and their occurrence in the samples from ScalandronebayX indicates that the site was detected as not deamidated, while D indicates that the site was observed in the deamidated form. X-D indicates that both deamidated and non deamidated form were detected.

ScalandroneCollagene alpha 2 (I)														
Position AA	CC 180	FMC 180	CC181	FMC181	CC182	RC182	CC183	FMC183	CC 184	FMC184	CC 186	FMC186	CC187	FMC187
80 Q														
111 Q					X	X	X	X	X-D			X		X
114 Q					X	X	X	X	X			X		X
125 Q					X	X-D	X	X	X-D			X		X
162 Q	D	X-D	D		X-D	X-D	X	X-D		X				
183 N														
191 Q														
207 N											X-D			
212 Q											X-D			
272 N	D	X-D	X-D		X-D	X-D	X-D	X-D	X-D	X-D	X-D	X	D	X
299 N		X-D			X	X	X-D		X-D		X		X-D	X-D
303 N		X			X	X	X-D		D		X-D		X-D	X-D
359 N						X	X		X		X			
369 Q						X	X		X-D		X			D
384 N	X-D	X-D	D		X-D	X-D	X-D	X	X	X-D	X	X		
435 N	X-D	X-D			X-D	X-D	X-D	X-D	X	X-D				X
458 N	X	X		X-D	X-D	X-D	X-D	X	X-D	X-D				X
491 N	X	X		D	X-D	X-D	X-D	X-D	X-D	D				
507 N														
527 N						X	X		D					
528 N						D	X		X					
531 Q						X	X		X					
537 Q						X	X		X					
540 Q						X	X		D					
546 Q							X		X					X
555 Q							X		X					X
620 N	X-D	X-D			X-D	X-D	X-D	X-D	X-D	X	X	X	X	X-D
668 N					X									
723 N	X				X-D	X	X-D	X	X-D	X-D	X-D			X-D
734 Q	X-D				X-D	X-D	X-D	X	X-D	X-D	X-D			D
750 N	X-D	X			X-D	X-D	X-D	X	D	X	X-D			X-D
768 N	X-D	X			X	X	X-D	X	X-D	X	X-D			X-D
825 Q		D			X-D	X-D	X	X-D						
864 Q	X-D	X-D	D		X	X-D	X-D	X	X-D	X-D	X-D	X-D	X-D	X-D
918 N	X-D	X-D			X-D	X-D	X-D	X	X-D	X-D		X		
929 N					X	X			X-D					

932 N					X	D			D					
941 Q														
953 N								X-D	X			X	X	
977 N					D					X-D				
1002 Q	D	X-D			X	X-D	X-D	X		D				
1026 N							X							
1029 Q							X							
1041 Q							X-D							
1083 Q														
1086 Q														

Table A7: Samples from Herculaneum. Percentage ratios between detected deamidation and detected sites (DD/DS) and ratio between detected deamidation and hypothetical deamidation sites (DD/HDS) were calculated for each sample for human Collagen alpha 1(I) and Collagen alpha (I)

Herculaneum																							
Protein		CC104	FPC106	CC107	FMC107	CC108	FMC108	CC109	FMC109	FMC110	CC111	FMC111	CC112	FMC112	CC113	FPC113	CC114	FPC114	CC 115	FMC115	CC116	FMC116	Average
Collagene alpha 1 (I)	DD/DS (%)	50	75	41	48	59	38	46	50	33	68	37	76	79	76	50	25	61	17	55	50	8	50
	DD/HDS (%)	5	15	22	29	41	24	29	27	15	51	24	61	56	39	27	15	34	2	15	7	2	26
Collagene alpha 2 (I)	DD/DS (%)	0	17	67	50	79	63	64	58	20	72	50	64	71	65	55	50	67	0	38	29	40	48
	DD/HDS (%)	0	2	26	17	24	22	20	15	4	39	15	39	37	24	13	24	30	0	17	9	9	18

Table A8. Samples from Pompeii. Percentage ratios between detected deamidation and detected sites (DD/DS) and ratio between detected deamidation and hypothetical deamidation sites (DD/HDS) were calculated for each sample for human Collagen alpha 1(I) and Collagen alpha (I)

Pompeii				
Protein		A3	D3	Average
Collagene alpha 1 (I)	DD/DS (%)	38	20	29
	DD/HDS (%)	29	5	17
Collagene alpha 2 (I)	DD/DS (%)	50	10	30
	DD/HDS (%)	28	2	15

Table A9. Samples from Scalandrone bay. Percentage ratios between detected deamidation and detected sites (DD/DS) and ratio between detected deamidation and hypothetical deamidation sites (DD/HDS) were calculated for each sample for human Collagen alpha 1(I) and Collagen alpha (I)

Scalandrone bay																
Protein		CC 180	FMC 180	CC181	FMC181	CC182	RC182	CC183	FMC183	CC 184	FMC184	CC 186	FMC186	CC187	FMC187	Average
Collagene alpha 1 (I)	DD/DS (%)	61	56	57	75	45	76	63	28	61	50	56	13	33	18	49
	DD/HDS (%)	41	37	10	15	34	54	49	22	46	27	34	5	5	7	28
Collagene alpha 2 (I)	DD/DS (%)	79	67	100	100	54	55	50	37	66	73	64	11	67	53	63
	DD/HDS (%)	24	22	9	4	28	35	35	15	41	24	20	2	9	20	20

Appendix B

Table B1. Samples used in the database construction. A random numbering and a letter that indicate the binders are shown: milk (M), eggs (E) and animal glue (G). For each sample, the pictorial pigment and some specific characteristics such as the state of aging, special pre-treatment that has been used in the classical proteomic approaches and the mixing with other protein-based binders were reported. In apex is indicated, if present, the reference number corresponding to the paper where results of singles sample are shown.

Sample	Pigment	Characteristic	Pre-treatment	Mass Spectrometry analysis
M1[1]	Minium ($Cu_2CO_3(OH)_2$)	-	-	MALDI-TOF; LC-MS/MS
M2[1]	Cinnabar (HgS)	-	-	MALDI-TOF; LC-MS/MS
M3[1]	-	-	-	MALDI-TOF; LC-MS/MS
M4[1]	$CaCO_3$	-	-	MALDI-TOF; LC-MS/MS
M5	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
M6	Yellow ochre ($Fe_2O_3H_2O$)	-	-	LC-MS/MS
M7	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$	-	-	LC-MS/MS
M8	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$	-	-	LC-MS/MS
M9	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$	-	-	LC-MS/MS
M10	-	-	-	LC-MS/MS
M11[1]	Minium ($Cu_2CO_3(OH)_2$)	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
M12[1]	Cinnabar (HgS)	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
M13[1]	-	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
M14[1]	$CaCO_3$	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
M15	Red ochre (Fe_2O_3)	-	-	MALDI-TOF; LC-MS/MS
M16[1]	$CaCO_3$	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
M17[1]	Azurite ($Cu_3(CO_3)_2(OH)_2$)	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS

M18[1]	Cinnabar (HgS)	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
M19[1]	Minium ($Cu_2CO_3(OH)_2$)	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
M20[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M21[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M22[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M23[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M24[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M25[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M26[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M27[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M28[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M29[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M30[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M31[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M32[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M33[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M34[2]	-	Buddha's decoration	Urea 6M	LC-MS/MS
M35[3]	Red ochre (Fe_2O_3)	-	N-Glycosidase F	MALDI-TOF; LC-MS/MS
M36[3]	Red ochre (Fe_2O_3)	-	N-Glycosidase F	MALDI-TOF; LC-MS/MS

M37[4]	-	-	-	LC-MS/MS
M38[4]	-	-	-	LC-MS/MS
M39[4]	-	Camposanto Monumentale (Pisa)	-	LC-MS/MS
M40[5]	Hematite (α-Fe_2O_3)	-	-	LC-MS/MS
M41[5]	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
M42[5]	Malachite ($Cu_2CO_3(OH)_2$)	Paraloid B72	-	LC-MS/MS
M43[5]	Red pigment (GA-07)	-	-	LC-MS/MS
E1	-	-	-	LC-MS/MS
E2	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$		-	LC-MS/MS
E3	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
E4	Yellow ochre ($Fe_2O_3H_2O$)	-	-	LC-MS/MS
E5	White St. John's ($CaCO_3$)	-	-	LC-MS/MS
E6[5]	Hematite (α-Fe_2O_3)$_2$)	-	-	LC-MS/MS
E7	-	-	-	MALDI-TOF; LC-MS/MS
E8	-	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
E9	$CaCO_3$	-	-	MALDI-TOF; LC-MS/MS
E10	$CaCO_3$	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
E11[3]	Cinnabar (HgS)	-	N-Glycosidase F	MALDI-TOF; LC-MS/MS
E12	Cinnabar (HgS)	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
E13[3]	Minium ($Cu_2CO_3(OH)$	-	N-Glycosidase F	MALDI-TOF; LC-MS/MS
E14	Minium ($Cu_2CO_3(OH)$	Artificially aged sample	-	MALDI-TOF; LC-MS/MS
E15[3]	Azurite ($Cu_3(CO_3)_2(OH)_2$)	-	N-Glycosidase F	MALDI-TOF; LC-MS/MS
E16	Azurite ($Cu_3(CO_3)_2(OH)_2$)	Artificially	-	MALDI-TOF;

		aged sample		LC-MS/MS
E17[3]	Red ochre (Fe_2O_3)	Mixed with milk	N-Glycosidase F	LC-MS/MS
E18	Red ochre (Fe_2O_3)	Artificially aged sample	-	LC-MS/MS
E19[3]	Red ochre (Fe_2O_3)	-	N-Glycosidase F	LC-MS/MS
E20[5]	GA04	-	-	LC-MS/MS
E21[6]	-	Gilding	-	LC-MS/MS
E22[6]	-	Gilding	-	LC-MS/MS
E23[6]	-	Gilding	-	LC-MS/MS
E24[6]	-	Gilding		LC-MS/MS
E25	Cn/tg Cinabro (HgS)	Fat tempera and linseed oil	-	LC-MS/MS
E26	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$	-	-	LC-MS/MS
E27	Verdigri (Cu)	UV	-	LC-MS/MS
E28	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
E29	White St. John's ($CaCO_3$)	-	-	LC-MS/MS
E30	Yellow ochre ($Fe_2O_3H_2O$)	-	-	LC-MS/MS
E31	Biacca $(PbCO_3)_2 \cdot Pb(OH)_2$	-	-	LC-MS/MS
E32	-	-	-	LC-MS/MS
E33[5]	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
E34[5]	Hematite (α-Fe_2O_3)	Trypsin digestion in Microwave	-	LC-MS/MS
E35[5]	Malachite ($Cu_2CO_3(OH)_2$)	Paraloid B72	-	LC-MS/MS
E36[5]	Malachite ($Cu_2CO_3(OH)_2$)	Paraloid B72 and Trypsin digestion in Microwave	-	LC-MS/MS
E37	Azurite ($Cu_3(CO_3)_2(OH)_2$)	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
E38	Minium ($Cu_2CO_3(OH)$	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS

E39	Cinnabar (HgS)	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
E40[3]	Red ochre (Fe_2O_3)	Mixed with animal glue	N-Glycosidase F	MALDI-TOF; LC-MS/MS
E41	$CaCO_3$	Mixed with animal glue	-	MALDI-TOF; LC-MS/MS
E42[3]	Purple pigment	-	N-Glycosidase F	LC-MS/MS
E43[3]	DSFL6	Old gilding sample	-	LC-MS/MS
G1	-	-	Urea 6M	MALDI-TOF; LC-MS/MS
G2	Minium ($Cu_2CO_3(OH)_2$)	-	Urea 6M	MALDI-TOF; LC-MS/MS
G3[3]	Yellow ochre ($Fe_2O_3H_2O$)	-	Urea 6M	MALDI-TOF; LC-MS/MS
G4	Cinnabar (HgS)	-	Urea 6M	MALDI-TOF; LC-MS/MS
G5	$CaCO_3$	-	Urea 6M	MALDI-TOF; LC-MS/MS
G6	Azurite ($Cu_3(CO_3)_2(OH)_2$)	-	Urea 6M	MALDI-TOF; LC-MS/MS
G7[3]	Yellow ochre ($Fe_2O_3H_2O$)	-	Urea 6M and N-Glycosidase F	MALDI-TOF; LC-MS/MS
G8	Minium ($Cu_2CO_3(OH)_2$)	Mixed with milk	Urea 6M	MALDI-TOF; LC-MS/MS
G9[3]	Yellow ochre ($Fe_2O_3H_2O$)	Mixed with milk	Urea 6M	MALDI-TOF; LC-MS/MS
G10	Cinnabar (HgS)	Mixed with milk	Urea 6M	MALDI-TOF; LC-MS/MS
G11	$CaCO_3$	Mixed with milk	Urea 6M	MALDI-TOF; LC-MS/MS
G12	Azurite ($Cu_3(CO_3)_2(OH)_2$)	Mixed with milk	Urea 6M	MALDI-TOF; LC-MS/MS
G13	Minium ($Cu_2CO_3(OH)_2$)	Mixed with egg	Urea 6M	MALDI-TOF; LC-MS/MS
G14	Yellow ochre ($Fe_2O_3H_2O$)	Mixed with egg	Urea 6M	MALDI-TOF; LC-MS/MS

G15	Cinnabar (HgS)	Mixed with egg	Urea 6M	MALDI-TOF; LC-MS/MS
G16	$CaCO_3$	Mixed with egg	Urea 6M	MALDI-TOF; LC-MS/MS
G17	Azurite ($Cu_3(CO_3)_2(OH)_2$)	Mixed with egg	Urea 6M	MALDI-TOF; LC-MS/MS
G18[4]	-	Camposanto Monumentale (Pisa)	-	LC-MS/MS
G19[4]	-	-	-	LC-MS/MS
G20[4]	-	-	-	LC-MS/MS
G21[5]	Hematite (α-Fe_2O_3)	-	-	LC-MS/MS
G22[5]	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
G23	Blue pigment	-	-	LC-MS/MS
G24	-	Gilding	-	LC-MS/MS
G25	Yellow ochre ($Fe_2O_3H_2O$)	-	-	LC-MS/MS
G26	-	-	-	LC-MS/MS
G27	Malachite ($Cu_2CO_3(OH)_2$)	-	-	LC-MS/MS
G28	White St. John's ($CaCO_3$)	-	-	LC-MS/MS
G29	Biacca ($PbCO3*Pb(OH)_2$)	-	-	LC-MS/MS

1. Vinciguerra, R., De Chiaro, A., Pucci, P., Marino, G. & Birolo, L. Proteomic strategies for cultural heritage: form bones to paintings. *Microchem. J.* **126,** 341–348 (2015).
2. Lluveras-Tenorio A., Vinciguerra R., Galano E., Blaensdorf C., Emmerling E., Colombini M.P., Birolo L., Bonaduc1 I. GC/MS and Proteomics to unravel the painting history of the lost Giant Buddhas of Bāmiyān (Afghanistan). *PlosOne*, (2017) in press
3. Vinciguerra, R. *et al.* Deglycosylation Step to Improve the Identification of Egg Proteins in Art Samples. *Anal. Chem.* 10178–10182 (2015). doi:10.1021/acs.analchem.5b02423
4. Leo G, Bonaduce I., Andreotti A., Marino G., Pucci P., Colombini M.P., Birolo L. Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. *Anal. Chem.* **83,** 2056–2064 (2011).
5. Leo G., Cartechini L., Pucci P., Sgamellotti A., Marino G., Birolo L. Proteomic strategies for the identification of proteinaceous binders in paintings. *Anal. Bioanal. Chem.* **395,** 2269–2280 (2009).
6. Gambino M., Cappitelli F., Cattò C., Carpen A., Principi P., Ghezzi L., Bonaduce I., Galano E., Pucci P., Birolo L., Villa F., Forlani F.. A simple and reliable methodology to detect egg white in art samples. *J Biosci.* 2013 Jun;38(2):317-28. doi: 10.1007/s12038-013-9321-z.

Table B2. Animal glue containing samples: Numbering of samples refers to Table B1. When available, detection by MALDI-TOF and LC-MSMS are given. X indicates detection of the peptide, D detection of the peptide in the monodeamidated form, 2D detection of the peptide with 2 deamidations. In the statistical analysis to prepare the graphic of occurrence (Figure 5 and Figure S3), detected peptides have been considered only once, regardeless the mass spectrometric tecnhnique, and deamidation. Modified peptides (Oxidation, pyroGlu formation, and hydroxylation of proline and lysine) were considered separately.

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	5330.27	487-492 R.GDPGSR.G		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X											
	5461.35	487-492 R.GDPGSRG + Hydroxy(P)				X		X		X				X																																		
	5468.27	442-447 K.GDTGAK.G																										X		X																	
	5586.29	505-510 K.GPAGER.G		X		X		X		X		X		X		X		X										X		X																
	6168.28	115-120 K.GDTGPR.G + Hydroxy(P)		X		X		X		X																X																					
	6328.34	304-309 R.GLDGER.G																										X		X																	
	6444.34	304-309 R.GLDGER.G + Hydroxy(P)		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X											
	6447.30	1026-1032 R.IDGPAGK.G + Hydroxy(P)																												X		X																
	6708.32	676-684 R.GPRGER.G + Hydroxy(P)		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X		X											
	7002.34	586-593 K.GAAGERGK.A + Hydroxy(P)									X																																					
	7222.38	220-227 R.GPPGPPGK.N + Hydroxy(P)														X																															
	7330.38	303-309 K.EGSKGPR.G																																													
	7338.38	220-227 R.GPPGPPGK.N + Hydroxy(P,K)		X		X																																									
	7617.40	511-519 R.GAPGPAGPK.G + Hydroxy(P)		X		X						X																		X		X															
	7633.44	733-741 R.GAAGLPGPK.G + Hydroxy(P)				X		X		X		X										X								X		X		X													
	7685.38	277-285 K.GDAGPAGPK.G + Hydroxy(P)												X																																	
	7699.43	733-741 R.GAAGLPGPK.G + 2 Hydroxy(P)																																													

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	838.44	244-252 R.GPPGPDGAR.G	1			X		X		X						X		X		X		X		X		X		X																			
	838.44	1084-1092 R.GPAGPDGPR.G	1	X	X	X		X		X-O	X	X		X-O	X	X		X				X	X-O	X		X				X	X-O			X-O								X	X		X	X	X
	851.42	269-276 R.GFSGLDGAK.G			X		X			X	X	X		X		X			X	X				X								X		X		X						X	X		X	X	X
	852.43	244-252 R.GPPGPDGAR.G+Hydroxy(P)	1			X		X		X	X	X		X		X		X						X	X	X				X		X		X-O								X	X		X	X	X
	852.43	1084-1092 R.GPAGPDGPR.G+Hydroxy(P)	1	X		X	X-O	X	X				X	X	X		X	X						X	X	X																		X		X		
	854.41	362-380 K.GEGGPDGPR.G	1	X		X		X		X		X		X		X		X				X		X		X				X														X		X		
	867.42	269-276 R.GFSGLDGAK.G+Hydroxy(K)																				X																										
	868.43	1084-1092 R.GPAGPDGPR.G+2Hydroxy(P)	1		X											X	X			X	X					X																					
	868.43	244-252 R.GPPGPDGAR.G+2Hydroxy(P)	1																	X						X																					
	882.52	959-966 R.GVVGLPGQR.G	1											D				D		D				D						X		X		X-O	X-O										
	886.40	362-380 K.GEGGPDGPR.G+2Hydroxy(P)	1			X						X						X		X				X		X																					
	886.43	361-369 R.GSEGPDGVR.G	1	X	X	X		X		X		X		D		X		X		X		X		X		X						X										X		X		
	888.51	959-966 R.GVVGLPGQR.G+Hydroxy(P)	1	X	X-O		X-O			X-O	X	X		D	X			X	X-O	D	X	X	X-O	X	X-O	X		X		X	X-O		X		X-O	X-O						X	X		X	X	
	911.45	141-151 R.GPPGSAAGPPK.D		X		X		X		X		X																				X															
	945.43	984-993 K.QGPSGAAGER.G	1								X																															D						
	1020.49	676-684 R.GERGPPGER+Hydroxy(P)																																		X												
	1088.54	493-504 R.GPPGADGVAGPK.G+Hydroxy(P)			X		X	X	X		X			X			X		X		X		X						X	X		X	X									X	X	X	X	X	X

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	1089.48	430-441 K.GNSGEPGAPGSK.G+Hydroxy(K)	1			D				D																																					
	1095.56	310-321 R.GPPGAPGPAGAR.G+2Hydroxy(P)		X		X		X		X				X								X																			X					
	1095.53	733-744 R.GAAGLPGPK.GDR.G		X		X		X				X		X														X		X																
	1089.55	751-762 K.GADGAPGKDGVR.G																				X																			X					X
	1104.53	493-504 R.GPPGADDQVAGPK.G+2Hydroxy(P)																X																												
	1105.57	685-696 R.GVDGPPGPAGPRG.G+Hydroxy(P)	1	X		X		X		X		X		D		X-D						X-D						X		X													X		X	
	1111.56	310-321 R.GPPGAPGPAGAR.G+3Hydroxy(P)																										X																		
	1111.58	733-744 R.GAAGLPGPKGDR.G+Hydroxy(P)																												X																
	1115.54	751-762 K.GADGAPGKDGVR.G+Hydroxy(P)										X																X		X												X	X			
	1115.54	586-597 K.GAAGEPGKAGERG.G+Hydroxy(P)																												X																
	1116.49	1014-1025 R.EGAPGAEGSPGRDG.D+2Hydroxy(P)										X																																		
	1121.57	685-696 R.GVDGPPGPAGPRG.G+2Hydroxy(P)	1	X				X		X		X				X		X										X		X		X													
	1127.58	733-744 R.GAAGLPGPKGDR.G+2Hydroxy(P)												X								X																								
	1130.59	253-264 R.GLPGTAGLPGMK.G+Hydroxy(P)+Oxi(M)																																		X							X			
	1146.58	253-264 R.GLPGTAGLPGMK.G+2Hydroxy(P)+Oxi(M)																												X																
	1160.64	588-611 R.GVPGPPGAVGPAGK.D																X																												
	1161.57	574-585 R.GDAGVMGPPGPK.G+Hydroxy(P) or Hydroxy(K) or Oxi(M)	1			X-D								D		X-D				X-D				X-D								X-D		X-D		X-D	D	D						X	X		
	1162.57	253-264 R.GLPGTAGLPGMK.G+3Hydroxy(P,K)+Oxi(M)																																												

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	1176.64	599-611 R.GVPGPPGAVGPAGK.D+Hydroxy(P)																																			X										
	1177.57	574-585 R.GQAGVMGFPGPK.G + 2 Hydroxy (P, K) or Hydroxy (P or K)+Oxi(M)	1	X-O		D		D		D		X-O		X		D		X-O		X-O		X-O		X-O						X-O		X-O		X-O		X-O	X-O	O			D	X-O		X	X		
	1192.63	599-611 R.GVPGPPGAVGPAGK.D+2 Hydroxy (P, K)				X								X	X									X				X															X					
	1193.56	574-585 R.GQAGVMGFPGPK.G+2 Hydroxy (P, K) and Oxi (M) or+3 Hydroxy (P, K)	1	X		X		X		X		X	O	X-O				O		X-O		X		X-O						X-O				X-O		X	X	O										
	1208.63	599-611 R.GVPGPPGAVGPAGK.D+3 Hydroxy (P)																																														
	1224.68	965-983 R.GVVGLPGQRGER.G	1																	D																													
	1240.68	969-983 R.GVVGLPGQRGERG.G+Hydroxy (P)	1					D						D				D				D		D				X	X	X-O		D	X	X-O							X		X	X			
	1242.60	538-551 K.GLTGSPGSPGPDGK.T+Hydroxy (P)										X	X									X								X		X																
	1258.59	538-551 K.GLTGSPGSPGPDGK.T+2 Hydroxy (P)				X				X		X	X									X								X		X		X				X	X								
	1281.62	781-795 K.GEAGPSGPAGPTGAR.G																																						X								
	1296.66	970-983 R.GPPGLDGPSGEPGK.Q												X																																		
	1312.65	970-983 R.GPPGLPGPSGEPGK.Q+Hydroxy (P)		X		X		X						X				X		X		X		X						X		X								X							
	1328.65	970-983 R.GPPGLPGPSGEPGK.Q+2 Hydroxy (P)		X				X				X		X				X		X		X		X						X		X		X						X		X	X				
	1332.65	415-429 R.GPSGPQGPSGPPGPK.G+Hydroxy (P, K)	1			X-O						X-O		X-O								X-O								X-O		X-O		X-O		X-O							X		X				
	1344.64	970-983 R.GPPGLPGPSGEPGK.Q+3 Hydroxy (P)				X						X		X				X				X		X						X		X										X					
	1348.65	415-429 R.GPSGPQGPSGPPGPK.G+Hydroxy (P, K)	1																			X						X				X																
	1427.70	865-881 R.GSMGPPGATGPPGAAGR.V														X													X																			
	1435.68	472-486 R.GEPGPAGLPGPPGERGG+3 Hydroxy (P)		X		X		X		X		X		X		X		X		X		X		X		X				X		X		X		X							X		X			

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	1459.69	885-891 R.GSAGPPGATGFPGAAGRLV+2Hydroxy(P)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	1496.75	171-186 K.STGISVPGPMGPSGPR.G																																													
	1512.75	171-186 K.STGISVPGPMGPSGPRGG+Hydroxy(P)																						X																			X				
	1546.75	1093-1107 R.GDKGETGEQGDRGIK.G	1			X			X																																					
	1560.81	1066-1083 R.GETGPAGPAGPPGPVGAR.G			X		X				X	X	X	X	X				X	X	X	X	X						X				X							X	X				X	X	
	1561.80	1152-1167 K.DGLNGLPGPIGPPGPR.G+3Hydroxy(P)	1	D		D		D		D		D		D				D		D		D		D										D					D							
	1563.78	397-410 K.GAIWGAPGAWGAPGPPGAIR.G+2Hydroxy(P)	1											D																					D												
	1588.77	397-414 K.GAIWGAPGIWGAPGPPGAIR.G+3Hydroxy(P)	1	X+D		D		D		D		X+D		D				D		D		D		D		D						D		D		D					D					
	1588.74	430-447 K.GNSGEPGAPGSKGDTGAK.G	1	X		X		X		X		X		X		X		X		X		X		X		X		X		X																
	1590.81	763-789 R.GLTGPIGPPGPAGAPGDK.G+2Hydroxy(P)										X												X																	X						
	1601.83	115-131 K.GDTGPRGPAGPPGPIGPPGPR.D																																										X				
	1605.83	594-611 K.AGERGVVPGPPGAVGPAGK.D+2Hydroxy(P)					X		X		X		X		X				X		X		X		X		X		X																		
	1606.81	763-780) R.GLTGPIGPPGPAGAPGDK.G +3Hydroxy(P,K)												X																																		
	1615.78	487-504 R.GDPGSRGPPGADGVAGPK.G+2Hydroxy(P)											X							X		X							X		X																
	1618.73	430-447 K.GNSGEPGAPGSKGDTGAK.G+2Hydroxy(P)	1																				D																								
	1624.80	745-762 R.GDAGPKGADGAPGKDGVR.G																																										X			
	1627.74	228-243 K.MGDDGEAGKPGRPGERGLPG+Hydroxy(P)	1																				D																								
	1636.84	733-750 R.GAAGLPGPKGDRGDAGPK.G+Hydroxy(P)																														X															

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	1637.82	594-611 K.AGERGVVPGPPGAVGPACK.D + 4 Hydroxy (P)														X																															
	1652.83	733-750 R.GAAGLPGPKGDRGDAGPK.G + Hydroxy (P, K)																				X																									
	1655.80	520-537 K.GSPGEAGPPGEAGLPGAK.G + 3 Hydroxy (P, K)					X		X		X		X		X		X				X		X					X	X			X	X									X	X			
	1655.81	493-510 R.GPPGADGVAGPKGPAGER.G + Hydroxy (P)																										X				X	X										X	X			
	1671.80	520-537 K.GSPGEAGPPGEAGLPGAK.G + 4 Hydroxy (P, K)																				X																									
	1671.81	493-510 R.GPPGADGVAGPKGPAGER.G + 2 Hydroxy (P)																																	X												
	1687.80	493-510 R.GPPGADGVAGPKGPAGER.G + 3 Hydroxy (P)																												X			X														
	1687.85	469-486 R.GARGEPGPAGLPGPPGER.G + Hydroxy (P)																										X		X		X	X							X							
	1690.78	612-630 K.DGEAGAQGPPGPAGPAGER.G	1																																												
	1706.77	612-630 K.DGEAGAQGPPGPAGPAGER.G + Hydroxy (P)	1												D								X+D																								
	1710.75	286-303 K.GEPGSPGENGAPGQMGPR.G + Hydroxy (P) or Oxi	2	X		X				X		X		X		X																														
	1718.84	469-486 R.GARGEPGPAGLPGPPGER.G + 3 Hydroxy (P)																												X		X			X								X	X			
	1728.74	286-303 K.GEPGSPGENGAPGQMGPR.G + 2 Hydroxy (P) or + Hydroxy (P) and Oxi (M)	2	X		X		X		X		X		X		X							X						X		X		X												
	1726.93	152-169 K.DGLNGLPGPIGPPGPRGR.T	1	X		X		X		X		X		X		X							X						X		X		X												
	1742.74	286-303 K.GEPGSPGENGAPGQMGPR.G + 3 Hydroxy (P)	2																			D																								
	1743.85	862-882 K.GARGSAGPPGATGFPGAAGR.V + 2 Hydroxy (P)																																									X			
	1744.77	1014-1032 R.EGAPGAEGSPGPDGSPGAK.G + 3 Hydroxy (P)																				X																								
	1753.82	362-383 K.GEGPQGPRGSEGPQGVRLG + 2 Hydroxy (P)	2	X		X		X		X		X		X																X		X		X										

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	1758.72	286-303 K.GEPGSPGENGAPGQMGPR.G + 4 Hydroxy (P)	2						2D				2D																			
	1769.82	362-389 K.GEGGPGQPGRSGEPGQGVR.G + 3 Hydroxy (P)	2	X		X		X		X							X		X												
	1784.87	994-1013 R.GPPGPMGPPGLAGPPGESGR.E																														
	1800.87	994-1013 R.GPPGPMGPPGLAGPPGESGR.E + Hydroxy (P) or Oxi (M)																														
	1812.88	882-902 R.VGPPGPSGNAGPPGPPGPAGK.E + 3 Hydroxy (P, K)	1		X-D		D		D		X-D		X-D		X-D		X		X-D		X-D		X-D						X		
	1816.86	994-1013 R.GPPGPMGPPGLAGPPGESGR.E + 2 Hydroxy (P) or + Hydroxy (P) and Oxi (M)			X		X		X		X		X		X					X											
	1828.88	882-902 R.VGPPGPSGNAGPPGPPGPAGK.E + 4 Hydroxy (P,	1	D									D	X-D						D		X									
	1832.86	994-1013 R.GPPGPMGPPGLAGPPGESGR.E + 3 Hydroxy (P) or + 2 Hydroxy (P) and Oxi (M)			X		X		X		X		X		X				X												
	1844.88	882-902 R.VGPPGPSGNAGPPGPPGPAGK.E + 5 Hydroxy (P)	1							D																			X		
	1848.85	994-1013 R.GPPGPMGPPGLAGPPGESGR.E + 4 Hydroxy (P) or + 3 Hydroxy (P) and Oxi (M)								X	X		X		X																
	1860.19	574-593 R.GDAGVMGFPGPKGAAGEPGK + 3 Hydroxy (P)	1																										X			
	1883.87	446-467 K.GEPGPTGIQGPPGPAGEEGK.R + 2 Hydroxy (P)	1					X-D	X					X																		
	1951.97	265-285 K.GHRGFSGLDGAKGDAGPAGPK.G									X	X	X		X																	
	1952.95	742-762 K.GDRGDAGPKGADGAPGKDGVR.G																	X		X											
	1962.93	472-492 R.GEPGPAGLPGPPGERGGPGSR.G + 4 Hydroxy (P)								X							X		X		X							X	X			
	1968.95	742-762 K.GDRGDAGPKGADGAPGKDGVR.G + Hydroxy (P)															X															
	1975.89	1062-1083 K.SGDRGETGPAGPAGPIGPVGAR.G		X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X				X	X			X	X	X	X	X
	2003.98	446-468 K.GEPGPTGIQGPPGPAGEEGKR.G + Hydroxy (P)	1					D																								

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	2019.37	448-468 K.GEPGPTGIQGPPGPAGEEGKRGLG + 2 Hydroxy (P, K)	1	X-D		X-D		X-D		X-D		X		X-D		X		X-D		X-D		X-D												X	X-D	X-D				D						
	2040.38	552-573 K.TGPPGPAGQDGRPGPPGPRGAR.G + 3 Hydroxy (P)	1					D										D																													
	2056.38	553-573 K.TGPPGPAGQDGRPGPPGPRGAR.G + 4 Hydroxy (P)	1	X-D		D		D		X-D		X-D		X-D				X-D		X-D		X-D				X				D				X-D	X	X-D	X-D		X			X	X			
	2065.36	1084-1104 R.GPAGPDGPPGPDGNKGETGEQGDRG.G	2																																X									X	X	
	2073.01	934-957 K.GAPGADGPAGAPGTPGPQGIAGQRG.G + Hydroxy (P)	2	2D				2D				D- 2D		D- 2D		2D		X- 2D		X-D		X- 2D		X- 2D											X-D	X-D			X							
	2081.96	1084-1104 R.GPAGPDGPPGPDGNKGETGEQGDRG.G + Hydroxy (P, K)	2																			D- 2D								X		X															
	2083.00	934-957 K.GAPGADGPAGAPGTPGPQGIAGQRG.G + 2 Hydroxy (P)	2	X- 2D		D												X-D 2D				X		X											X-D					X						
	2105.00	934-957 K.GAPGADGPAGAPGTPGPQGIAGQR.G + 3 Hydroxy (P)	2															X																												
	2107.04	676-696 R.GERGFPGERGVQGPPGPAGPR.G + 2 Hydroxy (P)	1																	X		X		X		X																				
	2117.39	802-825 R.GEPGPPGPAGFAGPPGADGQPGAK.G + Hydroxy (P, K)	1																																	D	D										
	2120.39	934-957 K.GAPGADGPAGAPGTPGPQGIAGQRG.G + 4 Hydroxy (P)	2									D												X																						
	2131.02	685-708 R.GVQGPPGPAGPRGANGAPGNDGAKG.G + 2 Hydroxy (P)	3													X		X								X																				
	2133.38	802-825 R.GEPGPPGPAGFAGPPGADGQPGAK.G + 3 Hydroxy (P)	1																					X-D																						
	2163.06	1038-1061 R.GETGPAGPPGAPGAPGPVGPAGK.S + 3 Hydroxy (P)										X												X																						
	2165.05	1038-1061 R.GETGPAGPPGAPGAPGAPGPVGPAGK.S + 4 Hydroxy (P, K)										X		X				X				X		X																						
	2188.37	709-732 K.GDAGAPGAPGSQGAPGLQGMPGER.G + 4 Hydroxy (P) or +3 Hydroxy (P) and Oxi (M)	2	X		D						X-D 2D		X-D		X		X		X		X		X													X	X							
	2214.37	709-732 K.GDAGAPGAPGSQGAPGLQGMPGER.G + 4 Hydroxy (P) and Oxi (M)	2	X		D								X-D 2D		X				X-D		X-D		X-D 2D						X-D					X					X				X		
	2216.06	910-933 R.GETGPAGPRGEVGPPGPPGPAGEK.G + 3 Hydroxy (P, K)						X		X				X				X		X		X								X		X													

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	2232.05	910-933 R.GETGPAGRPGEVGPPGPPGPAGEK.G+4				X	X					X	X	X	X			X	X											X	X				X									X			
	2247.08	469-492 R.GARGERGPAGLPGPPGERGGPGSR.G+4																										X	X	X	X			X												
	2248.05	910-933 R.GETGPAGRPGEVGPPGPPGPAGEK.G+4																										X	X																	
	2249.03	1036-1061 R.GETGPAGPPGAPGAPGPVGPAGK.G+8																														X	X													
	2255.07	970-993 R.GPPGLPGPSGEPGKQGPSGASGER.G+2	1																					D																							
	2258.10	574-597 R.GQAGVMGFPGPKGAAGEPGKAGERGK.G+2 Hydroxy (P) or +Hydroxy (P) and Oxi (M)	1																											X-D		X				X											
	2273.16	586-611 K.GAAGEPGKAGERGVPGPPGAVGPAGK.D+2																													X		X													
	2274.09	574-597 R.GQAGVMGFPGPKGAAGEPGKAGER.G+3	1																											X	X-D				X-D									X			
	2281.19	253-276 R.GLPGTAGLPGMKGHRGFSGLDGAK.G																													X	X														
	2284.06	370-396 R.GEPGPPGPAGAAGPAGNPGADGQPGAK.G+2	2	2D														2D				2D		2D				X				X	X		X											
	2290.09	574-597 R.GQAGVMGFPGPKGAAGEPGKAGER.G+3 Hydroxy (P, K) or +2 Hydroxy (P, K) and Oxi (M)	1																											X-D		X		X													
	2300.05	370-396 K.GEPGPPGPAGAAGPAGNPGADGQPGAK.G+3	2																			X		2D																							
	2307.14	835-861 K.GDAGPPGPAGPAGPPGPIGNVGAPGPK.G+3	1										X-D						X-D				X-D		X-D										X											
	2316.05	370-396 R.GEPGPPGPAGAAGPAGNPGADGQPGAK.G+4	2	X-D 2D									X- 2D		X				X- 2D		D		X-D 2D		X- 2D						X															
	2323.13	835-861 K.GDAGPPGPAGPAGPPGPIGNVGAPGPK.G+4	1																					D																							
	2347.10	220-243 R.GPPGPPGKNGDDGEAGKPGPPGPER.G+3	1																																											
	2363.10	220-243 R.GPPGPPGKNGDDGEAGKPGPPGPER.G+4 Hydroxy (P)	1												D								D																							
	2394.23	1066-1092 R.GETGPAGPAGPIGPVGARGPAGPQGPR.G+ Hydroxy (P)	1	X				X						X		X																														

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	2403.12	415-441 R.GPSGPQGPPGPPGPKGNSGEPGAPGSK.G + 3	2																	2D		2D																			X					
	2404.18	511-537 R.GAPGPAGPKGSPGEAGRPGEAGLPGAK.G + 4																													X													
	2410.22	1066-1082 R.GETGPAGPAGPIGPVGARGPAGPQGPR.G + 2	1	X		X		X		X		X		X				X																											
	2420.18	511-537 R.GAPGPAGPKGSPGEAGRPGEAGLPGAK.G + 4																												X															
	2420.19	493-519 R.GFPGADGVAGPKGPAGERGAPGPAGPK.G + 3																													X														
	2436.17	511-537 R.GAPGPAGPKGSPGEAGRPGEAGLPGAK.G + 6																												X				X							X				
	2454.24	1141-1167 R.GPPGSAGSPGKDGLNGLPGPIGPPGPR.G + 3	1		D		D		D						D						D						D																	
	2470.23	1141-1167 R.GPPGSAGSPGKDGLNGLPGPIGPPGPR.G + 4	1						D				D																						D	D									
	2481.21	1033-1061 K.GDRGETGPAGPPGAPGAPGPVGPAGK.S +		X										X				X		X		X		X				X				X													
	2486.13	1141-1167 R.GPPGSAGSPGKDGLNGLPGPIGPPGPR.G + 5	1	D		D		D		D		D		D				D		D		D		D				D				D		D	D									
	2493.11	277-303 K.GDAGPAGPKGEPGSPGENGAPGQMGPR.G + 3 Hydroxy (P) + 2 Hydroxy (P) and Oxi (M)	2	D-2D										2D				D-2D		D		D-2D		D				D				D													
	2497.21	1033-1061 K.GDRGETGPAGPPGAPGAPGPVGPAGK.S +		X		X		X		X		X						X				X		X				X				X		X							X	X		
	2508.10	277-303 K.GDAGPAGPKGEPGSPGENGAPGQMGPR.G + 4 Hydroxy (P) or + 3 Hydroxy (P) and Oxi (M)	2	D-2D		2D												2D		2D		D-2D						2D				D-2D													
	2513.20	1033-1061 K.GDRGETGPAGPPGAPGAPGPVGPAGK.S +		X		X				X		X						X		X		X		X				X		X		X									X	X		
	2524.25	882-909 R.VGPPGPSGNAGPPGPPGPAGKEGSKGPR.G + 3	1																															D										
	2525.008	277-303 K.GDAGPAGPKGEPGSPGENGAPGQMGPR.G + 4 Hydroxy (P, K) and Oxi (M)	2															2D				2D										X													
	2532.21	322-351 R.GNDGATGAAGPPGPTGPAGPPGPPGAVGAK.G + 2 Hydroxy (P, K)	1	D				D		D		X-D		X-D				D		D		D		D				D		X-D		D													
	2548.20	322-351 R.GNDGATGAAGPPGPTGPAGPPGPPGAVGAK.G + 3 Hydroxy (P)	1	X-D		D		D		D		D		D				X-D		X-D		X-D		X-D		D				X-D				X-D							X	X		

Protein	MH⁺ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha 1(I) (P02465)	2549.22	442-468 K.GDTGAKGEPGPTGIQGPPGPAGEEGKR.G + 2 Hydroxy (P)	1	X	D	X		X		X		X		X	X	X			D	X	X-D	X	X	X		X		X			D	X		X	X-D	X													
	2623.28	835-864 K.GDAGPPGPAGPAGPPGPIGNVGAPGPKGAR.G + 5 Hydroxy (P)	1																									X																					
	2651.36	1141-1169 R.GPPGSAGSPGKDGLNGLPGPIGPPGPRGR.T + 2 Hydroxy (P)	1					X																																									
	2673.26	1111-1140 R.GFSGLQGPPGPPGSPGEQGPSGASGPAGPR.G + Hydroxy	2	X				X				X		X		X								X										X															
	2689.26	1111-1140 R.GFSGLQGPPGPPGSPGEQGPSGASGPAGPR.G + 2 Hydroxy	2	X	X	X	X-D	X	D	X		X	X	X	X-D	X		X		X		X	D	X	X-D	X					X	X		X													X		
	2703.24	796-825 R.GAPGDRGEPGPPGPAGFAGPPGADGQPGAK.G + 4 Hydroxy (P, K)	1		X-D		X-D								D				X-D		X		X		X													X-D	X-D	X-D						X			
	2705.25	1111-1140 R.GFSGLQGPPGPPGSPGEQGPSGASGPAGPR.G + 3 Hydroxy	2	X	D	X	X	X	X-2D	X		X	X-D	X		X			D	X	X-D-2D	X			X-D-2D	X	D	X			X	X	X	X	X	X	X	X-D	X						X	X	X		
	2719.23	796-825 R.GAPGDRGEPGPPGPAGFAGPPGADGQPGAK.G + 5 Hydroxy (P, K)	1		X								X-D		X-D				X-D						D																								
	2775.27	984-1013 K.QGPSGASGERGPPGPMGPPGLAGPPGESGR.E + 4 Hydroxy	1																				X																										
	2853.41	763-795 R.GLTGPIGPPGPAGAPGDKGEAGPSGPAGPTGAR.G + 2 Hydroxy (P, K)		X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X		X	X	X	X	X		X	X	X	X	X	X	X								X	X	X		
	2864.39	598-630 R.GVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGER.G + 2	1																						X														X-D	X									
	2869.41	763-795 R.GLTGPIGPPGPAGAPGDKGEAGPSGPAGPTGAR.G + 3 Hydroxy (P, K)							X				X		X				X		X		X		X						X		X		X		X									X	X		
	2880.39	598-630 R.GVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGER.G + 3 Hydroxy (P, K)	1		X		X		X				X-D						X-D		X		X		X-D													X-D	X-D										
	2883.41	397-429 K.GANGAPGIAGAPGFPGARGPSGPQGPSGPPGPK.G + 3 Hydroxy (P)	2	X		X		X		X		X		X		X				X		X				X		X				X		X		X													
	2883.42	709-741 K.GDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPK.G	2	X		X		X		X		X		X		X				X		X				X		X				X		X		X													
	2883.46	487-519 R.GGPGSRGFPGADGVAGPKGPAGERGAPGPAGPK.G		X		X		X		X		X		X		X				X		X				X		X				X		X		X													
	2901.40	763-795 R.GLTGPIGPPGPAGAPGDKGEAGPSGPAGPTGAR.G + 5 Hydroxy (P)																																		X													
	3084.41	187-219 R.GLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPR.G + 4 Hydroxy (P) and Oxi (M)	2						D		X										X-D		X			X						X			X	X									X				

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF																												
				LC-MSMS																												
Collagen alpha 1(I) (P02465)	3000.41	187-226 R.GLPGPPGAPGPQGPPGEPGPAGSPMGPRLG+5 Hydroxy (P) and Oxi (M)	2	X		X		X-D				X-D		X			X-D			X								X	X		
	3116.40	187-226 R.GLPGPPGAPGPQGPPGEPGPAGSPMGPRLG+7 Hydroxy (P) or + 6 Hydroxy (P) and Oxi (M)	2	X		X					D			X-D						X											
	3180.47	361-396 R.GSEGPQGIRGEPGPPGPAGAAGPAGNPGADGQPGAK.G+4 Hydroxy (P)	3	3D						X-D				3D			X-D-2D			X								X	X		
	3224.43	637-732 R.GAAGIAGAPGMDGAKGDAGAPGAPGSQGAPGLQGMPGER.G+5 Hydroxy (P) or + 4 Hydroxy (P) and Oxi (M)	4			2D			2D			2D		2D-3D																		
	3240.40	637-732 R.GAAGIAGAPGMDGAKGDAGAPGAPGSQGAPGLQGMPGER.G+5 Hydroxy (P) and Oxi (M)	4	4D					3D-4D	4D		4D		3D																		
	3296.55	538-573 K.GLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPAPG.G+6 Hydroxy (P)	1															D	X		X								X		
	3320.51	268-303 R.GPSGLDGAKGDAGPAGPKGEPGSPGENGAPGQMGPR.G + 2 Hydroxy (P, K) and Oxi (M)	2													X	D-2D	D	D	D-2D											
	3344.51	268-303 R.GPSGLDGAKGDAGPAGPKGEPGSPGENGAPGQMGPR.G+3 Hydroxy (P) and Oxi (M)	2							D							D-2D			D-2D	D-2D										
	3360.60	222-360 R.GMDGATGAAGPPGPTGPAGPPGFPGAVGAKGEGGPGQPR.G. +3 Hydroxy (P, K)	2																										X			
	3501.59	796-834 R.GAPGDRGEPGPPGPAGFAGPPGADDGPGAKGEPGDAGAK.G+5 Hydroxy (P, K)	1	X					X-D		D			X-D						X												
	3616.69	430-469 K.GNIGEPGAPGSKGDTGAKGEPGPTGIQGPPGPPGPAGEEGKR.G +4 Hydroxy (P, K)	2															2D		2D												
	3634.82	370-444 R.GEPGPPGPAGAAGPAGNPGADGQPGAKGAAGAPGNKGAP GPPGARG.+4 Hydroxy (P, K)	3																			D										

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	584.38	1006-1021 R.GLPGLK.+Hydroxy(P,K)																																												
	691.27	353-368 K.GAESGNKDG.G	1																																											
	698.27	1007-1012 K.GEPGDKGG.+Hydroxy(P)								X																																				
	698.28	707-712 R.GGPGER.G.+Hydroxy(P)				X		X			X																																			
	628.34	216-220 R.GLPGER.G																				X								X		X		X												
	632.34	659-661 K.GETGLR.G		X		X				X	X																																			
	644.33	216-220 R.GLPGER.G.+Hydroxy(P)		X		X		X		X		X				X		X		X		X		X				X		X		X		X											
	647.35	737-742 K.GEPGTK.G																								X		X		X		X		X												
	774.34	497-504 K.GPSGDPGK.A																											X																	
	722.38	131-138 R.GPPGPPGK.A.+Hydroxy(P)														X																														
	738.38	131-138 R.GPPGPPGK.A.+Hydroxy(P,K)				X																																								
	742.38	404-411 R.GLPGADGR.A																										X																		
	743.40	644-652 R.GAAGPPGKGH.+Hydroxy(P)									X												X						X																	
	758.38	404-411 R.GLPGADGR.A.+Hydroxy(P)															X						X						X		X		X		X											
	784.44	914-920 K.EGLRGPR.G		X				X				X		X				X		X		X				X		X		X		X		X											
	785.39	821-828 R.GDQGPVGR.G	1		X	X-D				X	X	X				D						D							X		X		X-D		X-D	X		D							
	785.43	442-450 R.GATGPAGVR.G											X	X								X						X																	
	809.44	509-517 K.GHAGLAGAR.G		X		X	X	X	X	X		X		X				X		X		X				X				X															

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	840.46	165-183 R.GAVIGPQGAAR.G	1	x		x						x		x		x						x	x:D														x:D				D	x:D	x	x	x		x
	841.46	216-222 R.GLPGEPGR.V																										x		x																
	882.46	223-232 R.VGAPGPAGAR.G		x		x		x				x		x				x						x		x																			
	887.46	216-222 R.GLPGEPGR.V + Hydroxy (P)																												x																
	888.46	223-232 R.VGAPGPAGAR.G + Hydroxy (P)		x		x		x				x		x				x				x		x		x		x		x				x										
	888.46	995-1003 R.GPSGPQGIR.G	1			x							x:D				D	x				x:D				x				x						D									
	888.46	1062-1062 R.GPAGPSGPAGK.D						x				x										x							x		x														
	902.45	412-421 R.AGIVMGPAGSR.G														x						x							x		x				x									
	911.45	1062-1062 R.GPAGPSGPAGK.D + Hydroxy (P)		x		x		x		x		x														x						x												
	918.45	412-421 R.AGIVMGPAGSR.G + Hydroxy (P)						x						x		x						x							x		x				x										
	934.44	412-421 R.AGIVMGPAGSR.G + Hydroxy (P) and Oxi (M)																																				x								
	1067.56	644-655 R.GAAGIPGGKGEK.G + Hydroxy (P)																				x										x													
	1080.53	221-232 R.GPIVGAPGPAGAR.G + Hydroxy (P)																																										x			
	1084.56	341-352 R.GLVGEPGPAGSK.G + Hydroxy (P)											x									x							x																
	1096.50	609-619 R.GPSGPPGPDGNK.G + Hydroxy (P)	1	x		x		x				x		x														x		x				x										
	1096.56	818-828 R.GPPGADGQPVVGR.S	1	x		x		x				x																	x																
	1100.56	341-352 R.GLVGEPGPAGSK.G + Hydroxy (P, K)																										x							x										
	1111.50	609-619 R.GPSGPPGPDGNK.G + 2 Hydroxy (P)	1																									x																

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	1111.56	889-898 R.GPRGDQGPVGR.S + Hydroxy (P)	1																										X																	
	1116.53	497-508 K.GPSGDPGKAGEK.G + Hydroxy (P)																											X		X															
	1143.49	609-619 R.GPSGPPGPDGNK.G + 4 Hydroxy (P)	1	X																																										
	1170.56	662-673 R.GDDGSPGRDGARG + Hydroxy (P)																				X																									
	1183.57	85-97 K.GGGPGPMGLMGPR.G																																		X										
	1192.68	1066-1078 R.GDPGAVGPAGIR.G	1	X-O								X-O						X-O				X-O						X-O				X-O		X-O		X-O								X	X	X
	1198.67	86-97 K.GGGPGPMGLMGPR.G + Hydroxy (P) or Oxi (M)																								X				X																
	1201.58	405-416 R.GEPGNVGFPGPK.G + 2 Hydroxy (P, K)	1	X-O				X-O						D		D		D		X-O		D		D				X-O				X-O		X-O		X-O								X	X	
	1206.61	164-175 R.GPPGTPGLPGPK.G + 2 Hydroxy (P)										X		X										X				X		X																
	1208.67	1066-1078 R.GDPGAVGPAGIR.G + Hydroxy (P)	1	X-O		X-O				D		X-O		D				X-O		X-O		X-O		X-O				D		X-O		D		X-O		X-O				D		D		X		
	1222.61	164-175 R.GPPGTPGLPGPK.G + 3 Hydroxy (P, K)				X		X		X		X		X		X		X		X		X		X						X				X									X	X		
	1223.61	1062-1085 R.GPAGPSGPAGKDGR.I																				X								X	X													X	X		
	1238.61	164-175 R.GPPGTPGLPGPK.G + 3 Hydroxy (P, K)										X																																			
	1239.61	1062-1085 R.GPAGPSGPAGKDGR.I + Hydroxy (P)																														X															
	1251.68	326-340 R.GPPGPVGAAGATGAR.G																																X												
	1261.63	692-706 R.GEAGPAGPAGPAGPR.G										X		X								X														X								X		
	1267.67	326-340 R.GPPGPVGAAGATGAR.G + Hydroxy (P)		X		X		X		X		X				X		X		X		X		X				X		X		X		X		X						X			
	1283.67	326-340 R.GPPGPVGAAGATGAR.G + 2 Hydroxy (P)																X										X		X																

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites		G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17		G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
					MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	12897.63	449-462 R.GPPGSPGNIGPAGKE.+2 Hydroxy(P)	1			X-D		X-D				X-D		X-D		X-D	D		X-D		X-D		X-D						X-D				X-D		X-D	X-D	X-D						X	X		
	13130.85	398-411 R.GMPGSPGLPGADGPK.A	1															X				X				X		X		X																
	13387.69	656-669 K.GETGLPGDIGSPGRD.+Hydroxy(P)																																X													
	14427.73	572-586 R.GPRGEPGLPGPAGARGA.+2 Hydroxy(P)			X	X	X	X	X	X	X	X	X	X	X	X				X		X		X		X		X		X		X		X		X						X		X	X		
	14440.72	572-586 R.GPRGEPGLPGPAGARGA.+3 Hydroxy(P)													X																																	
	14457.68	776-792 R.GDDGPPGATGFPGAAGRT.+Hydroxy(P)				X										X								X		X				X																	
	14473.67	776-792 R.GDDGPPGATGFPGAAGRT.+2 Hydroxy(P)				X		X		X		X		X		X				X		X		X		X				X		X				X						X	X				
	14492.69	829-844 R.GGETGAGDPPGPVGEK.G+Hydroxy(P)				X								X					X				X		X																						
	14632.78	977-994 R.GEPGPAGAVGPAGAVGPRGLG.+Hydroxy(P)			X	X	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X	X	X				X		X		X		X						X					
	14632.82	164-178 R.GPPGTPGLPGPVGRGLG.+2 Hydroxy(P)					X		X					X		X																															
	14646.83	308-325 K.GAAGLPGVAGAPGLPGPRGA.+2 Hydroxy(P)					X				X												X																								
	14648.78	977-994 R.GEPGPAGAVGPAGAVGPRGLG.+2 Hydroxy(P)					X				X		X							X		X		X		X								X												
	14648.82	164-178 R.GPPGTPGLPGPVGRGLG.+3 Hydroxy(P)					X				X					X																															
	14660.81	814-828 K.EGLRGPRGSDGQPVGR.S	1																									X		D	X	X-D	D		D											
	14662.83	308-325 K.GAAGLPGVAGAPGLPGPRGA.+3 Hydroxy(P)				X	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X						X	X				
	14664.77	590-607 R.GPPGESGAAGPTGPIGSR.G																																			X										
	14664.77	977-994 R.GEPGPAGAVGPAGAVGPRGLG.+3 Hydroxy(P)																																												
	14664.82	164-178 R.GPPGTPGLPGPVGRGLG.+3 Hydroxy(P, K)																																											X	

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	1580.77	590-607 R.GPPGESGSAAGPTGPIGSR.G + Hydroxy (P)			X	X					X		X						X		X		X						X				X					X	X						
	1615.82	263-280 K.GELGPVGNPGPAGPAGPRG.I + Hydroxy (P)	1		X+D								X+D		X+D						X+D		X+D											X											
	1641.81	404-421 R.GLPGADGPAGIVMGPAGSR.G + Hydroxy (P)																											X		X															
	1648.77	139-154 K.AGEDGHPGKPGPPGERGR.G + 2 Hydroxy (P)													X				X		X				X																		X			
	1648.83	388-397 R.GSTGEIGPAGPPGPPGLR.G + 2 Hydroxy (P)			X		X		X		X		X		X		X						X		X				X												X				
	1657.81	404-421 R.GLPGADGPAGIVMGPAGSR.G + Hydroxy (P) and Oxi (M)																													X															
	1668.86	412-430 R.AGIVMGPAGSRGATGPAGYR.G																																												
	1686.88	644-661 R.GAAGIPGGKGEIGETGLR.G + Hydroxy (P, K)													X								X						X																
	1702.88	644-661 R.GAAGIPGGKGEIGETGLR.G + 3 Hydroxy (P, K)																					X																						
	1777.90	179-196 R.GHNGLDGLKGQPGAPGIKG.H + Hydroxy (P)	2														X				X				X																				
	1753.89	572-589 R.GIPGEPGLPGPAGARGERG.H + Hydroxy (P)			X		X		X				X																X		X															
	1769.89	572-589 R.GIPGEPGLPGPAGARGERG.H + 2 Hydroxy (P)			X		X		X		X						X	X											X		X		X								X				
	1782.82	431-448 R.GPNGDSGRPGEPGLMGPRG.I + 2 Hydroxy (P) or + Hydroxy (P) and Oxi (M)	1		D		D						D		D				D		D		D		D				D				D		D										
	1786.88	658-673 K.GETGLPGDIGSPGPDGAR.G							X																								X													
	1788.81	431-448 R.GPNGDSGRPGEPGLMGPRG.I + 3 Hydroxy (P) or + 2 Hydroxy (P) and Oxi (M)	1		D		D						D		D								D						D						D	D									
	1804.93	379-397 K.RGSTGEIGPAGPPGPPGLR.G + 2 Hydroxy (P)			X		X		X				X		X								X						X				X											
	1813.91	793-810 R.TGPPGPSGSGPPGPPGPAGK.E + Hydroxy (P, K)																											X												X				
	1829.90	793-810 R.TGPPGPSGSGPPGPPGPAGK.E + 2 Hydroxy (P, K)											X						X		X		X		X				X		X		X								X				

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	1645.90	793-803 R.TGPPGPSGSGPPGPPGPAGK.E + 4 Hydroxy(P,																X				X		X						X																	
	1661.89	793-803 R.TGPPGPSGSGPPGPPGPAGK.E + 4 Hydroxy(P,																																										X			
	1906.94	587-607 R.GERGPPGESGAAGPTGPIGSR.G																								X																					
	1922.93	587-607 R.GERGPPGESGAAGPTGPIGSR.G + Hydroxy(P)				X		X		X		X						X		X		X		X				X		X		X		X	X							X	X			
	1975.95	359-379 K.GEPGAVGQPGPPGPSGEEGKR.G + Hydroxy(P)	1																X								X																				
	1987.01	973-994 K.HGNRGEPGPAGAVGPAGAVGPRGPR.G + Hydroxy(P)	1		D								D		D				D		D		D																					X			
	2007.94	359-379 K.GEPGAVGQPGPPGPSGEEGKR.G + 3 Hydroxy(P)	1								X																																				
	2013.00	973-994 K.HGNRGEPGPAGAVGPAGAVGPRGPR.G + 2 Hydroxy	1		D												D		D		D		D						X		X		X		X								X				
	2028.06	155-175 R.GAVGPDGAARGPPGTPGLPGPK.G + 2 Hydroxy(P)	1	X		X		X		X		X																				X		X												
	2044.06	155-175 R.GAVGPDGAARGPPGTPGLPGPK.G + 3 Hydroxy(P)	1	X				X				X																																			
	2050.99	710-736 R.GEVGPAGPMGFAGPAGAAGDPGAK.G + Hydroxy	2										X-D					X-D																													
	2066.09	463-484 K.EGPVGLPGDGGPPGPIGPAGAR.G + Hydroxy(P)		X		X		X		X		X		X				X		X		X		X		X				X		X		X						X		X			
	2066.99	710-736 R.GEVGPAGPMGFAGPAGAAGDPGAK.G + 2	2																						D																						
	2072.08	463-484 K.EGPVGLPGDGGPPGPIGPAGAR.G + 2 Hydroxy(P)		X		X		X		X		X		X				X		X		X		X						X		X										X				
	2115.12	881-904 R.GLPGVAGSVGEPGPLGIAGPPGAR.G + 2 Hydroxy			X						X		X		X																												X		X		
	2131.11	881-904 R.GLPGVAGSVGEPGPLGIAGPPGAR.G + 3 Hydroxy		X		X		X		X		X		X		X			X		X		X		X		X			X		X		X							X		X		
	2233.10	380-403 R.GSTGEIGPAGPPGPPGLRGMPGSR.G + 3 Hydroxy	1																		D		D							D		X		D									X		X		
	2249.09	380-403 R.GSTGEIGPAGPPGPPGLRGMPGSR.G + 4 Hydroxy	1																																										X			

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	2226.07	399-421 R.GMPGSRGLPGADGRAGIVGFPAGSR.+4 Hydroxy(P)	1		X+D-2D			D+2D	X			D+2D		X+D-2D																		
	2261.03	947-972 R.GIYPGNAGPVGAAGAPGPQGPVGPVGKHR.+2 Hydroxy(P)	2			X+D-2D						D					X+D-2D			X+D-2D	X+D		X+D-2D				X	X		
	2285.16	793-817 R.TGPPGPSGSGGPPGPQGAGKEGLRGR.+3 Hydroxy(P)																			X								X		
	2300.15	793-817 R.TGPPGPSGSGGPPGPAGKEGLRGR.+4 Hydroxy(P,K)															X	X													
	2262.14	131-154 R.GPPGPPGKAGEDGHPGKPGPPGERGR.+3 Hydroxy(P)														X				X											
	2286.23	977-1003 R.GEPGPAGAVGPAGAVGPRGPSGPQGIR.G	1																	X											
	2286.25	281-307 R.GEVGLPGLSGPVIGPPGNPGANGLPGAK.G	2																	X											
	2293.16	715-739 R.GEVGPAGPNGFAGPAGAAGQPGAKGERG.+Hydroxy(P)	2				2D		2D																							
	2405.15	715-739 R.GEVGPAGPNGFAGPAGAAGQPGAKGERG.+Hydroxy(P,K)	2	D+2D	D	D	D	D+2D	D+2D		D+2D	D	D+2D	D+2D						2D											
	2410.27	1852-1878 R.GPAGPSGPAGKDGPRGGPGAVGPAGIPGAR.G.+Hydroxy(P)	1													X					X										
	2430.22	281-307 R.GEVGLPGLSGPVIGPPGNPGANGLPGAK.G.+4 Hydroxy(P)	2	D+2D	D+2D	D+2D	D+2D	D+2D	D+2D																							
	2440.09	353-378 K.GESGNKGEPGAVGQPGPPGPSGEEGK.R.+4 Hydroxy(P)	2													X		X		X											
	2440.21	494-520 R.GLPGADGRAGVMGPAGSRGATGPAGVR.G.+3 Hydroxy(P)														X		X													
	2440.27	164-187 R.GPPGTPGLPGFKGPRGMNGLDGLKGQ.+3 Hydroxy(P)	1													X															
	2446.22	281-307 R.GEVGLPGLSGPVIGPPGNPGANGLPGAKGQ.+5 Hydroxy(P,K)	2	D+2D	D+2D	D+2D	D+2D	D-2D	D+2D					2D																		
	2852.18	187-222 K.GEPGAPGENGTPGQTGARGLPGERGR.Y.+3 Hydroxy(P)	2																	X			X								
	2852.21	353-379 K.GESGNKGEPGAVGQPGPPGPSGEEGKR.G	2														X														
	2852.23	422-448 R.GATGPAGVRGPNGDSGRPGEPGLMGPRG.+Hydroxy(P) or Oxi(M)	1													X		X		X											

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	2549.22	422-448 R.GIATGPACVRGPNGDSGRPGEPGLMGPRG + 2 Hydroxy (P) or + Hydroxy (P) and Oxi (M)	1			x		x		x		x		x		x								x		x		x		x				x												
	2572.26	233-262 R.GSDGSVGPVGPAGPIGSAGPPGFPGAPGPKG + 2 Hydroxy (P; K)				x		x		x		x		x																																	
	2578.33	793-820 R.TGPPGPSGSGSPPGPQPAGKEGLRGPRG + 2 Hydroxy (P)														x												x		x				x													
	2598.09	353-379 K.GESGIWGEPGAVGPQPPGPSGEEGKRGLG + 3 Hydroxy (P)	2																			2D																									
	2598.24	746-776 K.GENGPVGFTGPVGAAGPSGPNGPPGPAGSRGLG + Hydroxy (P)	2	2D				2D				2D		2D				2D				D-2D		2D												2D											
	2599.26	233-262 R.GSDGSVGPVGPAGPIGSAGPPGFPGAPGPKG + 3 Hydroxy (P; K)				x				x		x		x								x																									
	2611.32	783-820 R.TGPPGPSGSGSPPGPQPAGKEGLRGPRGLG + 1 Hydroxy (P)																										x																		
	2651.28	944-972 K.GERGVPGNAGPVGAAGAPGPQGPVPGPVGKLH + 5 Hydroxy (P)	2				x																																						
	2666.25	707-736 R.GSPGPEGEVGPAGPNGPAGPAGAAGQPGAKG + 3 Hydroxy (P; K)	2																													2D														
	2673.25	590-619 R.GPPGESGAAAGPTGPIGSRGPGSGPPGPDGNK + 3 Hydroxy (P)	1	x				x						x		x														x		x														
	2689.24	590-619 R.GPPGESGAAAGPTGPIGSRGPGSGPPGPDGNK + 4 Hydroxy (P)	1								x						x		x							x				x																
	2705.23	590-619 R.GPPGESGAAAGPTGPIGSRGPGSGPPGPDGNK + 5 Hydroxy (P)	1								x				x		x									x		x																		
	2788.35	674-706 R.GAPGAAGAPGPAGAMGDRGEAGPAGPAGPAGPRG + 2 Hydroxy (P)	1		D		D		D		D		D		D				D		D		D				D						D		D	D									
	2794.37	8022-8061 K.GHNGLDGLPGLAGSHGDQGAPGAVGPAGPRG + 2 Hydroxy (P)	3				D						D-2D								2D				D																				
	2792.32	542-571 K.GEDGPAGPPGPGGLPGPAGTAGEAGKPGERG + 3 Hydroxy (P; K)	2		x		x		x		x		X-D												D																			x	x		
	2786.34	674-706 R.GAPGAAGAPGPAGAMGDRGEAGPAGPAGPAGPRG + 3 Hydroxy (P)	1																																	x										
	2802.34	674-706 R.GAPGAAGAPGPAGAMGDRGEAGPAGPAGPAGPRG + 4 Hydroxy (P)	1																																	x										
	2879.40	743-775 K.GPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG + 2 Hydroxy (P; K)	2	2D								2D		2D				2D				2D		2D																						

Protein	MH+ (theoretical)	Position in the sequence, Peptide sequence (with flanking residues) + Modification	Deamidation sites	G1		G2		G3		G4		G5		G6		G7		G8		G9		G10		G11		G12		G13		G14		G15		G16		G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29
				MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	MALDI-TOF	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS	LC-MSMS
Collagen alpha2(I) (P02465)	2926.43	973-1003 K.HGMPGEPGPAGAVGPAGAVGPRGPSGPPGQMRG+6 Hydroxy (P)	2						D-2D				X-D-2D		2D				D														D- 2D														
	3001.37	98-130 R.GPPGASGAPGPQGPQGPPGEPGQTGPAGARG+4 Hydroxy (P)	3														D						X-D												X-D	X-D										
	3008.42	707-739 R.GSPGPEGEVGPAGPMGPAGPAGAAGQPGAKGERG+3 Hydroxy (P, K)	2						D-2D						2D						D								D-2D				D													
	3017.36	98-130 R.GPPGASGAPGPQGPQGPPGEPGEPGQTGPAGARG+5 Hydroxy (P)	3																			X													X-D-2D								X			
	3017.50	845-877 K.GPSGEPGTAGPPGTPGPQGLLGAPGFGFLGLPGSRG+4 Hydroxy (P)	1									X-D-2D																																		
	3033.49	845-877 K.GPSGEPGTAGPPGTPGPQGLLGAPGFGFLGLPGSRG+5 Hydroxy (P)	1		X-D		X-D		X		X-D		X-D		X-D				X						X-D																					
	3040.47	388-430 R.GMPGSRGGLPGADGPAGVMGPAGSRGATGPAGVRG+5 Hydroxy (P)	1																																											
	3040.48	814-844 K.EGLRGPPGSDGQPGVKGPSGETGASGPPGFVGEK.G+2 Hydroxy (P)	1																																	X										
	3040.53	405-437 R.GEPGMMGFPGPKGSDGRAGEKGHAGLAGARG	1																																	X										
	3049.35	98-130 R.GPPGASGAPGPQGPQGPPGEPGEPGQTGPAGARG+7 Hydroxy (P)	3																																		X-D								X		
	3049.57	740-775 R.GTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG+ Hydroxy (P)	2																												2D		2D													
	3063.57	740-775 R.GTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRG+ 2 Hydroxy (P)	2																												2D		2D													
	3227.62	608-643 R.GPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERG+ 4 Hydroxy (P)	1		X						D				X		X-D						X-D		X-D										X-D								X	X		
	3243.51	608-643 R.GPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERG+ 5 Hydroxy (P)	1								X																																	X		
	3398.72	1006-1051 R.GLPGLKGHNGLQGLPGLAGHHGDQGAPGAVGPAGPRG+ 4 Hydroxy (P,K)	3														3D				D								3D				D													
	3646.74	341-379 R.GLVGEPGPAGSKGESGNKGEPGAVGPQGPGPPGPSGEEGKR G+4 Hydroxy (P, K)	2																		2D																									
	3990.00	281-325 R.GEVGLPGLSGPVGPPGNPGANGLPGAKGAAGLPGVAGA PGLPGPRG+8 Hydroxy (P)	2		D-2D				D-2D		2D				D-2D														2D				2D													

Table B3: Proteotypic peptides and selected transitions for Collagen alpha 1(I) and Collagen alpha 2(I) as obtained by Skyline software (http://sciex.com/products/software/skyline-software)

PROTEIN	PEPTIDE	PRECURSOR	PRODUCT	COLLISION ENERGY
ALPHA1 (I)	GVVGL**P**GQR	449,7589++	V [y8] - 841,4890+ V [y7] - 742,4206+ G [y6] - 643,3522+ L [y5] - 586,3307+ P [y4] - 473,2467+	16 16 16 16 16
	SGDRGETGPAGPAGPIGPVGAR (NO HYDROXY)	988,5009++	G [y15] - 1273,7011+ P [y14] - 1216,6797+ A [y13] - 1119,6269+ G [y12] - 1048,5898+ P [y11] - 991,5683+	36 36 36 36 36
	GPAGPQGPR (NO HYDROXY)	418,7223++	P [y8] - 779,4159+ A [y7] - 682,3631+ G [y6] - 611,3260+ P [y5] - 554,3045+ Q [y4] - 457,2518+	14 14 14 14 14
ALPHA2 (I)	IGQPGAVGPAGIR (NO HYDROXY) **PEP 1 ALPHA2 (I)**	596,8435++	Q [y11] - 1022,5742+ P [y10] - 894,5156+ G [y9] - 797,4628+ A [y8] - 740,4413+ V [y7] - 669,4042+	21 21 21 21 21
	GI**P**GPVGAAGATGAR **PEP 2 ALPHA2 (I) P1**	634,3413++	G [y12] - 984,5221+ P [y11] - 927,5007+ V [y10] - 830,4479+ G [y9] - 731,3795+ A [y8] - 674,3580+	22 22 22 22 22
	GIPG**P**VGAAGATGAR **PEP 2 ALPHA2 (I) P2**	634,3413++	G [y12] - 1000,5170+ P [y11] - 943,4956+ V [y10] - 830,4479+ G [y9] - 731,3795+ A [y8] - 674,3580+	22 22 22 22 22
	GI**P**GEFGL**P**GPAGAR **PEP 3 ALPHA2 (I) P1-P2**	714,3675++	P [y13] - 1257,6222+ G [y12] - 1144,5745+ E [y11] - 1087,5531+ F [y10] - 958,5105+ G [y9] - 811,4421+	25 25 25

			L [y8] - 754,4206+ P [y7] - 641,3365+ G [y6] - 528,2889+ P [y5] - 471,2674+ A [y4] - 374,2146+ G [y3] - 303,1775+	25 25 25 25 25 25 25 25
	GI**P**GEFGLPG**P**AGAR **PEP 3 ALPHA2 (I) P1-P3**	714,3675++	P [y13] - 1257,6222+ G [y12] - 1144,5745+ E [y11] - 1087,5531+ F [y10] - 958,5105+ G [y9] - 811,4421+ L [y8] - 754,4206+ P [y7] - 641,3365+ G [y6] - 544,2838+ P [y5] - 487,2623+ A [y4] - 374,2146+ G [y3] - 303,1775+	25 25 25 25 25 25 25 25 25 25 25 25
	GIPGEFGL**P**G**P**AGAR **PEP 3 ALPHA2 (I) P2-P3**	714,3675++	P [y13] - 1257,6222+ G [y12] - 1160,5695+ E [y11] - 1103,5480+ F [y10] - 974,5054+ G [y9] - 827,4370+ L [y8] - 770,4155+ P [y7] - 657,3315+ G [y6] - 544,2838+ P [y5] - 487,2623+ A [y4] - 374,2146+ G [y3] - 303,1775+	25 25 25 25 25 25 25 25 25 25 25 25

Figure B1: Graphical representation of the proteins that have been identified in 29 animal glue containing samples. The occurrence has been reported as percentage of the analysed samples. On the side the actual number of samples in which the protein has been identified is also reported. The UniProt code is reported in brackets.

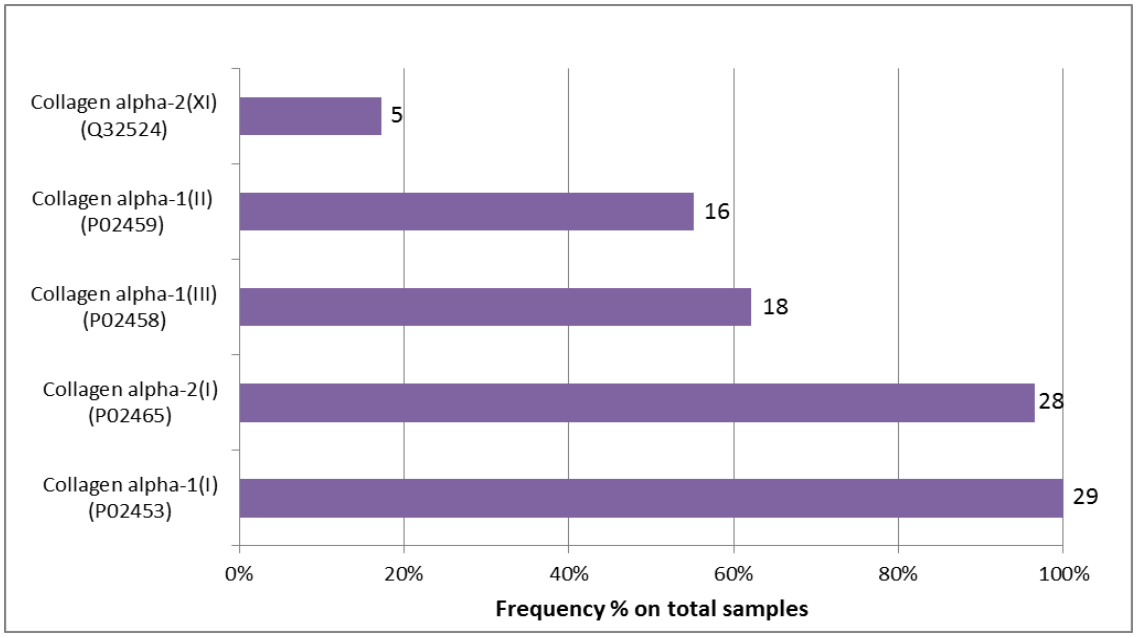


Figure B 2: Graphical representation of the proteins that have been identified in43 milk containing samples. The occurrence has been reported as percentage of the analysed samples. On the side the actual number of samples in which the protein has been identified is also reported. The UniProt code is reported in brackets.

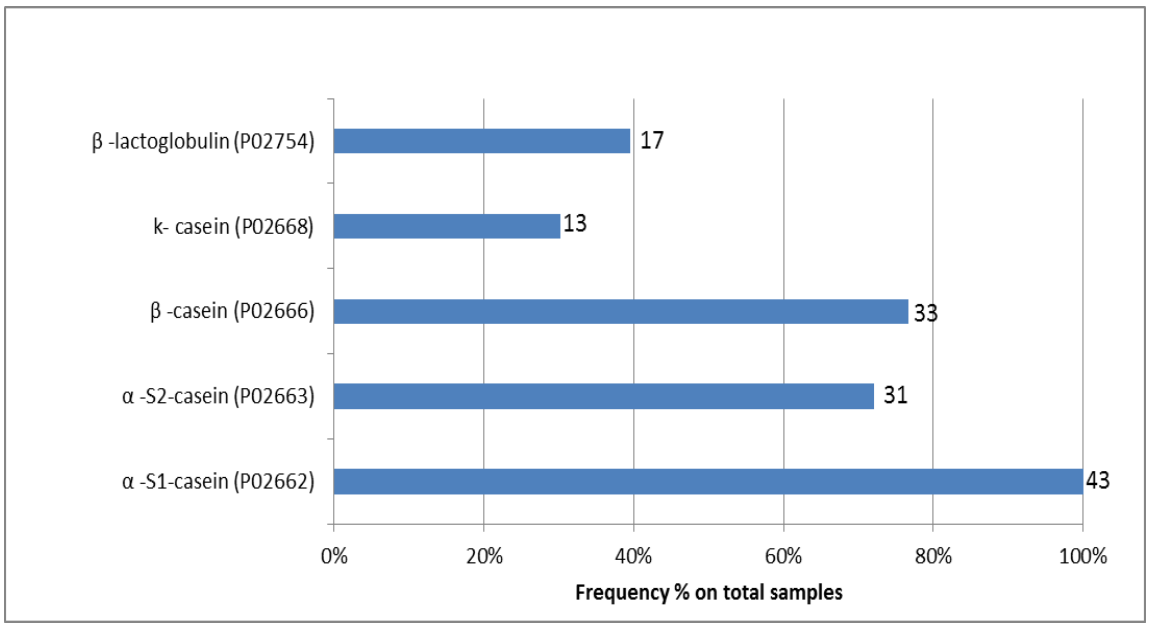


Figure B3: Graphical representation of the proteins that have been identified in43eggcontaining samples (27 samples albumen only, and 16 samples albumen plus yolk). The occurrence has been reported as percentage of the analysed samples. On the side the actual number of samples in which the protein has been identified is also reported The UniProt code is reported in brackets.

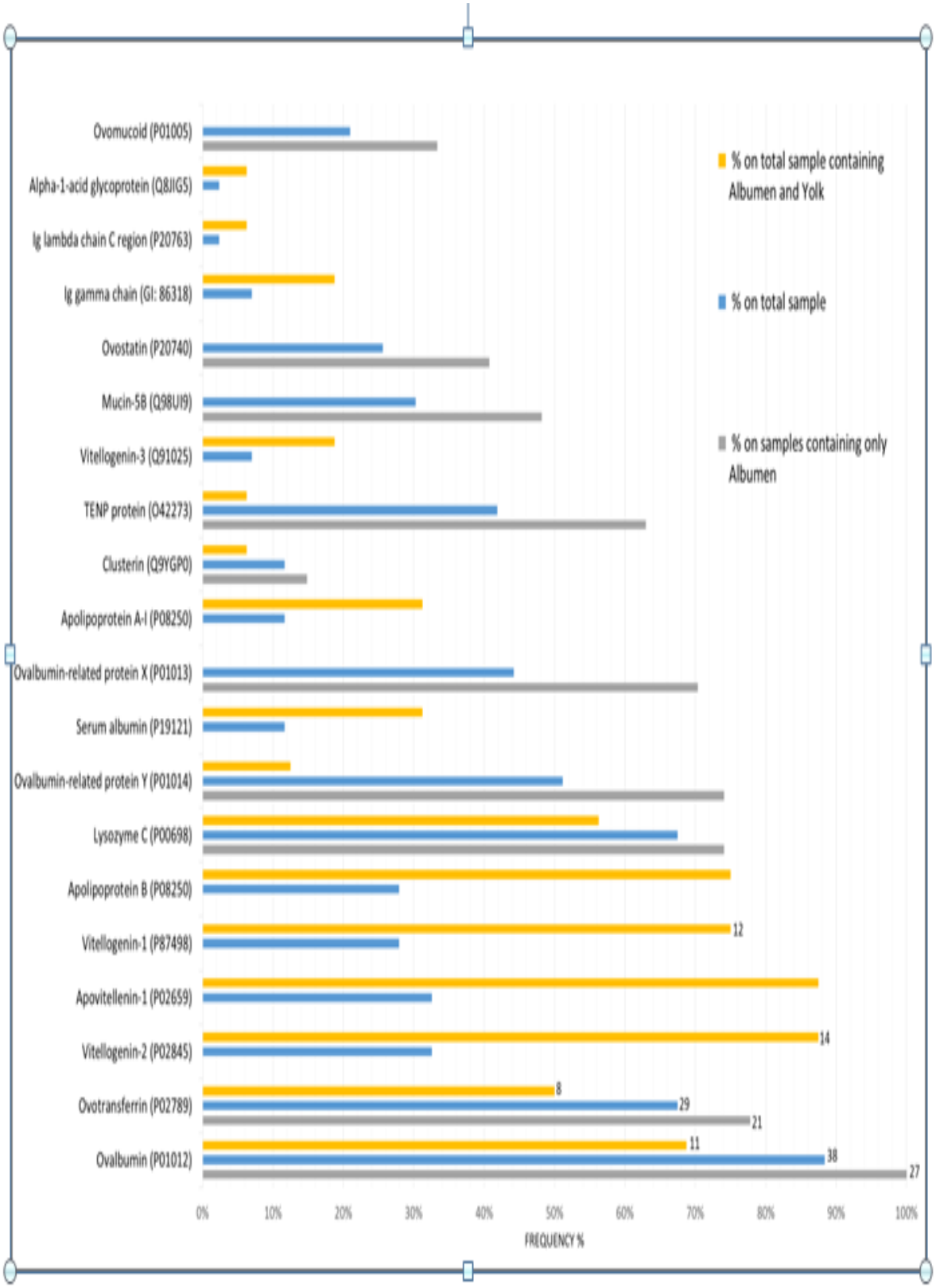


Figure B4. Observed frequency percentage for peptides of Collagen alpha 1(I) in the animal glue containing samples.

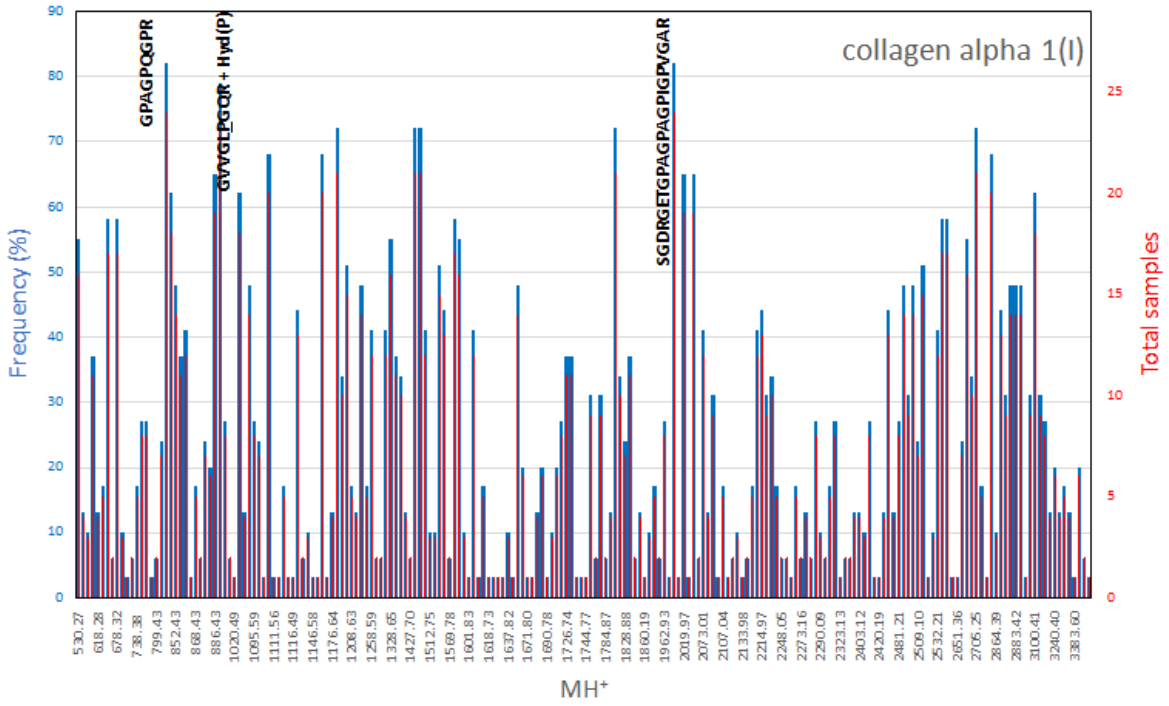
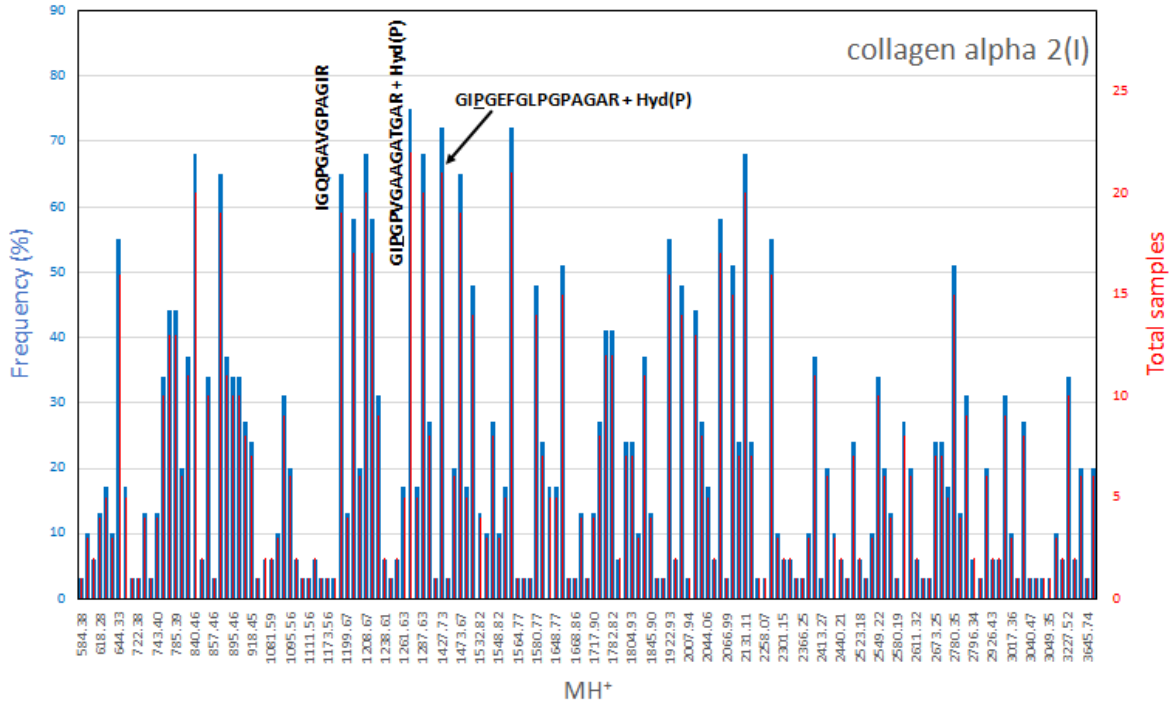


Figure B5. Observed frequency percentage for peptides of Collagen alpha 2(I) in the animal glue containing samples.



PhD Course Activity Summary

Candidate:Dr.Addolorata De Chiaro **Supervisor: Prof. Leila Birolo**

1) **Attended Courses:**

- Chimica fisica dei nanosistemi, 8 hours, April 2014n(Pof. Paduano)
- Tecniche di estrazione solido-liquido, 8 hours, June 2014 (Prof. Naviglio)
- Spettrometria di Massa, 10 hours, July 2014 (Prof. Pucci)
- Microscopia Raman, 8 hours, December 2014 (Prof. Vergara)
- Sintesi, struttura ed applicazioni di oligonucleotide naturali e modificati, 8 hours, February 2015 (Prof. Montesarchio)
- Glicoscienza, 8 hours, July 2015 (Prof.Parrilli and Dr.Bedini)
- Produzione ricombinante di protein naturali e mutanti, 8 hours, July 2015 (Prof. Duilio)
- Neutron scattering, 8 hours, October 2015 (Prof. Zorn)

2) **Attended Seminars:**

Title	Speaker	Date	Place
Studio delle basi molecolari dell' acidemia metilmalonica	Dr. Marianna Caterino	6/06/2014	CEINGE BIOTECNOLOGIE AVANZATE
Evidenze genetiche sulle prime espansioni umane dall'Africa	Prof. Guido Barbujani	23/10/2014	Università degli Studi di Napoli "Federico II"
Antibodies and Mass Spectrometry…collaborators or competitors?	Prof. Pietro Pucci	12/12/2014	CEINGE BIOTECNOLOGIE AVANZATE
Pharmaceutical Companies: external manufacturing and quality assurance.	Prof.Domenico De Masi	25/05/2015	Università degli Studi di Napoli "Federico II"

Alessandro Ballio e la chimica a Napoli negli anni '60	Dr. Andrea Carpentieri	18/06/2015	Università degli Studi di Napoli "Federico II"
Le transglutaminasi..dalle poliammine alle bioplastiche	Prof. Raffaele Porta	30/06/2015	Università degli Studi di Napoli "Federico II"
Introduction to bioinformatics	Dr. Remo Sanges	18/12/2015	Università degli Studi di Napoli "Federico II"
Biopesticides which target voltage-gated ion channels: efficacy and biosafety	Prof.AngharadM.R.Gatehouse	14/01/2016	Università degli Studi di Napoli "Federico II"
1,1,2,2-tetraethoxybutyne-Exciting to make and work with	Prof.Leiv K. Sydnes	14/01/2016	Università degli Studi di Napoli "Federico II"
Functionalized and artificial enzymes:newbio-derived catalysts	Prof.Therrytron	14/01/2016	Università degli Studi di Napoli "Federico II"
Secondary metabolites from higher land and sea plants in Chemical ecology and Chemosystematics	Prof. Christian Zidorn	27/01/2016	Università degli Studi di Napoli "Federico II"

3) Attended Integration Exams (for candidates not graduated in Chemical Science):

Title	Professor	Date
ChimicaFisica I	**Castronuovo**	**11/02/2015**
Chimica Inorganica	**Vitagliano**	**29/02/2016**

4) Visiting periods in Institutions different from University of Naples "Federico II":

Host Institution	Country	Start Date	End Date

5) Publications (include submitted and in preparation):

- Vinciguerra R., **De ChiaroA,** Pucci P, Marino G, Birolo L.Proteomic strategies for cultural heritage: From bones to paintings. Microchemical Journal, 126(2016) 341-348.
- Ventorino V, Ionata E, Birolo L, Montella S, Marcolongo L, **De Chiaro A**, Espresso F, Faraco V, Pepe O. Lignocellulose-adapted endo-cellulase producing Streptomyces strains for bioconversion of cellulose-based materials. Front Microbiol. 22 (2016) 1-15
- Lluveras A, Bonaduce I, **De Chiaro A**, Birolo L, Maria Perla Colombini MP. Identification of Asian materials used as binders. Work in preparation.
- Vinciguerra R, **De Chiaro A**, Illiano, A Vallone F, Bonaduce I, Amoresano A, Birolo L. A targeted proteomic approach for Cultural Heritage: identification of proteins in Paintings. Work in preparation.
- Liguori R, **De Chiaro A**, Birolo L, Faraco V. Screening of fungal strains for cellulolytic, xylanolytic and ligninolytic activities production and identification of the enzymes responsible for these activities and evaluation of Brewers'Spent Grain for both enzyme and sugar production. Work in preparation.

6) Attended congresses/workshops/summer schools/contribution:

- *Proteomicsstrategies for the analysis of ancientbones***Addolorata De Chiaro** , Mauro Finicelli Alessandra Cirillo, Roberto Vinciguerra, Marilena Cipollaro , Piero Pucci, Gianfranco Peluso, Leila BiroloItalianProteomicassociation IX congresso. Giugno 2014 –Napoli, Italia
- *Proreomics and cultural heritage:frombones*Roberto Vinciguerra, **Addolorata De Chiaro**, Piero Pucci, Gennaro Marino, Leila BiroloTechnart 2015. Aprile 2015- Catania, Italia
- *WhatProteomics can do for cultural heritage:frombones to paintings* Roberto Vinciguerra, **Addolorata De Chiaro**, Piero Pucci, Gennaro Marino, Leila Birolo1O Workshop Beni culturali e tecnologia L.O.S.A.I. Maggio 2015 –Napoli, Italia
- *Proteomics for enzyme discovery in industrial biotechnology***Addolorata De Chiaro**, Roberto Vinciguerra, Angela Amoresano, Leila BiroloTraining Course on Emerging Biotechnologies for Sustainable Waste Management and Biorefinery DevelopmentAprile 2016 –Napoli, Italia