UNIVERSITY OF PARDUBICE
FACULTY OF CHEMICAL TECHNOLOGY
Department of Analytical Chemistry

Marine Morvan

Proteomics Analysis of Aging Proteins

Theses of the Doctoral Dissertation

Pardubice 2023

Study program: **Analytical Chemistry**

Study field: **Analytical Chemistry**

Author: **M.Sc. Marine Morvan**

Supervisor: **Dr. Petr Česla**

Year of defense: **2023**

References

MORVAN, Marine. Proteomics analysis of ancient proteins. Pardubice, 2023, 192 pages. Dissertation thesis (Ph.D.). University of Pardubice, Faculty of Chemical Technology, Department of Analytical Chemistry. Supervisor Dr. Petr Česla.

Abstract

The analysis and characterization of aging and archaeological proteins are among the latest challenges in analytical chemistry. In the first part of this dissertation thesis, aging proteins were studied using LC-MS. The effects of aging on protein sequences, including amino acid racemization, post-translational modifications, and protein degradation, were studied. Subsequently, a chiral separation method was developer to determine the amino acid enantiomer rates. In the second part of this dissertation thesis, archaeological proteins were studied using nanoLC-MS. Proteomics, called paleoproteomics in this case, was developed as an alternative method for osteoarchaeology and genomics. Based on two sex-dependent forms of amelogenin protein preserved in teeth, both biological sexes were distinguished by nanoLC-MS because of differences in their protein sequences. The developed proteomic approach was designed to be minimally-invasive. This was confirmed by scanning the teeth before and after amelogenin extraction using both scanning electron microscope and micro-computer tomography.

Abstrakt

Analýza a charakterizace stárnutí a archeologických proteinů patří mezi nejnovější výzvy v analytické chemii. V první části této disertační práce byly studovány stárnoucí proteiny pomocí LC-MS. Byly studovány účinky stárnutí na proteinové sekvence, včetně racemizace aminokyselin, posttranslačních modifikací a degradace proteinů. Byla vyvinuta metoda chirální separace pro stanovení poměrů aminokyselinových enantiomerů. V druhé části práce byly archeologické proteiny studovány pomocí nanoLC-MS. Proteomika, v tomto případě zvaná paleoproteomika, byla vyvinuta jako alternativní metoda k osteoarcheologii a genomice. Na základě dvou forem proteinu amelogenin, který je pohlavně rozdílný, konzerovovaných v zubech, byly obě pohlaví rozlišeny metodou nanoLC-MS na základě rozdílů v sekvencích. Vyvinutý proteomický přístup byl navržen tak, aby byl minimálně invazivní. To bylo potvrzeno skenováním zubů před a po extrakci amelogeninu pomocí skenovacího elektronového mikroskopu a mikropočítačové tomografie.

Keywords

aging proteomics, amelogenin, amino acids, chiral separation, collagen, LC-MS, paleoproteomics, post-translational modifications, protein degradation

Klíčová slova

amelogenin, aminokyseliny, chirální separace, degradace proteinů, kolagen, LC-MS, paleoproteomika, posttranslační modifikace, proteomika stárnutí

Contents

Chapter 1

Introduction

1.1 Protein structures and functions

1.1.1 Proteins and their structures

Protein structure is described in four levels. The primary structure is defined by the distinctive amino acid sequence in the polypeptide chain. The secondary structure is the geometric arrangement of the polypeptide chain between several consecutive amino acids and is conditioned by the formation of hydrogen bonds between the amino and carbonyl groups of the peptide bond. The most common secondary structures include the α-helix and the β-sheet. The tertiary structure refers to the three-dimensional arrangement of the entire peptide chain. And the quaternary structure gives information on the arrangement of proteins, which are formed by two or more polypeptide chains (Figure 1.1). Proteins are classified into three families regardless of their structure *i.e.* fibrous as collagen, globular as amelogenin, and derived proteins.

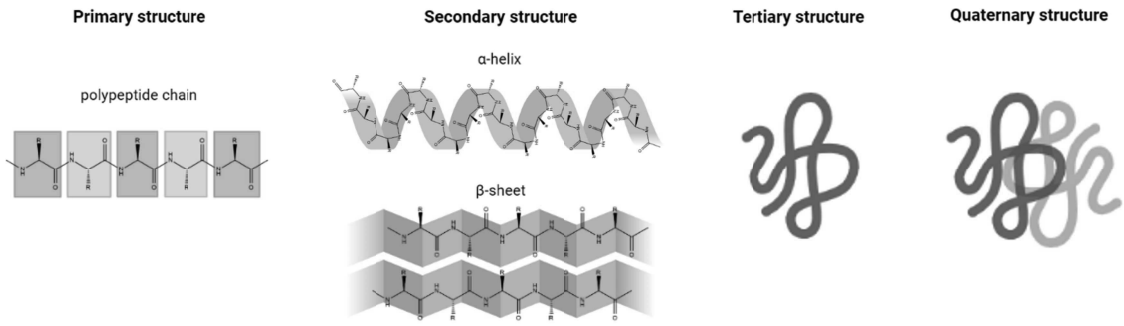

FIGURE 1.1: Structure of proteins
Created with BioRender.com

1.1.2 Proteins and their functions

Proteins are important biomolecules and can have different biological functions for proper body functioning. Most of them are contractile or mobile, catalytic, defense, nutrient, regulatory, storage, structural, transport, and toxin. *In vivo*, the protein turnover is unique for each protein. Protein turnover is defined by the *in vivo* balance between synthesis and degradation to maintain protein homeostasis. Due to their unique turnover, proteins have different half-lives, ranging from a few minutes to a century and can be classified as short-lived proteins (SLPs) or long-lived proteins (LLPs).

1.1.3 Proteins and their modifications

Post-translational modifications (PTMs) are the main change on the amino acid side chain and alter protein structure and function. *In vivo*, two mechanisms for the formation of PTMs coexist: enzymatic (ePTMs) and non-enzymatic (nPTMs) post-translational modifications [1]. First, ePTMs are the most abundant and generally appear after protein biosynthesis. Acetylation and methylation are among the most studied of them (Figure 1.2). Second, nPTMs can occur when a nucleophilic or redox-sensitive amino acid side chain spontaneously reacts with an electrophilic metabolite [1, 2]. Among the most studied, there are deamidation, formylation, oxidation, and sulfation (Figure 1.2). Finally, some PTMs can be generated with both processes, such as phosphorylation and hydroxylation, for example (Figure 1.2). Due to their spontaneity, nPTMs can be reversible or irreversible. Irreversible nPTMs, like carbonylation, glycation, and succination, can be produced by excessive oxidative and metabolic stress and are associated with age-related diseases, cancers, and diabetes [2]. This is also the case for the deamidation of asparagine and glutamine residues as an important nPTMS, in addition to releases and causes toxic ammonia accumulation in cells [3], causes the loss of activity and age-related alterations in proteins [4]. PTMs are also linked to protein turnover. In fact, the presence or absence of PTMs at global or specific-sites in the proteinogenic sequence could influence protein half-lives and turnovers [5]. This is the case of phosphorylation in proline residues and acetylation in proline and lysine (located in α-helix and β-sheet) residues could slow down the turnover compared to their non-modified counterparts [5]. In contrast, ubiquitination could accelerate protein

turnover [5].

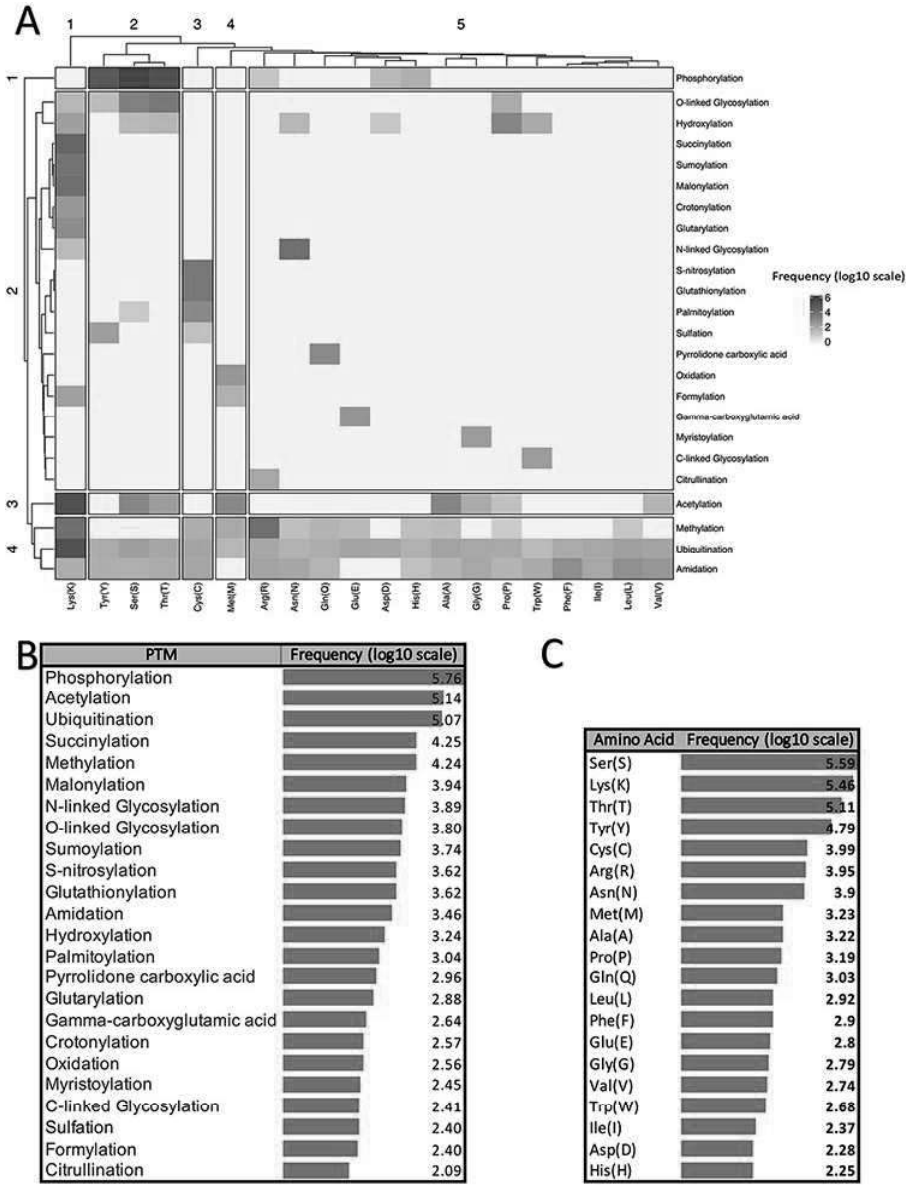

FIGURE 1.2: Major post-translational modifications reported

(A) Clustergram indicating the frequency of each PTM on different amino acids (B) Frequency of major PTMs (C) Frequency of each amino acid that was reported as a modified site. All frequencies are shown in log scale. Last update: October 2020. Reproduced with permission from [6]

1.2 Amino acids structures and functions

1.2.1 Amino acids and their structures

Amino acids consist of an amino group, an alpha carbon, and a carbonyl group for the common skeleton, and a different side chain (R) for each amino acid. Their stereochemistry is defined by the alpha carbon. The two enantiomers are identified in their L- and D-forms (Figure 1.3). These enantiomers have the same physicochemical properties, but their biological functions can be different.

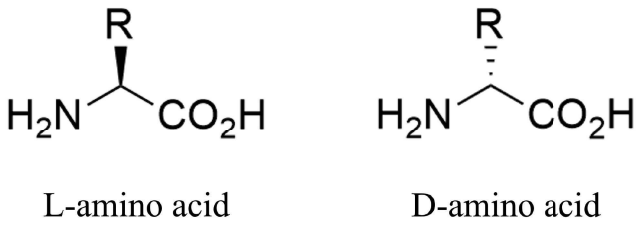

FIGURE 1.3: Structure of amino acid enantiomers

1.2.2 Amino acids and their functions

In nature, there are 20 different unmodified amino acids. Primarily found in their L-form in recent proteins, they are classed as essential, non-essential, or conditional. Essential amino acids must be taken in through food, while non-essential amino acids can be made by the body. Conditional amino acids are beneficial under certain circumstances, such as stress and illness. D-amino acids, as far as they are concerned, are not found in recent proteins. However, they can appear during aging and be implicated in aging dysfunctions and diseases [7]. Additionally, modified amino acids can also appear in proteins. They are formed by PTMs during protein synthesis. This is the case with the carboxylation of glutamate [8, 9] and the hydroxylation of proline and lysine [10, 11], as examples. Other modified amino acids are not proteinogenic, but play a crucial physiological role *i.e.* γ-aminobutyric acid as a neurotransmitter [12], and ornithine and citrulline in the urea cycle [13], as examples.

1.3 Analysis of ancient materials

1.3.1 Aging materials

The analysis of aging materials is at the interface of chemistry and biology. The complex structure of native proteins is often studied in molecular and structural biology, but rarely during the aging process. However, this study would allow us to better understand the *in vivo* aging mechanism of proteins and its molecular and biological consequences. The use of protein structural analysis methods, including cryogenic-electron microscopy, mass spectrometry, NMR spectroscopy, Raman spectroscopy, and X-ray crystallography, developed for recent proteins, can be fully transposed to the analysis of aging proteins. Nonetheless, new challenges may appear for the analysis of aging proteins, such as the change of solubility, the degradation of the sequence, and some sequence modifications such as amino acid racemization and post-translational modifications.

1.3.2 Archaeological materials

The analysis of archaeological materials is at the interface of chemistry, archaeology, and biology. Archaeological samples are rare and valuable, which is why few of them are available for biological or chemical analysis, as most of them are destructive methods. Biological analysis, such as genomics, is the most popular technique used on ancient samples. However, this analysis requires a large number of samples due to its degradation. Nevertheless, chemical analysis, such as proteomics, is an emerging approach and can be a complementary analysis method. The outgoing challenge in paleoproteomics is to be a less sample-consuming method while providing as much information as possible.

Chapter 2

Aims and objectives

The main aim of this dissertation thesis is the development of promising analytical techniques for proteomics applied to aging proteomics and paleoproteomics. This dissertation thesis is divided into four parts.

In the first part, recent advances in the chiral separation of amino acids will be described for the analysis of proteins and peptides. Knowledge surrounding the chirality of peptides and proteins during aging is lacking. The development of new chiral separation methods is essential for the study of the racemization of proteinogenic amino acids, which is an important effect of aging and is linked to some aging diseases. In this work, the performance of chromatographic and electrophoretic techniques for amino acid enantioseparation will be summarized. The different approaches to the derivatization of amino acids will also be summarized.

Then, in the second part, the analysis of aging collagens will be presented. In this work, the protein hydrolysis method plays a crucial role in determining the amino acid enantiomer rate in aging collagens. Indeed, protein hydrolysis conditions that do not influence natural amino acid racemization must be developed. Afterward, the percentage of amino acids in their D-forms, as well as their exact positions in the collagen sequence, will be able to be elucidated. This new analytical method will be applied to different organisms at different ages. Nevertheless, some troubles can be appearing during aging like the change of physiological and physicochemical properties, such as protein solubility, post-translational modifications, and sequence degradations. Taking into account these changes, this complete study will help us better understand the effects of aging on collagens.

Next, in the third part, a comparative review of osteoarchaeology, genomics, and proteomics approaches for sex estimation of ancient skeletons will be included. Based on sexual dimorphism, these three approaches can distinguish both sexes. In osteoarchaeology, morphological differences in the skeleton are determined by visual and metric methods. In genomics, DNA analysis allows the discriminating of both forms of the amelogenin gene into the sex chromosomes X and Y. In proteomics, the analysis of both forms of amelogenin protein encoded by both forms of amelogenin gene allows for determining the sexes according to their different proteogenic sequences. The efficiency of these three multidisciplinary methods, as well as their limitations in terms of the exploitability and consumption of samples, will be evaluated.

Finally, in the last part, a minimally-invasive paleoproteomics method for sex estimation will be described and applied to recent and ancient materials. Indeed, due to the rare and valuable aspects of archaeological materials, an efficient and less sample-consuming analysis method had to be developed. This complementary method will allow us to obtain a precise sex estimation when it was impossible by other methods i.e. osteoarchaeology and genomics. The minimally-invasive character of this method will be evaluated by scanning electron microscope and micro-computed tomography.

Chapter 3

Recent advances in chiral separation

This part of the dissertation thesis summarizes recent advances in chiral separation of amino acid enantiomers. Chiral separation is the last and most important challenge in analytical chemistry. Recent improvements in different separation techniques now allow precise detection and quantification of D-amino acids in complex biological materials. In addition, current improvements in derivatization chemistry contribute to the accurate detection of traces of free and proteinogenic amino acid enantiomers in different biological matrices. This chapter describes recent chromatographic and electrophoretic techniques coupled to mass spectrometry, and several derivatization reagents recently used for the enantioseparation of amino acid enantiomers.

First, liquid chromatography is the most widely used and developed technique. Indeed, several chiral columns are commercialized. These columns are composed of different types of chiral selectors, such as crown ethers, cyclodextrins, cyclofructans, ion exchange, macrocyclic glycopeptides, Pirkle type, polysaccharides, porous organic materials, and proteins [14]. These chiral selectors are linked to the surface of the stationary phase with a chemical spacer and have a different affinity with both enantiomers allowing a chiral separation.

Second, gas chromatography is not the most popular technique for the separation of amino acid enantiomers to date. As in liquid chromatography, stationary phases are functionalized with the chiral selector.

Third, capillary electrophoresis is an emerging technique and is orthogonal to liquid chromatography in terms of its separation mechanism. Indeed, the separation by capillary electrophoresis is performed according to the electrophoretic mobility of the analyte under applied voltage. For amino acid enantioseparation, different chiral selectors were recently used, such as crown ethers, cyclodextrins, and ligand exchanges. Two methods can be designed using these chiral selectors. The first and simplest approach is to add the chiral selector to the background electrolyte (BGE) as a pseudo-phase. The second and most advanced approach consists of the use of these chiral selectors as a dual-ligand. Dual-ligand was created in a combination of an immobilized chiral selector on the capillary surface, as a capillary coating, and the addition of a free chiral selector in the BGE, as a pseudo-phase.

Finally, the derivatization reaction allows the alkylation of a pure chiral reagent to the racemic compounds of interest to form a pair of diastereomers. This technique facilitates the isolation, separation, and detection of derivative analytes from biological matrices. The derivatization reaction takes place mainly on the amino group common to all amino acids, as *N*-alkylation, which makes it possible to analyze all of them simultaneously. In this case, particular attention should be paid to the amino group on the side chain of the lysine residue, which also undergoes the derivatization reaction. Besides, other functional groups can also be derivatized for a more selective analysis. In fact, the thiol group of the cysteine residue can be derivatized by a through *S*-alkylation. Some derivatization reagents are commercial; however, others can be synthesized to meet a precise requirement.

Chapter 4

Analysis of aging proteins

4.1 Aging proteins

Aging can cause different changes in protein sequences, such as loss of proteolytic capacity [15–17], increased surface hydrophobicity [15], the appearance of post-translational modifications such as oxidation [15, 18, 19], phosphorylation [18], methylation [18, 20], deamination [19, 21], and acylation (particularly acetylation [18], carbonylation [19], carboxymethylation [19]) [15], and racemization of amino acids [22–27]. These changes can alter the three-dimensional structure, and the biological activity of proteins, as well as induce dysfunctions and diseases.

4.1.1 Amino acid racemization

In vivo, natural racemization can occur during aging *via* an enzymatic or non-enzymatic process. First, *via* an enzymatic process, the amino acid isomerization is possible by an amino acid racemase. These racemases are classified into two sub-families: pyridoxal 5'-phosphate-dependent (AlaR, ArgR, AspR, HisR, LysR, and SerR as examples) and pyridoxal 5'-phosphate-independent (AspR, GluR, and ProR as examples) [7, 28–31]. These racemases can proceed to free amino acid isomerization before or during peptide elongation [32]. Second, *via* a non-enzymatic process, the amino acids isomerization is possible by a succinimidyl intermediate to form an intramolecular cyclization [33]. This racemization appears in proteinogenic amino acids. Both processes enrich proteins with D-amino acids, and the amount of this D-enantiomer found in aging proteins in healthy patients is progressive with age. However, in diseased patients, the percentage of D-amino acids is also progressive,

but in a higher proportion [23]. The most studied of them is D-aspartic acid in proteins and peptides located in the human body, such as the aorta and skin (elastin), brain (β-amyloid), and lens (α-crystallin). The presence of D-aspartic acid in these proteins and peptides is associated with arteriosclerosis, Alzheimer's disease, and cataracts. Their exact position in the sequence was recently detailed [7]. Other D-amino acids were also identified and located in various animals [7]. To determine the exact percentage of D-amino acids in aging proteins, hydrolysis conditions play a crucial role. Indeed, in HCl/H_2O conditions, a natural racemization of L-amino acids to their enantiomer can occur. During hydrolysis in a hydrogen environment, different racemization kinetics were observed according to the nature of the amino acids [34, 35]. To prevent this amino acids racemization, a DCl/D_2O condition is privileged [7, 36–38]. Indeed, under deuterated hydrolysis, the hydrogen on the alpha carbon was exchanged with a deuterium atom, considerably decreasing the racemization. This is why this hydrolysis method is preferred for age estimation to significantly reduce the age estimation error based on the percentage of D-amino acids [36].

4.1.2 Post-translational modifications

Post-translational modifications are biochemical modifications that occur on the side chain of amino acids. These modifications can be the appearance or disappearance of hydrophilic and/or hydrophobic groups during aging. Oxidation and dioxidation are the most important hydrophilic post-translational modifications. *In vivo*, age-related protein oxidation can occur on the amino acid skeleton by cleaving peptide bonds, and on their side chains [15]. These damages and modifications are performed by reactive oxygen species directly or by sub-products from sugar and lipid oxidations. Certain amino acids are more favorable to age-related oxidations, such as arginine, cysteine, glutamic acid, histidine, leucine, lysine, methionine, phenylalanine, proline, threonine, tryptophan, tyrosine, and valine [15]. Another hydrophilic post-translational modification commonly found in aging proteins is the deamidation of asparagine and glutamine. Aging and ancient proteins are particularly rich in deamidations [19, 21]. Phosphorylation and sulfonation on arginine, aspartic acid, cysteine, histidine, lysine, serine, threonine, and tyrosine are further effects of age-related and hydrophilic post-translational modifications. However,

special attention should be paid to the nature of the sample preparation buffer, such as phosphate and sulfate buffers, which can influence the amount of phospho and sulfo post-translational modifications. On the other hand, the number of carbonyl groups as hydrophobic post-translational modifications on lysine, serine, and threonine, such as acetyl, carbanyl, carboxy, carboxymethyl, carboxyethyl, and formyl, also increases almost exponentially with age [15]. The addition of these hydrophilic groups can increase the hydrophobicity of the protein surface.

4.1.3 Protein surface hydrophobicity

The change in protein surface hydrophobicity is correlated with the appearance of hydrophobic post-translational modifications and the disappearance of hydrophilic post-translational modifications [15]. The main consequences of this phenomenon are the reduction of the solubility of proteins in several solvents and their resistance to enzymatic treatments. Indeed, Miller *et al.* demonstrate that D-peptides are minimally or not at all cleaved by different commonly used enzymes, such as carboxypeptidase A, chymotrypsin, elastase, papain, pepsin, and trypsin [39]. This last point induces the use of successive enzymatic treatments. However, some enzymes recognized the D-amino acids, like the D-aspartyl endoproteinase.

4.2 Collagen

To date, 28 different types of collagen have been identified and numbered with Roman numerals (I – XXVIII) [40]. In animal bodies, including humans, collagen accounts for more than 25% of the total protein mass and is the main structural component of the extracellular matrix (ECM) present in tissues and organs [40, 41].

4.2.1 General structure of collagen

The three-dimensional structure common to all types of collagen is in the triple-helix form made up of polypeptide chains α, interspersed with non-triple helical domains (N- and C-terminal domains). These three α chains can be identical (homotrimers) or different (heterotrimers) [40]. The most common collagen is type I with two heterotrimer chains α 1 and 2 (Figure 4.1) that represents 90% of total collagens [41]. Collagen, as a fibrous protein, is one of the longest proteins and is useful for tissue formation. For example, the human collagen COL1A1 and COL1A2 sequences are composed of 1464 and 1366 amino acids, respectively. Due to this long proteinogenic sequence, collagen protein is composed of a large number of possible cross-linkings. These cross-linking are divided into two categories, *i.e.* intramolecular cross-linkings that stabilize the three-dimensional structure and intermolecular cross-linkings between two collagen proteins while forming fibers. During aging, the amount of collagen cross-linkings increases, leading to an increase in protein rigidity and a decrease in fiber tortuosity [42]. In addition, the presence of D-amino acids in aging collagen [43–46] impacts the design of the three-dimensional structure of the protein. In fact, D-amino acids reverse the rotation of the triple-helix which affects the relative orientation of the amino acid side chain and decreases the helix-helix interactions until the overall destabilization of the three-dimensional structure [47, 48].

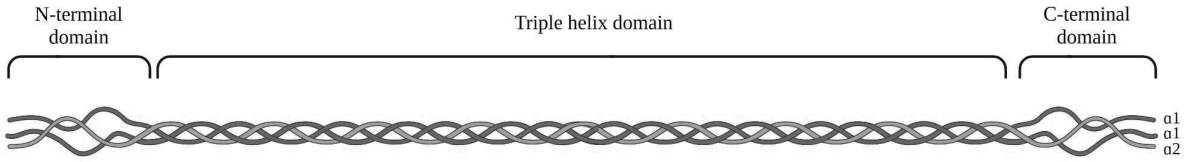

FIGURE 4.1: Representation of the collagen type I structure
Created with BioRender.com

4.2.2 Function of collagen

Collagens play a structural role and contribute to the mechanical properties, organization, and shape of tissues. They interact with cells *via* several receptor families and regulate their differentiation, migration, and proliferation. Some collagen has a restricted tissue distribution and hence specific biological functions [40].

4.3 Study of aging collagen

This part of the dissertation thesis presents the complete study of chiral and aging proteomics applied to collagens. Amino acid racemization, protein sequence degradation, and the evolution of post-translational modifications that occur during aging were studied on bovine and rat collagens. These results have shown for the first time the exact position of amino acids totally racemized in their D-form and the exact position of post-translational modifications. Regarding protein sequence degradation, a fifth of the sequence information was lost during aging. All results and conclusions are summarized in the following publication.

Chapter 5

Recent advances in sex estimation

This part of the dissertation thesis summarizes recent advances in the sex estimation of ancient skeletons for archaeological, anthropological, and forensic research. In forensic research, the study of head and neck bones, and teeth emerges due to their resistance to high temperatures [49, 50]. Sex estimation is fundamental for the characterization of ancient materials, considering that it is the first step in human identification before determining ancestry, height, and age [51, 52]. This estimate can be performed using three approaches based on sexual dimorphism: osteoarchaeology, genomics, and proteomics [53]. Osteoarchaeology and genomics are the two traditional methods, although they have limitations. Proteomics, which is called paleoproteomics in this case, presents itself as a new, more reliable, and less restrictive technique. This chapter describes these different methods for estimating sex and their advantages and limitations for the analysis of rare and precious archaeological samples.

Proteomics analysis is designed to analyze both forms of amelogenin protein encoded by both forms of the amelogenin gene. These two sexually distinct forms of the amelogenin protein are named AMELX and AMELY protein, respectively, from chromosomes X and Y. Like the amelogenin gene, the amelogenin protein differentiates its two forms with different amino acid sequences and lengths. Indeed, the AMELX protein (AMELX-2 isoform) is made up of 175 amino acids, while 192 amino acids for AMELY (AMELY-1 isoform) (Figure 5.1). Their amino acid sequence homology is 93 %. Major differences are the loss of a part of the proteinogenic sequence from position 19 to 34 in the AMELX-2 protein and the loss of the methionine residue at position 45. Minor differences appear with the exchange of 22 amino acid residues by others. Like the amelogenin gene, the encoded protein has different isoforms reported in (Figure 5.1). *In vivo*, the amelogenin protein is the main component of the dental organic enamel matrix. Enamel is one of the most calcified parts of the skeleton. In fact, the hydroxyapatite crystal represents 95% of the inorganic enamel matrix [54]. These calcium salts and derivatives are resistant to aging damage and can protect teeth and the proteins that constitute them for tens of thousands of years [55–58]. During enamel maturation, a natural proteolytic process fragments all proteins present in the tooth, such as ameloblastin (AMBN, 5%), amelogenin (AMELX/Y, 90%), enamelin (ENAM, approx. 3%), and matrix metalloproteinase-20 (MMP-20, approx. 2%), in various peptides (Figure 5.2) [54, 56, 59]. As the amelogenin protein is its main component, the resulting peptides originate mainly from amelogenin. For this reason, peptides from other proteins do not interfere with peptide analysis that is intended to estimate sex. In comparison, some studies applied an additional enzymatic treatment with trypsin to generate new peptides [55, 60–66]. Nevertheless, under the natural proteolytic activity, the average length of the peptide decreased with the age of the archeological sample, and the recovery of the peptide was higher [66]. Furthermore, due to its preponderance, only a small amount of archeological sample *i.e.* 50 mg, is required. Therefore, proteomics appears to be the method of choice for estimating sex. NanoLC-MS/MS with its higher sensitivity for peptide detections and identifications, performs seems to be the best analytical method [67]. The accuracy of the proteomics method is absolute, and it allowed us to solve the misclassified adult individuals and extended to sub-adult skeletal remains [68].

To conclude, this proteomics method presents the advantage of being the least sample-consuming method and being not contaminated when extracting. The next chapter will study the minimally-invasive character of this proteomics method.

```
AMELX-1     1  MGTWILFACL LGAAFAMPLP PHPGHPGYIN FSYE              VL  36
AMELX-2     1  MGTWILFACL LGAAFAMP                               VL  20
AMELX-3     1  MGTWILFACL LGAAFAMPLP PHPGHPGYIN FSYENSHSQA INVDRTALVL  50
AMELY-1     1  MGTWILFACL VGAAFAMPLP PHPGHPGYIN FSYE              VL  36
AMELY-2     1  MGTWILFACL VGAAFAMPLP PHPGHPGYIN FSYENSHSQA INVDRIALVL  50

AMELX-1    37  TPLKWYQS-I RPPYPSYGYE PMGGWLHHQI IPVLSQQHPP THTLQPHHHI  85
AMELX-2    21  TPLKWYQS-I RPPYPSYGYE PMGGWLHHQI IPVLSQQHPP THTLQPHHHI  69
AMELX-3    51  TPLKWYQS-I RPPYPSYGYE PMGGWLHHQI IPVLSQQHPP THTLQPHHHI  99
AMELY-1    37  TPLKWYQSMI RPPYSSYGYE PMGGWLHHQI IPVVSQQHPL THTLQSHHHI  86
AMELY-2    51  TPLKWYQSMI RPPYSSYGYE PMGGWLHHQI IPVVSQQHPL THTLQSHHHI  100

AMELX-1    86  PVVPAQQPVI PQQPMMPVPG QHSMTPIQHH QPNLPPPAQQ PYQPQPVQPQ  135
AMELX-2    80  PVVPAQQPVI PQQPMMPVPG QHSMTPIQHH QPNLPPPAQQ PYQPQPVQPQ  119
AMELX-3   100  PVVPAQQPVI PQQPMMPVPG QHSMTPIQHH QPNLPPPAQQ PYQPQPVQPQ  149
AMELY-1    87  PVVPAQQPRV RQQALMPVPG QQSMTPTQHH QPNLPLPAQQ PFQPQPVQPQ  136
AMELY-2   101  PVVPAQQPRV RQQALMPVPG QQSMTPTQHH QPNLPLPAQQ PFQPQPVQPQ  150

AMELX-1   136  PHQPMQPQPP VHPMQPLPPQ PPLPPMFPMQ PLPPMLPDLT LEAWPSTDKT  185
AMELX-2   130  PHQPMQPQPP VHPMQPLPPQ PPLPPMFPMQ PLPPMLPDLT LEAWPSTDKT  169
AMELX-3   150  PHQPMQPQPP VHPMQPLPPQ PPLPPMFPMQ PLPPMLPDLT LEAWPSTDKT  199
AMELY-1   137  PHQPMQPQPP VQPMQPLLPQ PPLPPMFPLR PLPPILPDLH LEAWPATDKT  186
AMELY-2   151  PHQPMQPQPP VQPMQPLLPQ PPLPPMFPLR PLPPILPDLH LEAWPATDKT  200

AMELX-1   186  KREEVD                                               191
AMELX-2   180  KREEVD                                               175
AMELX-3   200  KREEVD                                               205
AMELY-1   187  KQEEVD                                               192
AMELY-2   201  KQEEVD                                               206
```

FIGURE 5.1: Different proteinogenic sequence of AMELX and AMELY isoforms

Uniprot: Q99217-1 (AMELX-1), Q99217-2 (AMELX-2), Q99217-3 (AMELX-3), Q99218-1 (AMELY-1) and Q99218-2 (AMELY-2)

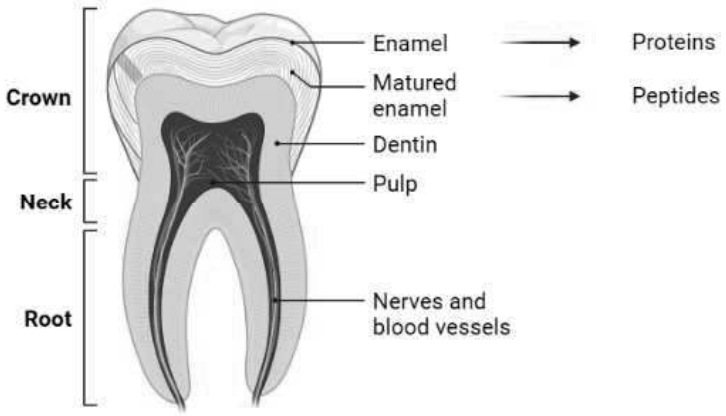

FIGURE 5.2: Composition of tooth

Created with BioRender.com

Chapter 6

Analysis of archaeological proteins

The main challenge for the analysis of archaeological proteins is to use a smaller quantity of samples, as this is possible with a minimum impact on the archaeological materials. In this chapter, the minimally-invasive character of the proteomics method for sex estimation will be described, applied, and evaluated.

6.1 Proteomics minimally-invasive method

The Proteomics method consists of analyzing the amelogenin protein contained in tooth enamel by nanoLC-MS and distinguishes both forms of AMELX and AMELY. Each tooth was cleaned with ultrapure water before undergoing chemical treatment on the tooth crown where the enamel is located (Figure 5.2). Low-concentrated H_2O_2 (3%, 200 μL, 30 s) was applied to demineralize the tooth surface to remove calcium phosphate salts and thereby prepare the tooth surface for chemical treatment. The tooth crown was then rinsed with ultrapure H_2O before proceeding to two successive chemical etching steps using low-concentrated HCl (5%, 200 μL, 2 min). The first etching allows the removal of enamel containing amelogenin in its protein form. Only the second etching solution containing amelogenin peptides from matured enamel was collected, concentrated, and prepared for proteomics analysis.

6.2 Evaluation of the minimally-invasive method

Previous studies have investigated the effect of highly concentrated HCl etching (37%, 8h). The results showed that the structure of the tooth surface was progressively and seriously deteriorated until complete dissolution within 8h [69]. Other strongly acidic conditions (H_2SO_4, 75%) were applied to study the progressive loss of enamel volume *i.e.* 0 % at 2h, 33% at 6h, 40% at 24h, and 84% at 96h [70]. Another study showed that a loss of av. 6.1 μm/min of enamel under phosphoric acid conditions (50%, 3 min) [71]. However, in the case of this study, low concentrations of H_2O_2 (3%) and HCl (5%) were used [72]. These concentrations are lower than the H_2O_2 concentration used in cosmetic tooth whitening (6%) in accordance with European directives [73].

Although the acid concentration and exposure time used in this study were much lower than those used in previous studies, their impact on archaeological samples must be measured. To evaluate the minimally-invasive character of the method, different tests were carried out: scanning electron microscope and micro-computed tomography.

6.3 Study of archaeological amelogenin

This minimally-invasive proteomics method was primarily applied to teeth from two control groups with known age and sex. The first control group was composed of 30 teeth from the recent adult population, whereas the second group was composed of 30 teeth from adult individuals autopsied in the past century [74]. The performance of this proteomics method was validated, with absolute accuracy, on teeth from recent and sub-recent adult individuals of known sex within both control groups. This method was then applied to archaeological teeth. Fifteen teeth from adults were selected because of their divergent estimates from previous studies [75]. In addition, 32 teeth from non-adult individuals were selected because of the impossibility of sex estimation using traditional morphological methods [76, 77]. Finally, scanning electron microscope and micro-computed tomography were used to evaluate the minimally-invasive proteomics method.

Chapter 7

Conclusion and perspectives

This dissertation thesis contributes to the development of promising analytical techniques for aging proteomics and paleoproteomics.

In the first section, the structural consequences of the natural aging process on peptide and protein sequences were studied.

Firstly, the main effect of aging on peptides and proteins is the racemization of amino acids. This racemization can occur *via* an amino acid racemase (enzymatic process) or *via* a succinimidyl intermediate (non-enzymatic process). By the enzymatic process, the racemization proceeds to free amino acids before or during peptide elongation, whereas by a non-enzymatic process, racemization proceeds on proteinogenic amino acids. However, knowledge surrounding amino acid racemization in terms of abundance, exact position, and detection is lacking. When protein turnover is low, the accumulation of D-amino acids in human or animal bodies can cause damage. Indeed, the correlation between free D-amino acids and D-amino acid-containing peptides and proteins and various age-related diseases and disorders was reported in this work. The hydrolysis of peptides and proteins is the first and most crucial step in determining and detecting D-amino acids in the proteinogenic sequence. Indeed, a deuterium environment is required to preserve the intact amino acid racemization rate and to limit the natural racemization that occurs during hydrolysis. Then, the development of a chiral separation method as the second crucial step is essential for the detection of D-amino acids. In this dissertation thesis, the performance of the chromatographic and electrophoretic techniques for amino acid enantioseparation were summarized. In chromatography, such as liquid, supercritical fluid, and gas chromatography, a chiral selector is

linked to the stationary material to create a chiral stationary phase. Conversely, in capillary electrophoresis, the chiral selector is added to the BGE as a pseudo-phase or a dual-ligand. In this review of the literature, different types of chiral selectors were summarized and compared. The results showed that crown ether is the most efficient chiral selector for separating underivatized amino acid enantiomers using liquid chromatography, supercritical fluid chromatography, and capillary electrophoresis. In addition, zwitterionic and macrocycle antibiotics are also among the most efficient chiral stationary phases in liquid chromatography. For gas chromatography, cyclofructan was the most efficient chiral selector. However, with all these enantioseparation techniques, the most challenging task is the separation of positional isomers: L-Ile, L-Leu, and their D-counterparts. Both Crownpak CR(+) and CR(-) offered the best enantioseparation. To facilitate the isolation, separation, and detection of amino acids, a derivatization reaction can be performed as an alkylation of a pure chiral reagent to form a pair of diastereomers. Traditionally, the derivatization reaction occurs on the amino group common to all amino acids by *N*-alkylation. Most of these derivatization reagents are commercial *i.e.* (+)- or (-)-FLEC and (*S*)-NIFE. Additional synthetic compounds have also been used, such as (*R*)- or (*S*)-BiAc and OTPTHE. The main advantage of the derivatization reaction on the amino group is the simultaneous analysis of all amino acids. In addition, for a more selective analysis, other specific derivatization reagents, such as NEM and NPEM, have been developed to link the thiol function on the cysteine residue by *S*-alkylation.

Secondly, aging collagens have been studied at different ages and from different organisms. This study has shown that the protein structure undergoes many changes during aging. The first change, amino acid racemization, occurred progressively with age. However, this racemization is not uniform. Indeed, some amino acids are more favorable for racemization according to their nature, position on the protein sequence, and three-dimensional environment. Using the novel chiral amino acid separation method developed in this study, the % of each D-amino acid was determined. The results showed a % D-amino acid-age correlation. As a result of a combination with peptide mapping, the exact positions of totally racemized D-amino acids in their D-forms were elucidated. Second structural change, post-translational modifications evolve during the aging process. In fact,

the number of hydrophilic groups resulting from oxidation, dioxidation, deamidation, phosphorylation, and sulfation reactions decreased, whereas the number of hydrophobic groups resulting from acetylation, carbamylation, carboxylation, carboxymethylation, carboxyethylation, formylation, and methylation reactions increased. A combination with peptide mapping made it possible to determine their exact positions on the sequence. These changes also progress with age and are the primary theories to explain the reduction of proteolysis and increase the hydrophobicity of aging proteins. Indeed, changes in solubility and proteolytic capacity have been observed in aging collagens. Four successive enzymatic treatments with pepsin, trypsin, proteinase K, and chymotrypsin were necessary to digest and solubilize the peptides that arise from aging collagen, whereas only the first two of them were used for recent collagen. The third change was the proteinogenic sequence, which degraded over time. With the comparison of peptide mapping at different ages, the results showed the loss of one-fifth of the information sequence in aging collagens.

This work is essential for aging proteomics and contributes to a better understanding of the effects of *in vivo* aging mechanisms on protein sequences. For the first time, complex structural changes in proteins during the aging process were studied using LC-MS. Next, for future work, all protein characterization methods, including cryogenic-electron microscopy, mass spectrometry, NMR spectroscopy, Raman spectroscopy, and X-ray crystallography, should be applied to aging proteins. The complementarity of these methods will confirm the stereochemistry of amino acids, post-translational modifications, and protein degradations. The evolution of post-translational modifications and the three-dimensional protein structures during aging should be added to protein databases. Other tools, such as AlphaFold using artificial intelligence, can predict the three-dimensional structure of proteins. It can also be applied to aging proteins to predict more favorable specific sites for amino acid racemization. Then, this analytical method can be transposed to all aging proteins and those associated with age-related diseases.

In the second section, paleoproteomics was used for the sex estimation of ancient skeletons.

The paleoproteomics analysis presents itself as a complementary method for sex estimation. The amelogenin protein, encoded by the amelogenin gene, is more resistant to age-related damage over tens of thousands of years, and is more conserved than its gene counterpart. In fact, the amelogenin protein is preserved in teeth and is the main component of enamel. Similar to the gene, both sex-dependent protein forms were distinguished by nanoLC-MS owing to differences in their amino acid compositions and sequence lengths. The major differences were the loss of 16 amino acids from positions 19 to 34 and the methionine residue at position 45 in the AMELX-2 isoform. Other minor differences appeared with the exchange of 22 amino acid residues with others throughout the sequence. The combination of these proteinogenic modifications contributes to a sequence homology of 93%. For the extraction of amelogenin protein from the tooth, low concentrations of H_2O_2 (3%) were applied to prepare the surface of the tooth to remove calcium phosphate salts. This concentration was lower than the concentration present in commercial cosmetics and hygiene products. Etching of tooth was then performed in two steps using low-concentrated HCl (5%). The first acidic etch allows the removal of the tooth enamel surface where the intact amelogenin protein is located. The second etch allows for the collection of amelogenin peptides in the matured enamel. Enzymatic treatment was not necessary because of the natural proteolytic process in the matured enamel. As the amelogenin protein is the main component of enamel, the resulting peptides in matured enamel are also the majority. Other peptides resulting from ameloblastin, enamelin, and matrix-metalloproteinase-20 did not interfere with the detection of amelogenin peptides by nanoLC-MS. The performance of this proteomics method was validated with 100% accuracy in both control groups consisting of recent and sub-recent adult individuals of known sex. This method was then applied to the two groups of archaeological teeth. The first group comprised 15 teeth from adults, who were selected because of their divergent estimates from previous studies. The second group consisted of 32 teeth from non-adult individuals who were selected because of the impossibility of sex estimation using traditional morphological methods. Sex was successfully estimated for archaeological adult and juvenile teeth in both unknown groups. The accuracy of the proteomics

method is absolute, which allowed us to solve misclassified adult individuals and extend them to sub-adult skeletal remains. To evaluate the minimally-invasive character of the proteomics method, 20 teeth of recent and sub-recent individuals were scanned before and after the chemical treatment to observe microscopic changes using scanning electron microscope and micro-computed tomography. The results showed a loss of approximately 10% of enamel using a scanning electron microscope and a loss of only 2% of dentin using micro-computed tomography.

This work is essential in paleoproteomics and contributes to archaeological, anthropological, and forensic research, with minimal impact on archaeological material. This proteomics method developed has the advantage of being absolute for sex estimation and has the least sample consumption without contamination. The minimally-invasive nature of the proteomics method was evaluated for the first time. For future work, the sex estimation of adult and juvenile individuals with complete or fragmented skeletons can be performed using proteomics analysis. Additionally, all human remains from museum or university collections can benefit from this sex estimation method and may resolve the misclassification ambiguities. Since proteins are more resistant than the corresponding genes, this analytical method can be applied to other proteins encoded by other genes, as an alternative to aDNA analysis.

Bibliography

[1] E. Q. Jennings, K. S. Fritz, J. J. Galligan, *Molecular Aspects of Medicine* **2022**, *86*, 101053.

[2] R. Harmel, D. Fiedler, *Nature chemical biology* **2018**, *14*, 244–252.

[3] M. Schneider, I. W. Marison, U. Von Stockar, *Journal of biotechnology* **1996**, *46*, 161–185.

[4] J. H. McKerrow, *Mechanisms of Ageing and Development* **1979**, *10*, 371–377.

[5] J. Zecha, W. Gabriel, R. Spallek, Y.-C. Chang, J. Mergner, M. Wilhelm, F. Bassermann, B. Kuster, *Nature communications* **2022**, *13*, 165.

[6] S. Ramazi, J. Zahiri, *Database* **2021**, *2021*.

[7] M. Morvan, I. Mikšík, *Separations* **2021**, *8*, 112.

[8] C. Vermeer, *Biochemical journal* **1990**, *266*, 625.

[9] M. L. Cancela, N. Conceição, V. Laizé, *Advances in Nutrition* **2012**, *3*, 174–181.

[10] L.-L. Hu, S. Niu, T. Huang, K. Wang, X.-H. Shi, Y.-D. Cai, *PLoS One* **2010**, *5*, e15917.

[11] M. Morvan, I. Mikšík, *Analytica Chimica Acta* **2023**, *1262*, 341260.

[12] D. L. Martin, K. Rimvall, *Journal of neurochemistry* **1993**, *60*, 395–407.

[13] S. Matsumoto, J. Häberle, J. Kido, H. Mitsubuchi, F. Endo, K. Nakamura, *Journal of human genetics* **2019**, *64*, 833–847.

[14] Y. Zhao, X. Zhu, W. Jiang, H. Liu, B. Sun, *Molecules* **2021**, *26*, 1145.

[15] R. L. Levine, E. R. Stadtman, *Experimental gerontology* **2001**, *36*, 1495–1502.

[16] G. Krishnamoorthy, R. Selvakumar, T. P. Sastry, A. B. Mandal, M. Doble, *Biochemical engineering journal* **2013**, *75*, 92–100.

[17] J. W. Silzel, G. Ben-Nissan, J. Tang, M. Sharon, R. R. Julian, *Analytical Chemistry* **2022**, *94*, 15288–15296.

[18] A. L. Santos, A. B. Lindner, *Oxidative Medicine and Cellular Longevity* **2017**, *2017*.

[19] I. Mikšík, P. Sedláková, S. Pataridis, F. Bortolotti, R. Gottardo, *Protein Science* **2016**, *25*, 2037–2044.

[20] R. J. Truscott, J. Mizdrak, M. G. Friedrich, M. Y. Hooi, B. Lyons, J. F. Jamie, M. J. Davies, P. A. Wilmarth, L. L. David, *Experimental eye research* **2012**, *99*, 48–54.

[21] S. S. Atavliyeva, P. V. Tarlykov, *Experimental Biology* **2021**, *89*, 4–14.

[22] N. Fujii, T. Kawaguchi, H. Sasaki, N. Fujii, *Biochemistry* **2011**, *50*, 8628–8635.

[23] M. Y. S. Hooi, R. J. Truscott, *Age* **2011**, *33*, 131–141.

[24] M. Y. S. Hooi, M. J. Raftery, R. J. Truscott, *Experimental Eye Research* **2013**, *106*, 34–39.

[25] B. Lyons, A. H. Kwan, J. Jamie, R. J. Truscott, *The FEBS journal* **2013**, *280*, 1980–1990.

[26] N. Fujii, T. Takata, N. Fujii, K. Aki, H. Sakaue, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2018**, *1866*, 840–847.

[27] S. Ha, T. Kinouchi, N. Fujii, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2020**, *1868*, 140410.

[28] J. J. Bastings, H. M. van Eijk, S. W. Olde Damink, S. S. Rensen, *Nutrients* **2019**, *11*, 2205.

[29] R. C. Strauch, E. Svedin, B. Dilkes, C. Chapple, X. Li, *Proceedings of the National Academy of Sciences* **2015**, *112*, 11726–11731.

[30] T. Miyamoto, T. Moriya, H. Homma, T. Oshima, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2020**, *1868*, 140461.

[31] S. Du, M. Wey, D. W. Armstrong, *Chirality* **2023**.

[32] C. Ollivaux, D. Soyez, J.-Y. Toullec, *Journal of peptide science* **2014**, *20*, 595–612.

[33] N. Fujii, H. Sakaue, H. Sasaki, N. Fujii, *Journal of Biological Chemistry* **2012**, *287*, 39992–40002.

[34] H. Frank, W. Woiwode, G. Nicholson, E. Bayer, *Liebigs Annalen der Chemie* **1981**, *1981*, 354–365.

[35] J. Csapá, Z. Csapó-Kiss, L. Wágner, T. Tálos, T. G. Martin, S. Folestad, A. Tivesten, S. Némethy, *Analytica chimica acta* **1997**, *339*, 99–107.

[36] G. Yasunaga, S. Inoue, T. Bando, T. Hakamada, Y. Fujise, *Marine Mammal Science* **2022**.

[37] T. Miyamoto, H. Homma, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2018**, *1866*, 775–782.

[38] M. Danielsen, C. Nebel, T. K. Dalsgaard, *Foods* **2020**, *9*, 309.

[39] S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, W. H. Moos, *Drug Development Research* **1995**, *35*, 20–32.

[40] S. Ricard-Blum, *Cold Spring Harbor Perspectives in Biology* **2011**, *3*, 004978–004978.

[41] S. Köster, H. M. Evans, J. Y. Wong, T. Pfohl, *Biomacromolecules* **2008**, *9*, 199–207.

[42] A. K. Gulevsky, I. I. Shcheniavsky, *Biotechnologia Acta* **2020**, *13*, 42–61.

[43] E. Gineyts, P. A. C. Cloos, O. Borel, L. Grimaud, P. D. Delmas, P. Garnero, *Biochemical Journal* **2000**, *345*, 481–485.

[44] N. Verzijl, J. DeGroot, S. R. Thorpe, R. A. Bank, J. N. Shaw, T. J. Lyons, J. W. Bijlsma, F. P. Lafeber, J. W. Baynes, J. M. TeKoppele, *Journal of Biological Chemistry* **2000**, *275*, 39027–39031.

[45] P. A. C. Cloos, C. Fledelius, *Biochemical Journal* **2000**, *345*, 473–480.

[46] J. L. Bada, R. A. Schroeder, R. Protsch, R. Berger, *Proceedings of the National Academy of Sciences* **1974**, *71*, 914–917.

[47] N. K. Shah, B. Brodsky, A. Kirkpatrick, J. A. Ramshaw, *Biopolymers: Original Research on Biomolecules* **1999**, *49*, 297–302.

[48] V. Punitha, S. S. Raman, R. Parthasarathi, V. Subramanian, J. R. Rao, B. U. Nair, T. Ramasami, *The Journal of Physical Chemistry B* **2009**, *113*, 8983–8992.

[49] D. M. S. Couto, N. C. D. Gallassi, S. L. Gomes, V. Ulbricht, J. S. Pereira Neto, E. Daruge Junior, L. Francesquini Junior, *Journal of Forensic Dental Sciences* **2019**, 73–77.

[50] Á. Azevedo, M. L. Pereira, S. Gouveia, J. N. Tavares, I. M. Caldas, *Forensic Science Medicine and Pathology* **2019**, *15*, 191–197.

[51] D. Franklin, A. Cardini, A. Flavel, M. K. Marks, *International journal of legal medicine* **2014**, *128*, 861–872.

[52] E. M. Mostafa, A. H. El-Elemi, M. A. El-Beblawy, A. E.-W. A. Dawood, *Egyptian Journal of Forensic Sciences* **2012**, *2*, 81–88.

[53] I. Mikšík, M. Morvan, J. Brůžek, *Journal of Separation Science* **2023**, 2300183.

[54] E. Cuéllar-Rivas, M. C. Pustovrh-Ramos, *Revista Facultad de Odontología Universidad de Antioquia* **2015**, *27*, 154–176.

[55] G. J. Parker, J. M. Yip, J. W. Eerkens, M. Salemi, B. Durbin-Johnson, C. Kiesow, R. Haas, J. E. Buikstra, H. Klaus, L. A. Regan, et al., *Journal of Archaeological Science* **2019**, *101*, 169–180.

[56] R. H. Ziganshin, N. Y. Berezina, P. L. Alexandrov, V. V. Ryabinin, A. P. Buzhilova, *Biochemistry (Moscow)* **2020**, *85*, 614–622.

[57] L. Zhu, H. Liu, H. E. Witkowska, Y. Huang, K. Tanimoto, W. Li, *Frontiers in Physiology* **2014**, *5*, 268.

[58] J. Moradian-Oldak, J. Tan, A. G. Fincham, *Biopolymers: Original Research on Biomolecules* **1998**, *46*, 225–238.

[59] V. Uskoković, F. Khan, H. Liu, H. E. Witkowska, L. Zhu, W. Li, S. Habelitz, *Archives of oral biology* **2011**, *56*, 1548–1559.

[60] I. M. Porto, H. J. Laure, R. H. Tykot, F. B. de Sousa, J. C. Rosa, R. F. Gerlach, *European journal of oral sciences* **2011**, *119*, 83–87.

[61] I. M. Porto, H. J. Laure, F. B. de Sousa, J. C. Rosa, R. F. Gerlach, *Journal of Archaeological Science* **2011**, *38*, 3596–3604.

[62] G. A. Castiblanco, D. Rutishauser, L. L. Ilag, S. Martignon, J. E. Castellanos, W. Mejía, *European journal of oral sciences* **2015**, *123*, 390–395.

[63] V. C. Wasinger, D. Curnoe, S. Bustamante, R. Mendoza, R. Shoocongdej, L. Adler, A. Baker, K. Chintakanon, C. Boel, P. S. Tacon, *Proteomics* **2019**, *19*, 1800341–1800352.

[64] R. Haas, J. Watson, T. Buonasera, J. Southon, J. C. Chen, S. Noe, K. Smith, C. Viviano Llave, J. Eerkens, G. Parker, *Science Advances* **2020**, *6*, eabd0310.

[65] C. Froment, M. Hourset, N. Sáenz-Oyhéréguy, E. Mouton-Barbosa, C. Willmann, C. Zanolli, R. Esclassan, R. Donat, C. Thèves, O. Burlet-Schiltz, et al., *Journal of proteomics* **2020**, *211*, 103548–103559.

[66] F. Welker, J. Ramos-Madrigal, P. Gutenbrunner, M. Mackie, S. Tiwary, R. Rakownikow Jersie-Christensen, C. Chiva, M. R. Dickinson, M. Kuhlwilm, M. de Manuel, et al., *Nature* **2020**, *580*, 235–238.

[67] L. Yi, P. D. Piehowski, T. Shi, R. D. Smith, W.-J. Qian, *Journal of Chromatography A* **2017**, *1523*, 40–48.

[68] J. Brůžek, I. Mikšík, A. Pilmann Kotěrová, M. Morvan, S. Drtikolová Kaupová, F. Santos, A. Danielisová, E. Zazvonilová, B. Maureille, P. Velemínskỳ, *Preprint* **2023**.

[69] M. Raj, K. Boaz, N Srikant, *Journal of Forensic Dental Sciences* **2013**, *5*, 7.

[70] A. Thurzo, V. Jančovičová, M. Hain, M. Thurzo, B. Novák, H. Kosnáčová, V. Lehotská, I. Varga, P. Kováč, N. Moravanský, *Molecules* **2022**, *27*, 4035.

[71] R Lehman, C. L. Davidson, *American Journal of Orthodontics* **1981**, *80*, 73–82.

[72] N. A. Stewart, R. F. Gerlach, R. L. Gowland, K. J. Gron, J. Montgomery, *Proceedings of the National Academy of Sciences* **2017**, *114*, 13649–13654.

[73] S. Dias, A. Mata, J. Silveira, R. Pereira, A. Putignano, G. Orsini, R. Monterubbianesi, D. Marques, *Materials* **2021**, *14*, 7597.

[74] P. Pachner, *Sexual differences in human pelvis*, Česká akademie věd a umění, Praha, **1937**.

[75] E. Zazvonilová, P. Velemínský, J. Brůžek, *Archeologické Rozhledy* **2020**, *72*, 67–101.

[76] P. Velemínský, Š. Bejdová, L. Bigoni, H. Brzobohatá, S. Drtikolová Kaupová, P. Brukner Havelková, A. Ibrová, V. Sládek, P. Stránská, J. Velemínská, E. Zazvonilová, J. Brůžek in *Great Moravian Elites From Mikulčice*, (Ed.: L Poláček), Czech Academy of Sciences, Institute of Archaeology, Brno, **2021**, pp. 385–428.

[77] S. Drtikolová Kaupová, J. Brůžek, J. Hadrava, I. Mikšík, M. Morvan, L. Poláček, L. Půtová, P. Velemínskỳ, *Archaeological and Anthropological Sciences Preprint* **2022**.

Scientific valorization

1. Articles in peer-reviewed journals

S. Kaupová*, J. Brůžek, J. Hadrava, I. Mikšík, **M. Morvan**, L. Poláček, L. Půtová, P. Velemínský, Early life histories of Great Moravian children carbon and nitrogen isotopic analysis of dentine serial sections from the Early Medieval population of Mikulčice (9^{th}-10^{th} centuries AD, Czechia), *Archaeological and Anthropological Sciences*, **2022**, *under review*. - Preprint DOI:10.21203/rs.3.rs-1913554/v1 - IF = 2.2

J. Brůžek*, I. Mikšík*, A. Pilmann Kotěrová, **M. Morvan**, J. Dašková, P. Velemínský, S. Drtikolová Kaupová, F. Santos, J. Velemínská, A. Danielisová, E. Zazvonilová, B. Maureille, Undertaking sex assessment of human remains within cultural heritage: Applicability of a minimally-invasive methods for the proteomic sex estimation from enamel peptides, *Journal of Cultural Heritage*, **2023**, *in revisions*. - Preprint DOI:10.2139/ssrn.4439221 - IF = 3.1

I. Mikšík*, **M. Morvan**, J. Brůžek, Peptide analysis of tooth enamel - a sex estimation tool for archaeological, anthropological, or forensic research, *Journal of Separation Science*, **2023**, 46, 2300183. - DOI:10.1002/jssc.202300183 - IF = 3.1

M. Morvan*, I. Mikšík, The chiral proteomic analysis applied to aging collagens by LC-MS: Amino acid racemization, post-translational modifications, and sequence degradations during the aging process, *Analytica Chimica Acta*, **2023**, 1262, 341260. - DOI:10.1016/j.aca.2023.341260 - IF = 6.2

M. Morvan*, I. Mikšík*, Recent Advances in Chiral Analysis of Proteins and Peptides, *Separations*, **2021**, 8 (8), 112. - DOI:10.3390/separations8080112 - IF = 2.6

I. Mikšík*, Š. Kubinová, **M. Morvan**, K. Výborný, A. Tatar, V. Král, K. Záruba, D. Sýkora, Analysis of chondroitin/dermatan sulphate disaccharides using high-performance liquid chromatography, *Separations*, **2020**, 7 (3), 49. - DOI:10.3390/separations7030049 - IF = 2.6

L. Leclercq, **M. Morvan**, J. Koch, C. Neusüß, H. Cottet*, Modulation of the electoosmotic mobility using polyelectrolyte multilayer coatings for protein analysis by capillary electrophoresis, *Analytica Chimica Acta*, **2019**, 1057, 152-161. - DOI:10.1016/j.aca.2019.01.008 - IF = 6.2

2. Journal cover page

M. Morvan*, I. Mikšík, Determination of Protein and Peptide chirality, *Separations*, **2021**, 8 (8) - IF = 2.6

3. Oral communications

M. Morvan*, I. Mikšík, Chiral and aging proteomics applied to bovine and rat collagen, PhD meeting – Institute of Physiology CAS, Prague (Czech Republic), November 1^{st}-2^{nd}, 2022.

M. Morvan*, I. Mikšík, Chiral separations for aging proteomics, PhD meeting – Institute of Physiology CAS, Prague (Czech Republic), May 16^{th}, 2022.

M. Morvan*, I. Mikšík, Chiral proteomics analysis applied to aging collagens, PhD meeting – University of Pardubice, Pardubice (Czech Republic), January 13^{th}, 2022.

L. Leclercq*, S. Bekri, **M. Morvan**, J. Koch, C. Neusüß, H. Cottet, Polyelectrolyte multilayer coatings for the separation of proteins by capillary electrophoresis: influence of polyelectrolyte nature, ITP 2019 - Toulouse (France), Spetember 1^{st}-4^{th}, 2019.

M. Morvan*, L. Leclercq, H. Cottet, Polyelectrolyte multilayer coatings for protein analysis by capillary electrophoresis, DSBC group scientific meeting - University of Montpellier, Montpellier (France), December 5^{th}, 2016.

4. Posters

I. Mikšík*, J. Brůžek, A. Kotěrová, **M. Morvan**, J. Dašková, P. Velemínský, F. Santos, J. Velemínská, A. Danielisová, E. Zazvonilová, B. Maureille, Use of a minimally-invasive method for the proteomic sex estimation from human tooth enamel, APCE-CECE-ITP-IUPAC 2022 - Siem Reap (Cambodia), November 6^{th}-10^{th}, 2022.

M. Morvan*, I Mikšík, Chiral amino acids analysis and aging proteomics applied to collagens, Analytics 2022 - Nantes (France), September 5^{th}-8^{th}, 2022.

M. Morvan*, I Mikšík, Chiral separation for proteomics in aging collagens, MSB 2022 - Liège (Belgium), July 3^{rd}-6^{th}, 2022.

M. Morvan*, I Mikšík, Chiral Amino Acid and Peptide Separations for Proteomics Applied to Aging Collagens, HPLC 2022 - San Diego (USA), June 18^{th}-23^{rd}, 2022.

I. Mikšík*, J. Brůžek, A. Kotěrová, **M. Morvan**, J. Dašková, P. Velemínský, F. Santos, J. Velemínská, A. Danielisová, E. Zazvonilová, B. Maureille, Use of Proteomic Analysis for Sex Determination in Human Tooth Enamel, HPLC 2022 - San Diego (USA), June 18^{th}-23^{rd}, 2022.

M. Morvan*, I Mikšík, Chiral proteinogenic amino acid analysis applied to aging collagen, Annual congress of the Institute of Physiology – Nesuchyně (Czech Republic), November 6^{th}-8^{th}, 2021.

S. Berki, **M. Morvan**, J. Koch, L. Salzer, C. Neusüß, L. Leclercq*, H. Cottet*, Polyelectrolyte multilayer coatings for the separation of proteins and monoclonal antibodies by capillary electrophoresis, University of Montpellier - Montpellier (France), February 26^{th}, 2020.

*corresponding or presenting author

Scientific formations

Proteomics data interpretation
Workshop - June 8^{th}, 2023
Proteomic Section of Czech Society for Biochemistry and Molecular Biology, Czech Republic

Introduction to histology: exploration of the tissues of the human body
Massive Open Online Course - certificate - February 23^{rd}, 2023
University of Liège, Belgium

Multi-platform metabolomics coverage
Workshop - February 15^{th}, 2023
Swiss Metabolomics Society, Switzerland

Journey to the hart of living things with X-rays: the crystallography
Massive Open Online Course - certificate - January 11^{th}, 2022
Paris Saclay University, France

Advances mass spectrometry applied to cultural heritage
Summer school - certificate - June 16^{th}-18^{th}, 2021
University of Bordeaux, France

Writing for publication in English for Czech academics and researchers in Chemistry
Workshop - February 4^{th}, 2021
University of Pardubice, Czech Republic