Proteomics analysis of aging proteins

Marine Morvan

The University of Pardubice
Faculty of Chemical Technology

Proteomics analysis of aging proteins

Doctoral Thesis

2023 M.Sc. Marine Morvan

UNIVERSITY OF PARDUBICE

DOCTORAL THESIS

Proteomics analysis of aging proteins

Author:
M.Sc. Marine MORVAN

Supervisor:
Dr. Petr ČESLA

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

Faculty of Chemical Technology, University of Pardubice

August 18, 2023

Declaration of Authorship

I declare:

The thesis entitled "Proteomics analysis of aging proteins" is my own work. All literary sources and information that I used in the thesis are referenced in the bibliography.

I have been acquainted with the fact that my work is subject to the rights and obligations arising from Act No. 121/2000 Sb., on Copyright, on Rights Related to Copyright and on Amendments to Certain Acts (Copyright Act), as amended, especially with the fact that the University of Pardubice has the right to conclude a license agreement for the use of this thesis as a school work under Section 60, Subsection 1 of the Copyright Act, and that if this thesis is used by me or a license to use it is granted to another entity, the University of Pardubice is entitled to request a reasonable fee from me to cover the costs incurred for the creation of the work, depending on the circumstances up to their actual amount.

I acknowledge that in accordance with Section 47b of Act No. 111/1998 Sb., on Higher Education Institutions and on Amendments to Other Acts (Act on Higher Education Institutions), as amended, and the Directive of the University of Pardubice No. 7/2019 Rules for Submission, Publication and Layout of Theses, as amended, the thesis will be published through the Digital Library of the University of Pardubice.

In Pardubice on August 18, 2023

Marine MORVAN

UNIVERSITY OF PARDUBICE

Abstract

Faculty of Chemical Technology, University of Pardubice

Doctor of Philosophy

Proteomics analysis of aging proteins

by Marine MORVAN

Abstract: The analysis and characterization of aging and archaeological proteins are among the latest challenges in analytical chemistry. Different techniques have been developed to analyze recent proteins but have not been fully applied to aging proteins. This contributes to the lack of knowledge regarding the aging mechanism of proteins. In addition, the analysis of archaeological proteins, which are much less studied than ancient DNA but more resistant to degradation processes, is a promising candidate for archeological as well as forensic research.

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. Collagens from different organisms of different ages were used for this purpose. To determine the exact rate of amino acid racemization, protein hydrolysis conditions were used in a deuterium environment to exclude natural racemization during hydrolysis. Subsequently, a chiral separation method was developed to determine the amino acid enantiomer rates. The results showed that these rates were progressive in the D-form of amino acids, according to age. The evolution of post-translational modifications was also progressive with age and is the primary theory to explain the reduction of proteolysis and increases the hydrophobicity of aging proteins. Furthermore, peptide mapping comparison of collagens at different ages showed that natural sequence degradation occurs during aging and causes a loss of one-fifth of the protein sequence information.

In the second part of this dissertation thesis, archaeological proteins were studied using nanoLC-MS. Proteomics, called paleoproteomics in this case, can be applied to ancient samples for archaeological, anthropological, and forensic research. Sex estimation is fundamental for the characterization of skeletal remains as it is the first step in human identification. Osteoarchaeology and genomics are the traditional methods used for this estimation; however, they have limitations and are not absolute. In this study, paleoproteomics was developed as an alternative method. 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 proteogenic sequences. This method is more reliable, and less restrictive than previous methods. As teeth remains as archaeological samples are rare and valuable, few are available for destructive biological and chemical analyses. 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-computed tomography.

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

UNIVERSITY OF PARDUBICE

Abstract

Faculty of Chemical Technology, University of Pardubice

Doctor of Philosophy

Proteomics analysis of aging proteins

by Marine MORVAN

Abstrakt: Analýza a charakterizace stárnutí a archeologických proteinů patří mezi nejnovější výzvy v analytické chemii. Byly vyvinuty různé techniky pro analýzu současných proteinů, ale nebyly plně aplikovány na stárnoucí proteiny. To přispívá k nedostatku znalostí týkajících se mechanismu stárnutí proteinů. Kromě toho analýza archeologických proteinů, které jsou mnohem méně studovány než starobylá DNA, ale jsou odolnější vůči degradačním procesům, je slibným kandidátem pro archeologický i forenzní výzkumu.

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. K tomuto účelu byly použity kolageny z různých organismů různého stáří. Pro stanovení přesné rychlosti racemizace aminokyselin byly použity podmínky hydrolýzy proteinu v prostředí deuteria k vyloučení přirozené racemizace během hydrolýzy. Následně byla vyvinuta metoda chirální separace pro stanovení poměrů aminokyselinových enantiomerů. Výsledky ukázaly, že tyto poměry byly v závislosti na věku progresivní u D-aminokyselin. Vývoj posttranslačních modifikací byl také progresivní s věkem a je primární teorií, která vysvětluje redukci proteolýzy a zvýšení hydrofobicity stárnutí proteinů. Navíc srovnání mapování peptidů kolagenů v různých věkových věcích ukázalo, že při stárnutí dochází k degradaci přirozené sekvence během stárnutí a způsobuje ztrátu jedné pětiny informací o proteinové sekvenci.

Ve druhé části této disertační teze byly archeologické proteiny studovány pomocí NanoLC-MS. Proteomika, v tomto případě zvaná paleoproteomika, může být aplikována na starobylé vzorky pro archeologický, antropologický a forenzní výzkum. Odhad pohlaví je zásadní pro charakterizaci kosterních pozůstatků, protože je to první krok v lidské identifikaci. Osteoarchaeologie a genomika jsou tradiční metody používané pro tento odhad, mají však omezení a nejsou absolutní. V této studii byla paleoproteomika vyvinuta jako alternativní metoda. Na základě dvou forem proteinu amelogenin, který je pohlavně rozdílný, konzervovaných v zubech, byly oba pohlaví rozlišeny metodou nanoLC-MS na základě rozdílů v jejich proteogenních sekvencích. Tato metoda je spolehlivější a méně omezující než předchozí metody. Protože zuby jsou jako archeologické vzorky vzácné a cenné, jen málokdy jsou k dispozici pro destruktivní biologické a chemické analýzy. 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.

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

UNIVERSITY OF PARDUBICE

Abstract

Faculty of Chemical Technology, University of Pardubice

Doctor of Philosophy

Proteomics analysis of aging proteins

by Marine MORVAN

Résumé: L'analyse et la caractérisation des protéines durant le vieillissement et archéologiques constituent des défis récents en chimie analytique. Bien que différentes techniques aient été développées pour analyser les protéines récentes, elles n'ont pas encore été complètement appliquées aux protéines durant le vieillissement. Ce manque de recherche contribue à une lacune de connaissances concernant les mécanismes du vieillissement des protéines. En outre, l'analyse des protéines archéologiques, bien qu'elles soient moins étudiées que l'ADN ancien, présente un niveau de conservation supérieur et constitue donc d'excellentes candidates pour contribuer à la recherche archéologique et médico-légale.

Dans la première partie de cette thèse de doctorat, une étude des protéines durant le processus de vieillissement a été réalisée par LC-MS. Les effets du vieillissement sur la séquence protéique, y compris la racémisation des acides aminés, les modifications post-traductionnelles et la dégradation des protéines, ont été examinés. Des échantillons de collagène provenant d'organismes et d'âges différents ont été utilisés à cette fin. Pour déterminer précisément le taux de racémisation des acides aminés, l'hydrolyse des protéines a été réalisée dans un milieu deutéré pour empêcher la racémisation naturelle pendant le processus d'hydrolyse. Une méthode de séparation chirale a ensuite été développée pour évaluer les ratios énantiomériques des acides aminés. Les résultats ont révélé que ces ratios augmentaient progressivement pour la forme D des acides aminés en fonction de l'âge. L'évolution des modifications post-traductionnelles a également montré une progression avec l'âge, ce qui pourrait expliquer la réduction de la protéolyse et l'augmentation de l'hydrophobicité des protéines durant le vieillissement. De plus,

la comparaison de la cartographie peptidique du collagène à différents âges a révélé une dégradation naturelle de sa séquence au fil du vieillissement, entraînant une perte d'environ un cinquième de son information.

Dans la deuxième partie de cette thèse de doctorat, une étude a été menée sur les protéines archéologiques en utilisant la nanoLC-MS. La paléoprotéomique, une branche de la protéomique adaptée aux échantillons anciens, a été appliquée dans le contexte de la recherche archéologique, anthropologique et médico-légale. L'estimation du sexe est d'une importance fondamentale dans la caractérisation des restes squelettiques, car elle constitue la première étape de l'identification humaine. Les méthodes traditionnelles telles que l'ostéoarchéologie et la génomique sont couramment utilisées pour cette estimation, mais elles présentent des limites et ne sont pas infaillibles. Dans cette étude, la paléoprotéomique a été développée comme une méthode alternative. Grâce à l'analyse par nanoLC-MS, deux formes de protéines d'amélogénine conservées dans les dents et variant en fonction du sexe biologique, ont été identifiées en raison de leurs différences au niveau de leurs séquences protéiques. Cette méthode est plus fiable et moins contraignante que les méthodes précédentes. Étant donné que les restes dentaires, en tant qu'échantillons archéologiques, sont rares et précieux, seuls quelques-uns d'entre eux sont disponibles pour des analyses biologiques et chimiques destructives. L'approche protéomique développée dans cette étude a été conçue pour être peu invasive. Cette caractéristique a été confirmée par l'analyse des dents avant et après l'extraction de l'amélogénine à l'aide d'un microscope électronique à balayage et d'une microtomographie assistée par ordinateur.

Mots-clés: amélogénine, acides aminés, collagène, dégradation des protéines, LC-MS, modifications post-traductionnelles, paléoprotéomique, protéomique durant le vieillissement, séparation chirale.

Acknowledgements

I would like to express my deepest gratitude to my previous supervisor **Pr. Ivan Mikšík**, from the Laboratory of Translational Metabolism of the Institute of Physiology of the Czech Academy of Sciences, in Prague, and the Department of Analytical Chemistry of the University of Pardubice, to accept me as a Ph.D. Student. Thank you for your wise advice, and for allowing me to participate in international projects and conferences.

I would like to thank **Dr. Petr Česla**, from the Department of Analytical Chemistry of the University of Pardubice, and **Dr. Tomáš Čajka**, from the Laboratory of Translational Metabolism of the Institute of Physiology of the Czech Academy of Sciences, in Prague, for becoming my supervisors for the last months of my doctorate.

I thank **Pr. Jaroslav Brůžek**, from the Department of Anthropology and Human Genetics of the Charles University, in Prague, and our colleagues for each enriching interdisciplinary collaboration.

My sincere thanks to **Dr. Dušan Koval**, from the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, in Prague, and all our colleagues for having warmly welcomed me in their laboratory as a visiting Ph.D. Student, and your help and support to complete my experiments.

I would also like to thank **Pr. Serge Rudaz**, from the Institute of Pharmaceutical Sciences of Western Switzerland of the University of Geneva, for welcoming me into his laboratory as an international visiting Ph.D. Student.

My sincere thanks to my French chemist colleagues **M.Sc. Rémi Coulon** (Ph.D. Student at Palacký University, Olomouc) and **M.Sc. Yann Jézéquel** (Ph.D. Student at Masaryk University, Brno) for their support and mutual aid in this Czech doctoral system, and when the situation in the Czech Republic was difficult for us. I wish you all the best for the end of your Ph.D. and future projects. To my French biologist colleague **M.Sc. Roman Coupeau** (Research Assistant at the Institute of Physiology, CAS), good luck and take advantage of all the opportunities.

I thank my students from the *Lycée français de Prague*, and their families for allowing me to teach Physics and Chemistry. Thank you for your trust, your warm welcome, your scientific curiosity, your keen interest in my thesis, and your support. Congrats to you on obtaining your *baccalauréat* with honors, and welcome to the world of university studies!

Last and most of all, I would like to express my deepest gratitude to my longtime friends in France for all their love, encouragement, and constant support throughout my doctorate in a foreign country. Thanks for believing in me. Special thanks to the French-speaking Ph.D. Student community on Discord, particularly the (bio)chemistry team, and expat members around the world. Thanks for your daily support, your expertise in programming and data treatment, and all scientific and non-scientific discussions. It was my great pleasure to meet you in the Czech Republic, France, and Switzerland, in personal and professional contexts.

Additionally, it is with growing gratitude that I extend my most sincere thanks to all those who have actively contributed to struggle the threats and intimidations that I have suffered during my doctorate. Thank you for the strength and courage you gave me. The accomplishment of this manuscript would not have been possible without your daily support and encouragement. We will meet again soon for new adventures around the world, much better and more benevolent, I wish it for us.

Scientific valorization

1. Articles in peer-reviewed journals

1.1. As part of the main research project

1.1.1. *Published*

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

1.1.2. *Submitted*

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

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

1.1.3. *Journal cover page*

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

1.2. As part of side projects

1.2.1. *Published*

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

1.3. Other contributions

1.3.1. *Published*

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

*corresponding author

2. Oral communications

2.1. As part of the main research project

2.1.1. *National conferences*

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.

2.2. Other contributions

2.2.1. *International conferences*

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.

2.2.2. *National conferences*

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.

*presenting author

3. Posters

3.1. As part of the main research project

3.1.1. *International conferences*

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.

3.1.2. *National conferences*

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.

3.3. Other contributions

3.3.1. *National conferences*

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 26th, 2020.

*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

Contents

List of Figures

List of Abbreviations

A.

aDNA	Ancient deoxyribonucleic acid
AMBN	Ameloblastin
AMELX	X form of amelogenin protein
AMELY	Y form of amelogenin protein
Ala	Alanine
AlaR	Alanine racemase
Arg	Arginine
ArgR	Arginine racemase
Asn	Asparagine
Asp	Aspartic acid
AspR	Aspartic acid racemase
Asx	Asparagine or aspartic acid

B.

BGE	Background electrolyte

C.

CE	Capillary electrophoresis
COL1A1	Collagen type I alpha 1
COL1A2	Collagen type I alpha 2
Cys	Cysteine

D.

DCl	Deuterium chloride
DFA	Discriminant function analysis
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate

D_2O	Deuterium oxide

E.

ECM	Extracellular matrix
ENAM	Enamelin
ePTMs	Enzymatic post-translational modifications

F.

(+)-FLEC	(+)-1-(9-fluorenyl)ethyl chloroformate
(-)-FLEC	(-)-1-(9-fluorenyl)ethyl chloroformate

G.

Gln	Glutamine
Glu	Glutamic acid
GluR	Glutamic acid racemase
Glx	Glutamine or glutamic acid
Gly	Glycine

H.

HCl	Hydrochloric acid
His	Histidine
HisR	Histidine racemase
Hyl	Hydroxylysine
Hyp	Hydroxyproline
H_2O_2	Hydrogen peroxide

I.

IEC	Ion-exchange chromatography
Ile	Isoleucine

L.

LC	Liquid chromatography
LC-MS	Liquid chromatography coupled to mass spectrometry
Leu	Leucine
LLPs	Long-lived proteins

Lys	Lysine
LysR	Lysine racemase

M.

Met	Methionine
micro-CT	Micro-computed tomography
MMP-20	Matrix metalloproteinase-20
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry

N.

nanoLC-MS	Nano-liquid chromatography coupled to mass spectrometry
NEM	*N*-ethyl maleimide
NMR	Nuclear magnetic resonance
NPEM	(*R*)-(+)-*N*-(1-phenylethyl) maleimide
nPTMs	Non-enzymatic post-translational modifications

O.

OTPTHE	*N*-[1-oxo-5-(triphenylphosphonium)pentyl]-(*R*)-1,3-thiazolidinyl-4-*N*-hydroxysuccinimide ester bromide salt

P.

PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pH	Potential hydrogen
Phe	Phenylalanine
Pro	Proline
ProR	Proline racemase
PTMs	Post-translational modifications

R.

(*R*)-BiAc	(*R*)-4-nitrophenyl-*N*-[2-(diethylamino)-6,6-dimethyl-[1,1-biphenyl]-2-yl] carbamate hydrochloride
RP	Reversed phase

S.

(*S*)-BiAc	(*S*)-4-nitrophenyl-*N*-[2-(diethylamino)-6,6-dimethyl-[1,1-biphenyl]-2-yl] carbamate hydrochloride
(*S*)-NIFE	(*N*)-(4-nitrophenoxycarbonyl)-L-phenylalanine-2-methoxyethyl ester
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SEM	Scanning electron microscope
Ser	Serine
SerR	Serine racemase
SLPs	Short-lived proteins

T.

Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine

U.

UV	Ultraviolet

V.

Val	Valine

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.

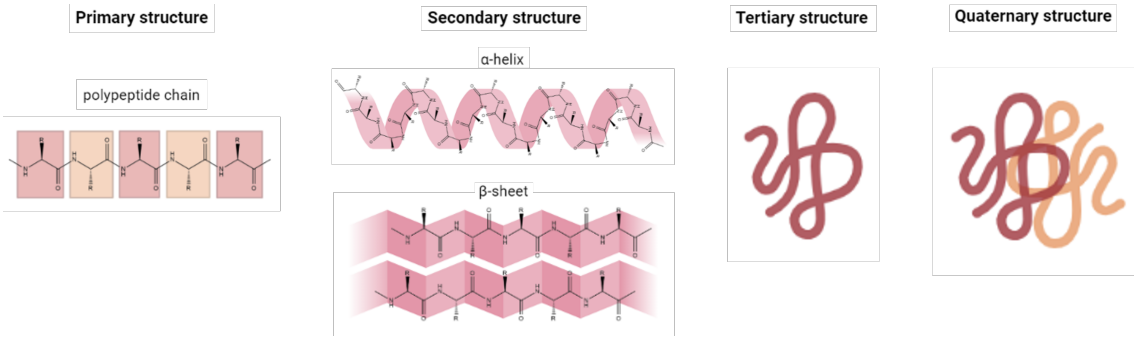


FIGURE 1.1: Structure of proteins
Created with BioRender.com

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].

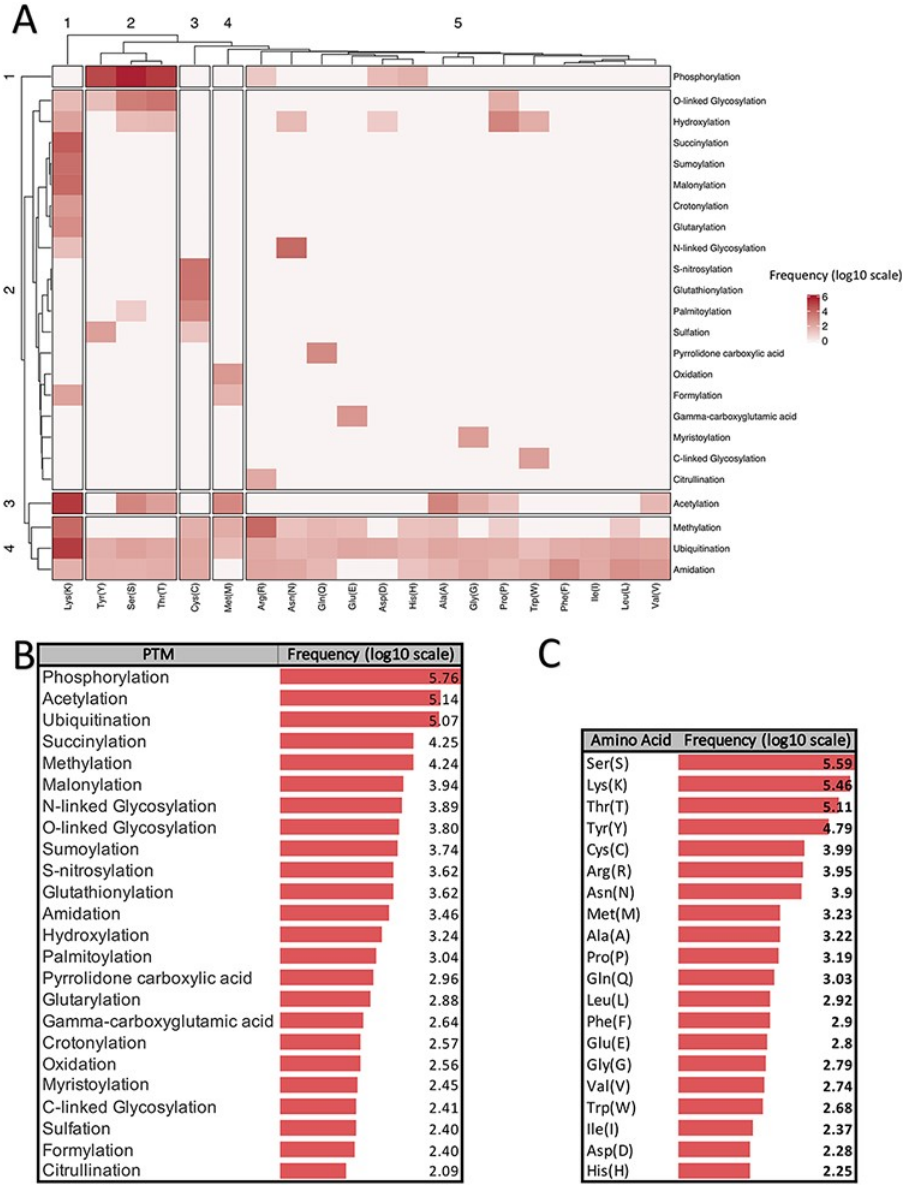


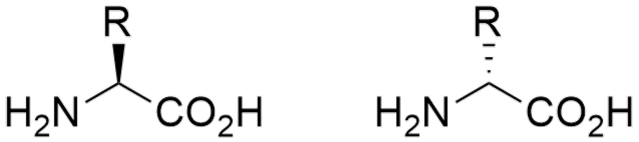
FIGURE 1.2: Major post-translational modifications reported

(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.



L-amino acid D-amino acid

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 biological materials

1.3.1 Proteomics

Proteomics is the analysis of proteins, including their resulting peptides, proteinogenic amino acids, and their modifications. Liquid chromatography coupled to mass spectrometry (LC-MS) is the most popular technique used for proteomics. In LC-MS, two separation modes are possible: normal and reversed phases. Reversed phase is commonly used for protein and peptide analysis using a chromatographic column that contains a non-polar stationary phase, such as the octadecyl chains column (C18). Proteomics is particularly used for the characterization, identification, function determination, localization, separation, and quantification of proteins or complex protein mixtures. The main applications are the discovery of disease biomarkers at different stages of disease and drug targets. In proteomics, bottom-up and top-down are the two traditional approaches (Figure 1.4). The bottom-up approach analyzes cleaved peptides after the enzymatic treatment of purified proteins or complex protein mixtures. Peptide mass spectra obtained after MS and MS/MS analyses were compared to the theoretical peptide masses calculated from proteomic databases. This comparison allowed for the identification of the peptide sequence and its position in the protein sequence. The top-down approach, which is less commonly used, allows the analysis of intact protein molecular ions and their fragments obtained during MS/MS analysis. This analytical method also allowed us to identify proteins after database comparison. Using both approaches, quantification and post-translational modifications (PTMs) mapping can be determined.

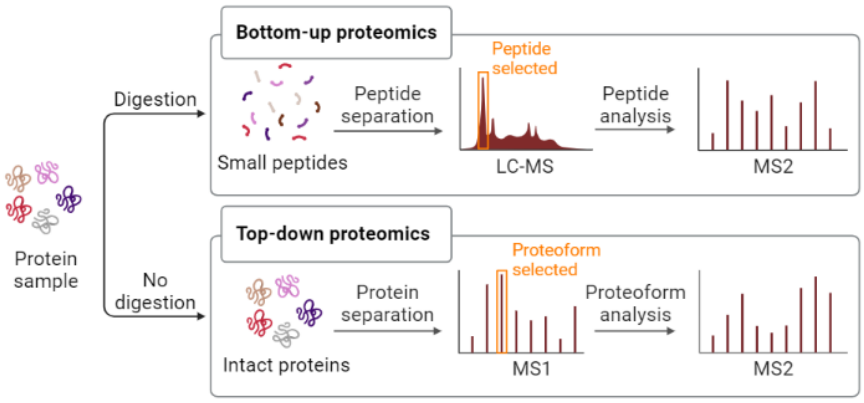


FIGURE 1.4: Bottom-up and top-dow proteomics
Created with BioRender.com

1.3.2 Chromatographic analysis

In chromatography, analytes are separated according to their different affinity for the stationary phase. Although the reversed phase (RP) mode is the most widely used in liquid chromatography (LC) for the separation of proteins and peptides, other different techniques can be used.

1.3.2.1 Affinity chromatography

Affinity chromatography allows for the separation of peptides and proteins based on their specific interactions with an immobilized ligand. Columns can be functionalized with different types of ligands [14] and used for different types of proteins [15].

1.3.2.2 Ion-exchange chromatography

Ion-exchange chromatography (IEC) separates molecules based on their charge. IEC columns can be functionalized with anions or cations of different natures for protein separation [16]. The IEC column can also be coupled to the RP column in two-dimensions for proteomics analysis [17, 18], or in three-dimensions coupled to hydrophobic interaction and RP chromatographies [19].

1.3.2.3 Size-exclusion chromatography

Size-exclusion chromatography (SEC) allows the separation of biological materials according to their overall size. Different columns with pores of a specific size can be selected based on the molecular size of peptides, proteins [20], and protein aggregates [21]. The association of successive SEC in series can be coupled to RP chromatography for proteomic applications. This two-dimensional coupling can achieve high-resolution separation of proteins in complex mixture [22, 23] and protein aggregates [23].

1.3.3 Electrophoretic analysis

1.3.3.1 Capillary electrophoresis

In capillary electrophoresis (CE), analytes are separated into fused-silica capillaries according to their isoelectric point. Because of its orthogonality to LC, CE can also be used for the analysis of amino acids [24], peptides [25], and proteins [25, 26], and can be coupled to mass spectrometry [27] for peptide mapping [28]. The advantage of CE over LC is its low buffer and sample consumption, which is useful for biomarker discovery [29]. However, the main problem encountered when using a fused-silica capillary for peptide and protein separations is the adsorption of these analytes on the capillary surface. To prevent adsorption, a capillary coating was added to the capillary wall. These coatings can be mono- or multilayer [30] and MS-compatible [31]. Additional advantages of capillary coating are separation performances including separation efficiency, intra- and inter-capillary repeatability and reproducibility [32–34].

1.3.3.2 Gel electrophoresis

Gel electrophoresis can also be used to separate proteins. Two mechanisms for protein separation are possible, according to their isoelectric points or molecular masses (Figure A.2). In native-PAGE, a pH gradient in the gel was formed using an applied high electric potential. This pH gradient allows for the separation of proteins according to their isoelectric points. In SDS-PAGE, a moderate voltage is sufficient to separate proteins based on their molecular mass. These two methods can also be combined to separate proteins into two dimensions.

1.4 Analysis of ancient materials

1.4.1 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.

1.4.2 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 [35]. 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.

Further details and a comparison of enantioseparation and derivatization techniques are gathered in the following publication and highlighted on the journal cover page.

Recent advances in chiral analysis of proteins and peptides
Marine Morvan, Ivan Mikšík
Separations, **2021**, *8*, 112.

Determination of Protein and Peptide chirality
Marine Morvan, Ivan Mikšík
Separations, **2021**, *8*, cover page.

Review

Recent Advances in Chiral Analysis of Proteins and Peptides

Marine Morvan [1,2,*] and Ivan Mikšík [1,2,*]

1 Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20 Prague 4, Czech Republic

2 Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 532 10 Pardubice, Czech Republic

* Correspondence: Marine.Morvan@fgu.cas.cz (M.M.); Ivan.Miksik@fgu.cas.cz (I.M.)

Citation: Morvan, M.; Mikšík, I. Recent Advances in Chiral Analysis of Proteins and Peptides, *Separations* **2021**, *8* (8), 112. https://doi.org/10.3390/separations8080112

Abstract: Like many biological compounds, proteins are found primarily in their homochiral form. However, homochirality is not guaranteed throughout life. Determining their chiral proteinogenic sequence is a complex analytical challenge. This is because certain D-amino acids contained in proteins play a role in human health and disease. This is the case, for example, with D-Asp in elastin, β-amyloid and α-crystallin which, respectively, have an action on arteriosclerosis, Alzheimer's disease and cataracts. Sequence-dependent and sequence-independent are the two strategies for detecting the presence and position of D-amino acids in proteins. These methods rely on enzymatic digestion by a site-specific enzyme and acid hydrolysis in a deuterium or tritium environment to limit the natural racemization of amino acids. In this review, chromatographic and electrophoretic techniques, such as LC, SFC, GC and CE, will be recently developed (2018–2020) for the enantioseparation of amino acids and peptides. For future work, the discovery and development of new chiral stationary phases and derivatization reagents could increase the resolution of chiral separations.

Keywords: chiral separation; proteins; peptides; D-amino acids

1. Introduction

Homochirality is omnipresent in biological processes and is essential for the development and maintenance of life. The phenomenon is well known in saccharides when mono-, di-, oligo- and polysaccharides are found in their D-form. Fructose, galactose and glucose are natural monosaccharides found in fruits and vegetables. Lactose, maltose and sucrose are natural disaccharides are made up of galactose and glucose, two glucoses, and glucose and fructose, respectively. They are found in mammalian milk (lactose), malted grains (maltose) and plants, fruits and vegetables (sucrose). Each of them plays a role in human health and diseases. Indeed, these sugars are essential and provide the energy necessary for the proper organ function. Nevertheless, excess blood sugar can cause metabolic disorders such as diabetes [1,2]. Oligo- and polysaccharides are also found in vegetables, fruits and grains. These saccharides could offer a promising hypoglycemic potential, without side effects on the human body [3]. Oligo- and polysaccharides are naturally present as glycosylation on protein sequences. These post-translational modifications can play a physiological role in the human body. The change of nature of glycosylation on protein sequences can modulate inflammatory responses, allow viral immune escape, promotes the onset of cancer cell metastasis and regulate apoptosis [4]. Natural or synthesized drugs are other important source of chiral compounds

necessary for life. Their (R)- and (S)-enantiomers can be beneficial, neutral or toxic for human life, with different pharmacology and pharmacokinetics, which is why their enantioseparation is an important bioanalytical challenge. Several separation techniques can be used, including liquid chromatography, gas chromatography, supercritical fluid chromatography and capillary electrophoresis [5].

Not so commonly described as chiral, proteins and peptides are found mainly in their L-amino acid form. Properties of each amino acid in peptides and proteins can play a role in the global properties and function of the peptides and proteins they form. The L- and D-amino acids have very similar chemical and physical properties but differ in their optical character. Peptides containing L-amino acids would be in the α-helical peptide by the left-handed helix rotation, and hypothetically all peptides containing D-amino acids could be in the α-helical peptide by the right-handed helix rotation [6]. The spatial architecture of L-peptides and L-proteins allows them to play an important role in enzymatic specificity and structural interaction. D-peptides and D-proteins are also biostable to proteolytic enzymes [7]. When a D-amino acid appears in a L-peptide or L-protein sequence, the orientation of the amino acid side chain is reversed [8]. This inversion can change the peptide or protein properties, such as affinity for solvents and their interaction with other proteins [9]. However, the reasons for the elimination of D-amino acids in all living organisms composed mainly of L-amino acids are not well known [10]. Nevertheless, thanks to recent technological advances, some D-amino acids in peptides and proteins have been detected in various living organisms, including humans, and they have been found to be generated by enzymatic or non-enzymatic post-translational isomerization, especially in elderly patients, diseased tissues and those with innate immune defense [11–16]. Therefore, peptides and proteins containing D-amino acids which are linked to various diseases can be used as novel disease biomarkers, but the amount of research and papers is still limited. Among some peptides containing D-amino acids, the most famous examples are the agatoxins, dermorphin and gramicidin S. Agatoxins are a class of neurotoxin peptides isolated from the Agelenopsis aperta spider venom. Specifically, the ω-agatoxin IV contains 48 amino acids on its sequence with a D-Ser residue at position 46. Its main biological action is to block exclusively the P-type calcium channels. It has no activity against T-type, L-type or N-type calcium channels present on the neuron membranes [17]. Dermorphin is a heptapeptide found in the skin secretions of two frogs Phyllomedusa rhodei and Phyllomedusa sauvagei and contains a D-Ala at the second position. Thanks to this D-Ala, dermorphin has an affinity and selectivity for μ-opioid receptors and has biological activity similar to morphine. However, this opioid activity is lost when the alanine is in its L-form [18]. Gramicidin S is a cyclodecapeptide containing two D-Phe with antibiotic activity against some bacteria, produced by the Gram-positive bacterium Bacillus brevis, and is used to treat wound infections [19].

The list of peptides containing D-amino acid discovered to date is summarized in Table 1, including the position of the D-amino acid. Proteinogenic D-amino acids may also be found in some proteins. Among the most famous examples, D-Asp has been found in the sequence of several proteins, such as elastin, myelin and β-amyloid; in the human aorta; and in the skin and brain tissue of patients with atherosclerosis and Alzheimer's diseases. Moreover, proteinogenic D-Asp, D-isoAsp, D-Asn, D-Glu, D-isoGlu, D-Ser and D-Thr have been found in the sequence of α- and β-crystallin contained in the human lens [12,20–23]. The isomerization of these amino acids under physiological conditions may change the higher-order structure of a protein, which in turn may have a role in age-related disorders such as cataracts [12]. Recently, Fujii et al. have proposed a relationship between protein aggregation and Asp isomerization, leading to the cataract formation [24].

Table 1. Position of D-amino acids in protein/peptide sequences.

D-amino Acids	Proteins/Peptides	Length of Amino Acid Sequence	Position on the Sequence	Ref.
D-Asp	Phosphophoryn	1129	undetermined	[25]
D-Asp	Elastin	786	undetermined	[26,27]
D-Ala	Ovalbumin	385	undetermined	[28]
D-Asp	Ovalbumin	385	undetermined	[28]
D-Glu	Ovalbumin	385	undetermined	[28]
D-Pro	Ovalbumin	385	undetermined	[28]
D-Ser	Ovalbumin	385	undetermined	[28]
D-Asp	Myelin	304	145	[29]
D-isoAsp	Myelin	304	34, 145	[29]
D-Asp, D-isoAsp	*β*B1-crystallin	252	211	[23]
D-Asp	IgGK light chain	214	151, 170	[15]
D-Asp	*β*B2-crystallin	205	4	[12]
D-Asp	*α*B-crystallin	175	36, 62, 140, 143	[11,30]
D-Ser	*α*B-crystallin	175	59, 66	[23]
D-Asn	*α*B-crystallin	173	undetermined	[21]
D-Asp	*α*B-crystallin	173	58, 84, 151	[20,23]
D-Glu, D-isoGlu	*α*B-crystallin	173	83	[23]
D-isoAsp	*α*B-crystallin	173	84	[23]
D-Ser	*α*B-crystallin	173	59, 62	[22]
D-Thr	*α*B-crystallin	173	undetermined	[21]
D-Tyr	Achatin-like neuropeptide	158	56, 86	[31]
D-Asp	Histone H2B	126	25	[32]
D-Asp	Osteocalcin	100	undetermined	[33]
D-Phe	Phenylseptin	67	50	[34]
D-Trp	*ω*-agatoxin IV	48	46	[17]
D-Asp	*β*-amyloid	42	1, 7, 23	[35]
D-Ser	*β*-amyloid	42	8, 26	[36]
D-Asp	IgG H5	27	24	[37]
D-Asp	IgG L2	24	12	[37]
D-*allo*-Ile	Brombinin H4	21	2	[38]
D-Phe	Gramicidin S	10	cyclopeptide	[39]
D-Asp	mAb heavy chain CDR2 peptide (51-59)	9	4	[40]
D-Phe	Hyperglycemic hormone	8	3	[41]
D-Trp	Contryphan	8	4	[42]
D-Ala	Dermorphin	7	2	[18]
D-Ala	Deltorphine	7	2	[43]
D-Met	Dermenkephalin	7	2	[44]
D-Asp	Type 1 collagen C-terminal telopeptide (1209-1214)	6	3	[45]
D-Asn	Fulicin peptide	5	2	[46]
D-Phe	Achatin I peptide	4	2	[47]

Amino acids can also be found in their free form in the human body and intervene in various diseases. Lee et al. are recently reported to free D-Ala has been found in the brain (white and gray matter, serum

and cerebrospinal fluid) for patients with Alzheimer's disease, in the plasma and serum for renal disease, in the serum for liver cirrhosis and hepatocellular carcinoma, in the urine for short bowel syndrome and in the nails for diabetes [48]. Free D-Asp and D-Trp are also found in chickens and mammals including humans and rats. The list of free D-amino acids discovered is summarized in Table 2 by Ayon et al., as well as D-amino acid-containing peptides and proteins, their location and associated roles [49].

Table 2. List of various D-amino acids in higher organisms, their location and associated roles. Reproduced with permission from the authors of [49].

D-Amino Acids	Proteins/Peptides/Free AA	Source	Associated Disease/Function	Ref.
D-Asp	Elastin	Aorta and skin (H)	Arteriosclerosis	[26,27]
	Myelin	Brain (H)		[50]
	β-amyloid	Brain (H)	Alzheimer's disease	[51]
	Free AA	Brain (H, R, C) Testis, adrenal and pineal glands (R)	Neuromodulatory effect Inhibit secretion of melatonin Increase testosterone production	[52–54]
D-Asp, D-Asn, D-Ser and D-Thr	*α*-crystallin	Lens (H)	Cataract	[21]
D-Ala	Dermorphin Deltorphine	Skin (F)	1000 times more analgesia than morphine due to presence of D-Ala	[18,43]
D-Met	Dermenkephalin	Skin (F)	Analgesia	[44]
D-Phe	Achatin I	Ganglia and atrium (S)	Enhances cardiac activity Excitatory action on muscles	[47]
	Hyperglycemic hormone	Sinus gland (L)	Increase glucose concentration in response to stress	[41,55]
	Phenylseptin	Skin (F)	Antimicrobial activity	[34]
D-*allo*-Ile	Brombinin H4	Skin (F)	Antimicrobial and antiparasitic activity	[38]
D-Asn	Fulicin peptide	Ganglia (S)	Enhance concentration of penis retractor muscle	[46]
D-Trp	Contryphan	Venom (CS)	Paralysis of fish prey by snails	[42]
	ω-agatoxin	Venom (SP)	Calcium channel blocker	[17]
	Free AA	Brain (M)	N-methyl D-aspartate (NMDA)/glycine receptor agonist	[52]

Notes: H: Human; R: Rat; C: Chicken; F: Frog; S: Snail; L: Lobster; CS: Cone Snail; SP: Spider; M: Mammals.

Many of these examples come from aging proteins and prove that the homochirality of amino acids in proteins is not guaranteed throughout life and can cause disorders in the human body as age progresses. Therefore, it is important to determine the protein sequence, in particular the position of the supposed D-amino acids.

2. Determination of Chiral Protein Sequences

Due to sequence modification, determining the exact protein sequences can be difficult. These modifications are caused via an enzymatic or a non-enzymatic process. Among these modifications, most frequently of them are the isomerization of amino acids and some post-translational modifications. Amino acid enzymatic isomerization is possible via an amino acid racemase. These racemases are classed into two families: pyridoxal 5'-phosphate-dependent (alanine racemase AlaR, arginine racemase ArgR, aspartate racemase AspR, histidine racemase HisR, lysine racemase LysR and serine racemase SerR), and pyridoxal 5'-phosphate-independent (aspartate racemase AspR, glutamate racemase GluR and proline racemase ProR) [16,56,57]. These racemases can proceed to free amino acid isomerization before or during peptide elongation [58]. On the other hand, the natural conversion of L-amino acid residues to its D-enantiomers in peptides and proteins occurs by non-enzymatic isomerization via a succinimidyl intermediate under physiological conditions. Figure 1 shows the spontaneous isomerization of L-Asp and the deamidation of L-Asn to D-Asp residues via an intramolecular cyclization. This conversion mechanism can also be applied from L-Glu and L-Gln to D-Glu residues. Certain specific amino acid sites seem more favorable to this isomerization and their position and steric environment on the sequence can influence this isomerization [20]. Besides, this non-enzymatic isomerization is associated with aging or disease in general [59]. A kinetic factor can also influence isomerization during ageing. Indeed, Hooi et al. have shown the percentage of racemization, under physiological conditions, in healthy and disease patients as a function of their age. For this study, the racemization of L-Asx, L-Ser, L-Thr, L-Phe, L-Glx and L-Leu was assessed. For healthy lens samples from 12-year-old patients, the L-amino acid racemization rates were 5.7%, 3.3%, 2.7%, 1.9% and 1.6%, respectively. For 80-year-old patients, these rates were 12.6%, 5.9%, 3.5%, 2.3%, 2.0% and 1.6%, respectively. Asx and Ser appear to be amino acids with faster isomerization throughout life. These two amino acids have also been studied in patients with cataracts. At 80 years of age, the racemization of L-Asx and L-Ser was 16.3% and 6.6%, respectively. Compared with healthy patients, these levels are significantly higher and confirm the influence of the disease on the increase in the racemization of amino acids contained in the proteins of the human body [21].

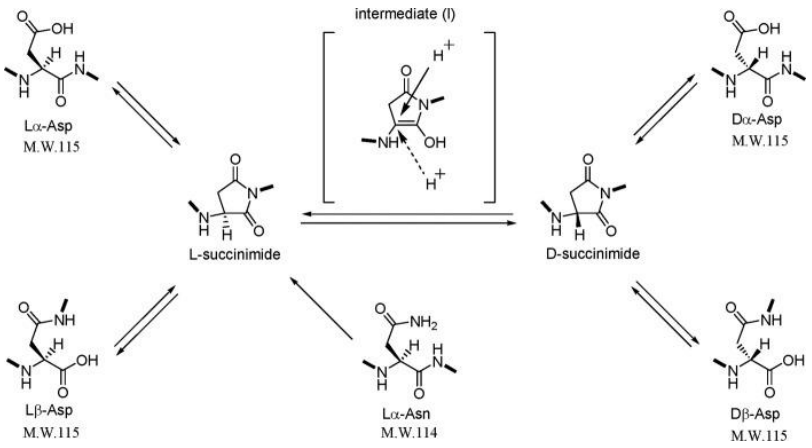


Figure 1. Possible reaction pathways for spontaneous isomerization of Asp and deamidation of Asn residues in protein. Reproduced with permission from the authors of [60].

However, in peptides and proteins, the conversion of L/D-amino is not uniform. Thus, it is necessary to examine each amino acid individually at each specific site susceptible to isomerization. Two strategies, as shown in Figure 2, can be used for the determination of chiral peptide and protein sequences: a sequence-dependent strategy and a sequence-independent strategy [14].

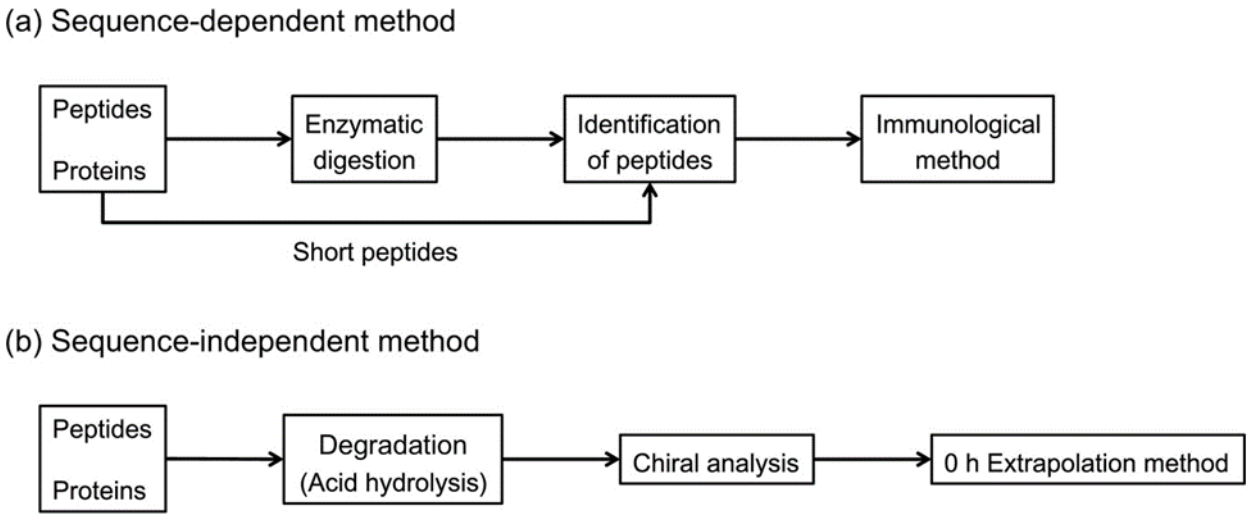


Figure 2. Flow charts for the determination of D-amino acid residues in peptides and proteins using (a) sequence-dependent and (b) sequence-independent analytical methods. Reproduced with permission from the authors of [14].

With the sequence-dependent strategy, i.e., peptide analysis, two methods can be applied to identify the isomerized position of the amino acid on the sequence. The first method consists of digesting the protein into several peptides by an enzyme. Each enzyme has its specific cleavage sites. Then, these peptides are separated by a chromatographic method, such as RP-HPLC, which is the most used. The collected peptide fractions are analyzed by mass spectrometry and their positions on the protein sequence are identified. These peptides are next hydrolyzed into amino acid residues and all react with a standard protected amino acid, such as Boc-L-cysteine [13] or Boc-L-Leucine [61]. This peptide

coupling makes it possible to obtain LL- and DL-dipeptides with different retention times and to distinguishes the L- and D-amino acids of the same mass. Nevertheless, the temperature, pH, metal ions and buffer can influence the efficiency of the peptide or protein enzymatic hydrolysis. Furthermore, recent studies have demonstrated the influence of chiral peptide sequence on enzymatic activity. Indeed, a D-amino acid on the specific cleavage site can prevent its recognition by the enzyme and block enzymatic activity. Moreover, the distance between a D-amino acid on the sequence and the specific site can also influence the degree of enzymatic cleavage [62]. The second method consists of homogenizing the protein in a natural buffer solution. After centrifugation, the water-soluble and water-insoluble fractions are collected. Both fractions are digested by an enzyme. Liquid chromatography coupled to mass spectrometry analysis is used for the separation of these digested peptides and their identification by sequencing coverage. Peptides containing different amino acid enantiomers have different retention times. Then, the retention time of each peptide is compared to that of standard synthesized peptides which several possibilities of enantiomer positions. With this approach, various enzymes or a combination of enzymes with different amino acid enantiomers at specific cleavage sites can be used. As an example, trypsin in association with endoproteinase Asp-N (cleavage at N-terminus L-Asp) [63], L-isoaspartyl methyltransferase (PIMT) (methylation of L-isoAsp) [14,32,35,37,63] and D-aspartyl endoproteinase (DEAP) also named paenidase (cleavage at C-terminus D-Asp) [64] can recognize only L-α-Asp, L-β-Asp and D-α-Asp residues respectively. Other proteases can be used to cleave proteins at specific sites, such as glutamyl endoproteinase Glu-C (cleavage at C-terminus L-Glu) for histone H2B [32] and human immunoglobulin G kappa chain [15] analyses.

The sequence-independent strategy, i.e., analysis of the content of whole amino acids in peptides and proteins consists of hydrolyze peptides and proteins, under acidic conditions and at high temperatures. D- and L-amino acids obtained were analyzed by enantioseparation. HCl/H2O conditions were more frequently used for the hydrolysis of biological material. However, in a hydrogenated environment, the natural conversion of L-amino acids to its enantiomer can take place. Kaiser et al. have shown the rate of racemization of free and proteinogenic amino acids by two different hydrolysis processes [65]. In vapor phase hydrolysis, the average rate of racemization over 11 amino acids for free amino acids was 3.44%, for amino acids of Bovine Serum Albumin (BSA) was 5.77%, and for amino acids of Lysozyme was 3.87%. The same liquid phase hydrolysis experiment showed that the average rate of racemization over 11 amino acids for free amino acids was 1.28%, for amino acids of BSA was 2.06%, and for amino acids of Lysozyme was 2.15%. To limit this natural conversion due to the presence of hydrogen, a DCl/D2O hydrolysis is frequently used [14,28,31,66]. Manning et al. also used a tritiated hydrochloric acid to correct for racemization during acid hydrolysis [67]. In a deuterated or tritiated environment, the hydrogen on the alpha carbon was exchanged with a deuterium or tritium atom and decreases considerably the racemization (Figure 3).

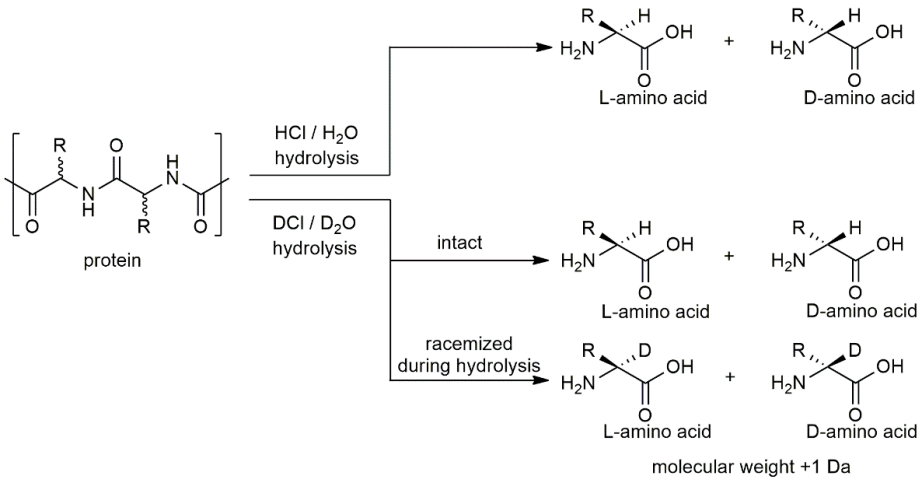


Figure 3. Amino acid racemization during protein hydrolysis under HCl/H2O and DCl/D2O conditions.

By combining these two previous strategies, another approach for determining the sequences of chiral proteins is possible. Indeed, Livnat et al. have reported the presence of two D-Tyr on GYFD and SYADSKDEESNAALSDFA peptides on achatin-like neuropeptide sequence from Aplysia californica. This approach takes place over three steps. First, neuropeptide was digested by an aminopeptidase M and resistant peptides were screening. Second, a DCl/D2O hydrolysis was used to hydrolyze each resistant peptide. Hydrolyzed amino acids were derivatized with Marfey's reagent and their chirality was determined by liquid chromatography coupled to mass spectrometry. Finally, when the presence of a D-amino acid was confirmed, the retention time of the endogenous peptide was compared to standard synthetic peptides. A schematic is shown in Figures 4 and 5 [31].

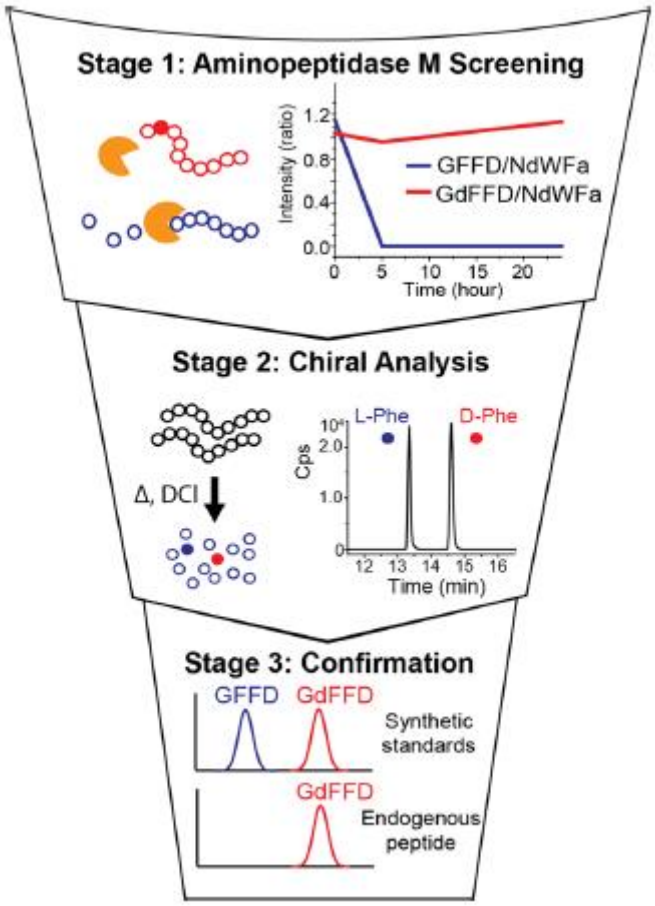


Figure 4. DAACP discovery funnel is capable of identifying DAACPs in three stages, as illustrated with GdFFD. In stage 1, MS-based detection of APM digestion is capable of identifying potential DAACPs in the screening phase of the discovery funnel. Here, GFFD, used as an example, is rapidly degraded after 5 h, whereas its DAACP counterpart, GdFFD, is not degraded after 24 h. Both are shown as a ratio to NdWF amide, a peptide that is known to resist degradation by APM. In stage 2, chiral analysis utilizes the MRM mode of MS to detect L- and D-amino acids in a peptide following acid hydrolysis and labeling. First, microwave-assisted vapor phase hydrolysis is carried out in DCl to break down peptides into their component amino acids. Importantly, DCl-based acid hydrolysis reduces detection of racemized residues in peptides. The amino acids are then labeled with Marfey's reagent, separated and detected using a triple quadrupole mass spectrometer. The result of this step is outlined using GdFFD, where a D-Phe is detected. In stage 3, confirmation of DAACPs, peptides are synthesized with the suspected chirality at each position and then compared to the endogenous peptides. Here, the retention time of the endogenous peptide matches that of the GdFFD synthetic standard, confirming its presence as a DAACP. Reproduced with permission from the authors of [31].

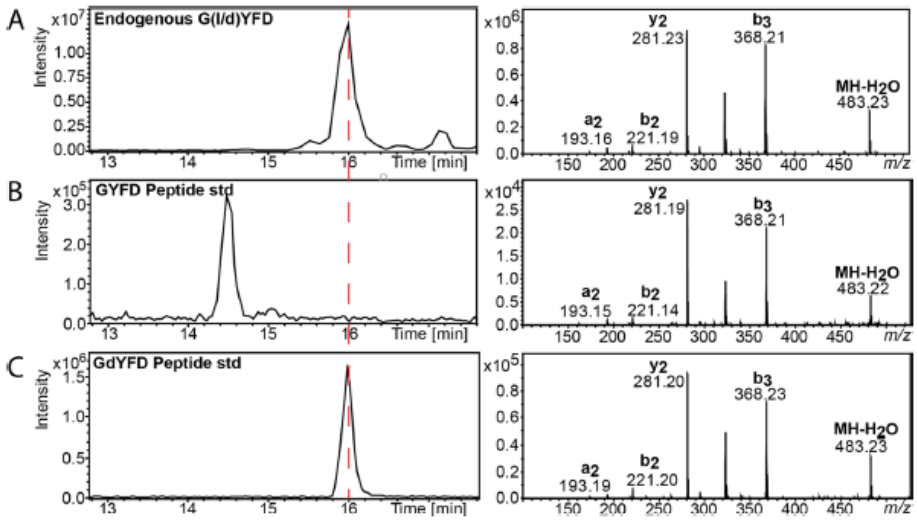


Figure 5. LC−MS/MS characterization of GdYFD, which is confirmed by comparing to the retention time of standards. (A) Left, LC−MS (base peak chromatogram) trace of endogenous GYFD after 48 h of APM digestion, with a retention time of 15.9 min. Right, the MS/MS fragmentation with fragment assignments is shown. (B) Left, an LC−MS trace of the all-L-amino acid synthetic GYFD, with a retention time of 14.5 min. Right, the MS/MS fragmentation with fragment assignments. (C) Left, an LC−MS trace of the synthetic DAACP GdYFD, with a retention time of 15.9 min. Right, MS/MS fragmentation with fragment assignments is shown. The matching retention time of the synthetic GdYFD standard with the endogenous GYFD demonstrates that the sequence for the endogenous peptide is in fact GdYFD. Reproduced with permission from [31].

3. Analysis of Chiral Amino Acids

The presence of D-amino acids is confirmed in several peptides and proteins with important biological functions. Improvements in chiral analytical methods now allow precise detection and quantification of these D-amino acids in complex biological materials. Chromatographic and electrophoretic techniques coupled to mass spectrometry are the most popular platforms, owing to advances in chiral stationary phase (CSP) and derivatization chemistry. Employing a derivatization reagent contributes to the accurate detection of trace levels of proteinogenic amino acid enantiomers in different matrices. The separation of chiral amino acids or even peptides is only possible by a chiral approach which can be created by two methods: the direct or indirect method, described below.

3.1. Direct Method

The direct method consists of separating the optical isomers using a chiral selector by one of two following techniques. First, chiral selectors can be attached due to a chemical spacer into the surface of the stationary phase or the capillary. This chiral selector has a different affinity with both enantiomers. This selectivity follows the three points Dalgliesh rule: three interaction points between the chiral selector and enantiomer [68], where all functional groups of the chiral selector, asymmetric centres and steric repulsion are involved. Second, chiral selectors can be added to the mobile phase or the background electrolyte, defined as chiral mobile phase additives (CMPA) [69]. In the case of ionic enantiomers, a chiral compound can add to the formation of pairs of ions with enantiomers. Otherwise, a stereospecific compound can add to interact with only one of the enantiomers. Many groups of chiral compounds can be utilized as chiral selectors for the enantioseparation of amino acids or peptides. The main of them are polysaccharides [70] and cyclodextrins. However, in this review, we will also evaluate

other chiral compounds as chiral selectors for the enantioseparation by liquid chromatography, supercritical fluid chromatography, gas chromatography and capillary electrophoresis.

The morphology of the chiral compounds used as a chiral selector can form three types of chiral stationary phase: fully-porous particles (FPP), superficially-porous particles (SPP) and non-porous particles (NPP) [71]. For the amino acid enantioseparation, the SPP columns appear to allow faster and more efficient separation than the FPP columns. Some examples will be described in the following section.

3.1.1. Liquid Chromatography

Enantioseparation by liquid chromatography needs a chiral environment. Several builder companies have developed commercial chiral columns. Zhao et al. have reported these columns with their target characteristics according to the type of chiral stationary phases and detailed are included in Table 3 [72].

Table 3. Target characteristics and commercial columns corresponding to chiral stationary phases (CSPs). Reproduced with permission from the authors of [72].

Types of CSPs	Basic Material	Target Characteristics	Commercial Column
Polysaccharides	Amylose or cellulose	Widely applicable, such as compounds containing amide groups, aromatic ring substituents, carbonyl nitro, sulfonyl, cyano, hydroxyl, amino and other groups, and amino derivatives	Chiralcel®OD, Chiralpak®IB, Lux®Amylose-1, Lux®i-Amylose-1
Cyclodextrins	*β*-cyclodextin	Widely applicable, such as hydrocarbon compounds, sterols, phenol esters and their derivatives, aromatic amines, polyheterocycles	B-DEXTM225, Astec Cyclobond®, I 2000 RSP, LiChroCART®250-4, ChiraDex®
Proteins	Enzymes, plasma proteins, receptor proteins	Water-soluble medicine	Chiralpak®HAS, Resolvosil®BSA, Chiralpak®AGT
Crown ethers	Macrocyclic polyester	Amino acids, amino alcohools, primary amines	Crownpak®CR(+)/CR(-), Chirosil®RCA(+)/RAC(-)
Pirkle type	Amine, amino alcohol, amino acid derivative compounds, anthrone derivatives	Widely applicable, CSPs designed by analyzing the target	Whelk-O1®, ULMO®, Chirex®

Ion exchange type	Cinchona alkaloids, sulfamic acid	N-protected amino acid, N-protected amino group, sulfamic acid, amino phosphoric acid, aromatic carbonyl acid	Chiralpak®QN-AX, Chiralpak®QD-AX, Chiralpak®ZWIX(+), Chiralpak®ZWIX(-)
Macrocyclic glygopeptides	Avomycin, Ristomycin A, Vancomycin, Teicoplanon and Teicoplanon aglycone	Widely applicable, such as amino acids, peptides, non-steroidal anti-inflammatory drugs	Astec® CHIROBIOTIC® V, Astec® CHIROBIOTIC® R, Astec® CHIROBIOTIC® TAG
Cyclofructans	Cyclic oligosaccharides	Primary amine, acid, secondary amine, tertiary amine, alcohol	Larihc® CF6-RN
Porous organic materials	Covalent organic framework, metal organic framework, metal organic cage, mesoporous silica	Halogenated hydrocarbons, ketones, esters, ethers, organic acids, alkylene oxides, alcohols and sulfoxides	/

Here, we will describe different chiral stationary phases, such as polysaccharides, cyclodextrins, crown ethers, Pirkle type, ion-exchange type and macrocyclic glycopeptides, for the chiral separation of amino acids and peptides.

3.1.1.1. Polysaccharide-based Chiral Stationary Phases

Chiral stationary phases based on polysaccharides are probably the most used owing to the many different columns for enantioseparations by liquid chromatography. Commercial cellulose- and amylose-based columns are most popular, however synthesized columns based on chitosan, xylan, curdlan, dextran and inulin may also be found [73]. Indeed, Lin et al. have evaluated 17 amylose and cellulose-based columns, both coated and immobilized, for chiral pharmaceutical analysis [74]. Different types of cellulose and amylose-based chiral selectors and their derivatives, the effect of substituents and carrier on properties of polysaccharide-based CSPs, and optimizations of the CSP and the mobile phase for separation of enantiomers by high-performance liquid chromatography (HPLC) have been summarized by Chankvetadze [75].

3.1.1.2. Cyclodextrin-based Chiral Stationary Phases

Cyclodextrins are frequently used for the separations of chiral compounds, as chiral stationary phases as well as novel polysaccharides. Many article reviews and book chapters generally describe α-, β-, γ- and modified-cyclodextrins for the separation of enantiomeric compounds by liquid chromatography [76,77]. More recently, Li et al. have described the synthesis of a triazole-bridged duplex CD-CSP and its application for the enatioseparation of Dns-Leu and Dns-Phe by HPLC [78]. Shuang et al. synthesized a new chiral stationary phase based on β-cyclodextrin for the enantioseparation of five derivatized and three underivatized DL-amino acid enantiomers by high-performance liquid chromatography. The results obtained with the new ethylenediamine dicarboxyethyl diacetamido-bridged bis(β-CD)-bonded

chiral stationary phase (EBCDP) were compared with those obtained with the native β-CD chiral stationary phase (CDSP). For each Dns-DL-amino acid enantiomer, the EBCDP has offered a better separation than the CDSP. The same trend was observed for the enantioseparation of underivatized DL-amino acids, with the addition of copper ion Cu (II) into the mobile phase, exhibiting a property of the chiral ligand exchange chromatography [79].

3.1.1.3. Crown Ether-Based Chiral Stationary Phases

Among the various crown ether-based columns, Crownpak CR-I (+) has been used in all recent studies. With this column, the retention time of D-amino acid was shorter than their L-amino acid enantiomers with a primary amino group [80,81]. For the separation of proteinogenic and non-proteinogenic amino acid enantiomers with the same molecular mass, Nakano et al. [82] have separated L-Ile, L-allo-Ile, L- and D-Leu while D-Ile and D-allo-Ile were co-eluted with the Crownpak CR-I(+) column. Using the Crownpak CR-I (−) column, D-Ile, D-allo-Ile and L-Leu were separated while L-Ile and L-allo-Ile were co-eluted (Figure 6) [82]. Meanwhile, Yoshikawa et al. have separated DL-Ile, DL-allo-Ile and DL-Leu with the Crownpak CR-I (+) column. Similar results were obtained for the enantioseparation of DL-Thr, DL-allo-Thr and DL-Hse [80]. Nakano et al. have separated L-Thr, L-allo-Thr and L-Hse while D-Thr, D-allo-Thr and D-Hse were co-eluted with the Crownpak CR-I (+) column. Using the Crownpak CR-I (−) column, D-Thr, D-allo-Thr and D-Hse were separated while L-Thr, L-allo-Thr and L-Hse were co-eluted (Figure 6) [82]. Yoshikawa et al., have separated only DL-Thr with the Crownpak CR-I (+) column, DL-allo-Thr and DL-Hse were not studied [80]. The exchange of columns Crownpak CR-I (+) and CR-I (−) reversed the elution order for each compound. However, for the secondary amino group, such as DL-Pro, the Crownpak CR-I (+) column did not offer separation while the Chiralpak ZWIX (−) column did (Figure 6) [82].

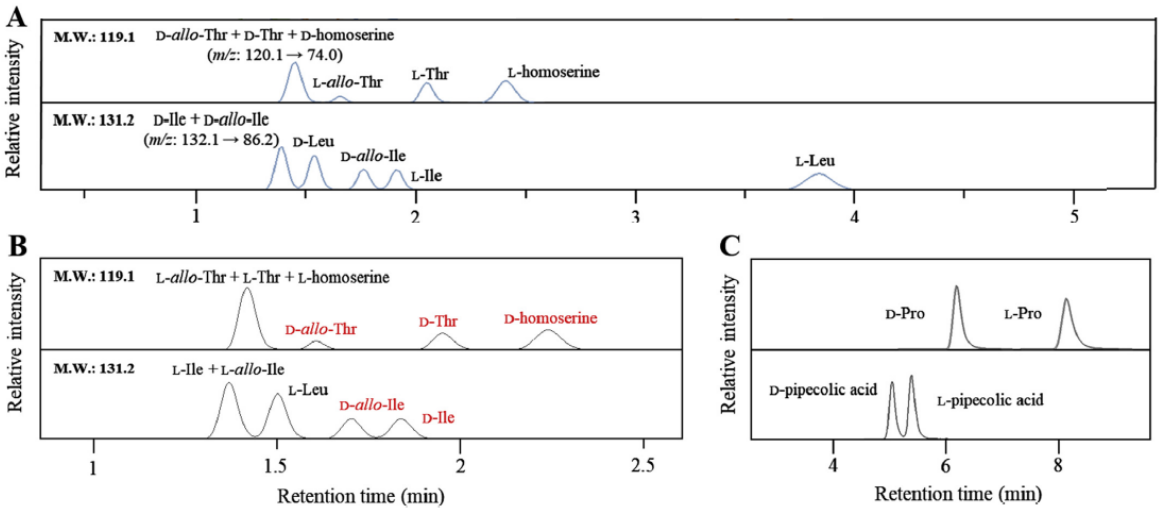


Figure 6. The chromatograms obtained by Crownpak CR-I (+), CR-I (−) and Chiralpak ZWIX (−). (A) Partial chromatograms of targeted compounds with the same molecular weight obtained by LC-MS/MS analysis using Crownpak CR-I (+). These chromatograms were obtained from a 1 nmol/mL mix standard solution (Thr, threonine; Ile, isoleucine; and Leu, leucine). The same color peaks were detected in the same MRM transition. Most compounds showed baseline separation while some compounds including D-allo-Thr, D-Thr, and D-homoserine or D-Ile and D-allo-Ile co-eluted. (B) Extracted ion chromatograms using Crownpak CR-I (−). The chromatogram for the MRM transition of m/z 120.1 to 74.0 (molecular weight: 119.1) and 132.1 to 86.2 (molecular weight: 131.2). The CR-I (−) column reversed the elution

order for each compound from the CR-I (+) column. (C) A part of chromatogram of secondary amines by using Chiralpak ZWIX (−). The ZWIX column, with zwitterionic molecules that incorporated both anion and cation-exchange functional groups, enabled the separation of secondary amines including proline (Pro). Adapted from the work in [82] with permission.

Concerning the enantioseparation of biological and synthesized peptides, the retention times of LXX or LXXX peptides were shorter than DXX or DXXX peptides using the Crownpak CR-I(+) column, whereas the migration order was reversed using the Crownpak CR-I(−) and ChiroSil RCA(+) columns. The chirality of the first amino acid seems to influence the retention time and different interactions and affinity with the column than its enantiomer [34,81]. In addition, the chirality of the second amino acid contained in the peptide sequence also influences the retention time. Indeed, peptides with a L-Phe (XLX) at the second position was more retained than peptide with a D-Phe (XDX). By using this column, the position of D-Phe in the tripeptide (Phe)3 from the amphibian antimicrobial DL-phenylseptin (Table 4) was elucidated [34]. The order of migration also changed according to the percentage of acetonitrile contained in the mobile phase [81]. In those cases, the Crownpak CR-I (+) column offered a better separation for a complex mixture of different peptide enantiomers.

Table 4. Comparison of peptide separation performances using the ChiroSil RCA (+), Crownpak (+) and Chrownpak (−) columns.

Samples	**Column**						**Ref.**
	ChiroSil RCA(+)		**Crownpak(+)**		**Crownpak(−)**		
Peptides	**Rs**	***α***	**Rs**	***α***	**Rs**	***α***	
LLL/DDD-$(Phe)_3$	-	-	-	5.58	-	1.28	[34]
LDL/DLD-$(Phe)_3$	-	-	-	-	-	-	
LLD/DDL-$(Phe)_3$	-	-	-	-	-	-	
LDD/DLL-$(Phe)_3$	-	-	-	-	-	-	
DLDL/LDLD-Tyr-Arg-Phe-Lys-NH_2	3.08	1.51	4.68	1.48			[81]
DDLL/LLDD-Tyr-Arg-Phe-Lys-NH_2	4.76	1.92	9.46	2.39			
DLLL/LDDD-Tyr-Arg-Phe-Lys-NH_2	3.51	1.61	10.62	2.56			
DLDD/LDLL-Tyr-Arg-Phe-Lys-NH_2	3.00	1.52	<0.50	1.04			
DDLD/LLDL-Tyr-Arg-Phe-Lys-NH_2	3.14	1.51	5.75	1.65			

Finally, for the enantioseparation of proteinogenic amino acids, the Crownpak CR-I (+) column with 3 µm particle size offered a better resolution than the same column with 5 µm particle size, especially for amino acids strongly retained. Indeed, for compounds that are poorly retained, such as DL-His, their resolutions were 1.2 and 1.5 with 5 µm and 3 µm particle size columns, respectively, and for DL-Lys which, was strongly retained, their resolutions were 4.1 and 6.1 with the 5 µm and 3 µm particle size

columns, respectively. Moreover, the separation time was reduced when the length of the column was shorter [80].

3.1.1.4. Brush-type or Pirkle-type Chiral Stationary Phases

Neutral synthetic chiral low-molecular-mass molecules are typically covalently linked to a silica support, which can have monosubstituted or trisubstituted silane groups through a spacer to create Pirkle-type chiral stationary phases [83,84]. The advantages of this type of column are the compatibility with a wide range of solvents used as mobile phase and stability against temperature and pressure [85]. Brush-type or Pirkle-type chiral stationary phases are classified into three groups: donors of π-electron/acceptor of π-electron (such as Whelk-O 1, Whelk-O 2, DACH-DNB and ULMO columns), acceptors of π-electron (such as Alpha burke 2, Leucine, Phenylglycine, Pirkle 1-J, β-Gem 10), and donors of π-electron (such as Naphtylleucine). Kohout et al. have studied the influence of the nature of Brush-types CSP in combination with a 9-O-tert-butylcarbamoylquinidine as a chiral selector. In that work, three different FPP—Daiso silica (CSP 1, 5 µm, 120 Å), Kromasil silica (CSP 2, 5 µm, 100 Å) and YMC silica (CSP 3, 5 µm, 120 Å)—were used. The averages of resolutions obtained for the enantioseparation of eight different analytes were promising (5.24, 4.21 and 4.79 for CSP 1, CSP 2 and CSP 3, respectively). The study showed the influence of the characteristics (nature, particle diameter and pore size) of the CSP used on the chromatographic performances [86]. Hsiao et al. have separated the DL-Phe enantiomers contained in physiological fluids of mammals, including humans, by 2D-HPLC. The D- and L-Phe was derivatized by 4-fluoro-7-nitro-2,1,3-benzoxadiazole (NBD-F) and separated by a reverse phase column for the first dimension and a Pirkle-type column with L-Leucine as a chiral selector for the second dimension. The same study showed the presence of 0.1% and 3.99% of D-Phe in the human plasma and urine, respectively. However, the authors planned further studies using this analytical method on various disease samples [87].

3.1.1.5. Ion- and Zwitterion-exchange-based Chiral Stationary Phases

Bäurer et al. have demonstrated the chiral chromatography characteristics of Chiralpak ZWIX (+) and ZWIX (−) columns based on zwitterionic quinine and quinidine carbamate selectors for the enantioseparation of proteinogenic amino acids by tandem liquid chromatography. For this work, molecular dynamics simulations, lipophilicity/hydrophilicity measurements of selectors, pH-dependent ζ-potential determinations and chromatographic characterization were studied for both columns [88]. Geibel et al. have demonstrated the Chiralpak ZWIX column with superficially-porous particles (SPP) allowed a better separation performance than a Chiralpak ZWIX column with fully-porous particles (FPP) [89].

Different anion- and zwitterion-exchange columns such as Chiralpak QN-AX, Chiralpak QD-AX, KSAAAX, Chiralpak ZWIX(+) and Chiralpak ZWIX(-) have been recently used for the enantioseparation of DL-amino acids and peptides, and were applied to biological samples in the discovery of new potential biomarkers for some diseases. Among these columns, Chiralpak QN-AX and Chiralpak ZWIX (+) have been used more, either alone or in tandem. With these two columns, free derivatized D-amino acids or D-amino acid-containing peptides at the first position were less retained on the column than their L-amino acid enantiomers. In contrast, with the Chiralpak QD-AX and Chiralpak ZWIX (-) columns, the migration order was reversed.

Zhu et al. used the Chiralpak QN-AX column for the enantioseparation of 13 Fmoc-DL-amino acids. The chromatographic performances (resolution and separation factor α) were compared to those obtained thanks to new synthesized chiral selectors based on 3,5-dimethylphenyl carbamoylated β-cyclodextrin connecting quinine (QN) or quinidine (QD) to create two chiral stationary phases (β-CD-QN- and β-CD-QD-based CSP). The Chiralpak QN-AX column allowed separation of 11 out of 13 Fmoc-DL-amino acids

with a resolution greater than 1.5, including seven Fmoc-DL-amino acids which had the best resolution (Asn, Thr, Ile, Leu, Met, Val and Phe) to compare to these two other experimental columns [90]. Bajtai et al. have applied this column to the enantioseparation of 20 dipeptide couples, but only seven of them obtained a good separation (LL/DD-Ala-Phg, LL/DD-Ala-Phe, LL/DD-Ala-Tyr, LL/DD-Ala-4-NO2-Phe, LL/DD-Ala-Trp, LL/DD-Phe-Ala and LL/DD-Leu-Leu). Nevertheless, 10 dipeptide couples have not been separated (LL/DD-Ala-Ala, LD/DL-Ala-Ala, LL/DD-Ala-β-Phe, LL/DD-Ala-Phe, LL/DD-Ala-hPhe, β-Ala-L/D-Phe, Gly-L/D-Phe, LD/DL-Phe-Ala, LL/DD-Lys-Phe and LD/DL-Leu-Leu) [91] (Table 5)

Table 5. Comparison of peptides separation performances using the ion-exchange (QN-AX and QD-AX) and zwitterionic (ZWIX (+) and ZWIX (-)) columns.

Samples	**Column**								**Ref.**
	Ion-exchange				**Zwitterionic**				
	QN-AX		**QD-AX**		**ZWIX(+)**		**ZWIX(-)**		
Peptides	**Rs**	***α***	**Rs**	***α***	**Rs**	***α***	**Rs**	***α***	
LL-Ala-Ala	0	1.00	1.07	1.75	0.29	1.03	1.24	1.35	
LD-Ala-Ala	0	1.00	0	1.00	0.97	1.14	0.21	1.03	
LL-Ala-Phg	2.35	1.46	2.44	1.52	0.94	1.13	1.15	1.17	
LL-Ala-*β*Phe	0	1.00	0	1.00	0.90	1.09	0	1.00	
LL-Ala-Phe	1.82	1.25	2.43	1.40	1.03	1.13	0.80	1.12	
LD-Ala-Phe	0	1.00	0	1.00	0.53	1.12	1.29	1.17	
LL-Ala-hPhe	0	1.00	1.56	1.30	0	1.00	0.90	1.20	
*β*Ala-L-Phe	0	1.00	1.05	1.15	4.36	1.45	2.50	1.34	
LL-Ala-Phe-OMe	-	-	-	-	0	1.00	0	1.00	
LL-Ala-Phe-NH_2	-	-	-	-	0	1.00	0.55	1.12	[91]
LL-Ala-Tyr	2.36	1.34	2.47	1.48	1.52	1.21	1.30	1.42	
LL-Ala-4-NO_2-Phe	1.88	1.25	0.82	1.09	1.96	1.28	2.00	1.46	
LL-Ala-Trp	2.36	1.51	2.35	1.58	7.09	2.21	3.50	1.96	
Gly-L-Phe	0	1.00	1.89	1.49	0	1.00	0.53	1.19	
L-Phe-Gly	0.63	1.14	1.52	1.28	3.21	1.33	2.00	1.26	
LL-Phe-Ala	4.36	1.52	4.73	2.01	2.55	1.55	2.94	1.40	
LD-Phe-Ala	0	1.00	0	1.00	2.71	1.44	0.97	1.17	
LL-Lys-Phe	0	1.00	2.14	1.71	1.95	1.28	0.94	1.12	
LL-Leu-Leu	3.20	2.19	4.00	2.37	1.40	1.15	0.63	1.17	

LD-Leu-Leu	0	1.00	0	1.00	1.77	1.24	0.53	1.16	

The Chiralpak ZWIX (+) column has been used by several groups for the enantioseparation of derivatized or underivatized common and uncommon DL-amino acids and dipeptide couples. Indeed, Horak et al. have compared the separation factor α for the enantioseparation of derivatized or underivatized common and uncommon DL-amino acids. This study has shown to the Chiralpak ZWIX (+) column offers a better separation to ACQ-DL-amino acids than underivatized amino acids, except for DL-Pro [92,93]. For this work, the authors do not study the DL-Cys. Nevertheless, Pucciarini et al. have developed a method for the separation of standard ACQ-DL-Cys, with a good resolution (Rs: 2.7). Then, the authors have applied this cysteine analysis method to human lung adenocarcinoma A549 cells (ATCC, CCL-185) samples, and the resolution obtained was similar to cysteine standards. The study allows showing the presence of D-Cys in biological samples [94]. These results are promising because it has been shown that D-Cys protects neurons against oxidative stress thanks to its hydrogen sulfide production properties and promotes dendritic development [95]. For the enantioseparation of underivatized dipeptides, Bajtai et al. have obtained a similar average separation factor α to the enantioseparation of underivatized amino acids in the study by Horak's group describe earlier. However, for some dipeptide, the authors have obtained good resolutions (β-Ala-L/D-Phe, LL/DD-Ala-Tyr, LL/DD-Ala-4-NO2-Phe, LL/DD-Ala-Trp, L/D-Phe-Gly, LL/DD-Phe-Ala, LD/DL-Phe-Ala, LL/DD-Lys-Phe and LD/DL-Leu-Leu) [91] for details see Table 5.

Kimura et al. used these two Chiralpak QN-AX and Chiralpak ZWIX (+) columns by chiral tandem liquid chromatography. For this study, the authors found the same enantiomer migration order (EMO) with both columns to one of these columns alone, as in the previous works described above. For amino acids with a primary amine, ACQ-D-amino acids were a faster elution than their counterpart ACQ-L-amino acids, except for DL-Pro which is a secondary amine and which has a reverse EMO. The results of resolution and separation factor α of the simultaneous 24 L-amino acids and trace D-Amino acids, which includes all proteinogenic amino acids and non-proteinogenic amino acids such as citrulline (Cit) and ornithine (Orn), contained in the human blood were not described. In this study, the blood of 305 women aged 65 to 80 who passed a Mini-Mental State Examination (MMSE), and classified into three groups: control, Mind Cognitive Impairment (MCI) and dementia was analyzed. This high sensitivity analytical method can detect a slight difference in the number of chiral amino acids and be developed to examine cognitive decline based on the difference detected. Indeed, the blood analysis of patients in the MCI group showed a higher proportion of D-Pro than the control group. Furthermore, the combination of D-Pro x D-Ser proportions improved the correlation with early cognitive decline described by MMSE. Finally, D-Pro and D-Ser were also found in the blood of the dementia group with a higher proportion than in the MCI group. These amino acids can also be used to monitor dementia. These results showed that the proportion of D-Pro in the blood was strongly associated with early cognitive decline and may represent a new index of brain health. This tandem method of analysis, in less to 40 min, might apply to practical medical evaluations to monitor the early risk of cognitive decline with high precision [96].

The Chiralpak ZWIX (−) has also been used for the enantioseparation of derivatized or underivatized DL-amino acids and dipeptide couples. For underivatized DL-amino acids, separation factors α obtained were better using the Chiralpak ZWIX (−) when compared to the Chiralpak ZWIX (+) column, except for Trp and Glu [92]. The same group also compared both columns for the enantioseparation of ACQ-DL-amino acids. Conversely, the Chiralpak ZWIX (+) column allowed to obtain better separation factors α than the Chiralpak ZWIX (−), except for Pro, aThr, His and Arg [93]. Bajtai et al. have also used Chiralpak ZWIX (+) to compare both zwitterionic columns for the enantioseparation of 20 dipeptide couples. The

separation factors α and the resolutions have shown that the Chiralpak ZWIX (+) column offered a better separation [91] see Table 5.

The Chiralpak QD-AX column has only been used by Bajtai et al. to compare the Chiralpak QN-AX column for enantioseparation of dipeptide couples [91]. The separation factors α and the resolutions have shown that the Chiralpak QD-AX column offered a better separation significantly (Table 5).

To compare these two zwitterion-exchange columns, the average separation factor α was better with the Chiralpak ZWIX (+) column for underivatized and ACQ-DL-amino acids (1.21 and 2.33, respectively) than with the Chiralpak ZWIX (−) column (1.29 and 2.03 respectively). In addition, the Chiralpak ZWIX (+) also offered a better dipeptide enantioseparation with an average resolution of 1.66 compared to 1.20 with the Chiralpak ZWIX (−) column. For these two anion-exchange columns, the Chiralpak QD-AX column offered a better separation of dipeptide enantiomers with a resolution average of 2.12 compared to 1.05 with the Chiralpak QN-AX column. Regarding the Fmoc-DL-amino acids enantioseparation, only the Chiralpak QD-AX column was used and offered an excellent resolution average of 2.94. Both Chiralpak ZWIX (+) and Chiralpak QD-AX columns are to be preferred for the derivatized and underivatized amino acids and small peptides enantioseparation.

Concerning the anion-exchange column KSAAAX, Furusho et al. have separated four NBD-DL-amino acids contained in the plasma of 25 patients with chronic kidney disease (CKD) using a 3D-HPLC system. Several studies agree on the potential correlation between kidney disease and the presence of D-amino acids [97]. The 3D-HPLC system equipped with one KSAAAX column in combination with a reversed-phase and enantioselective columns allows separation of all compounds according to their hydrophobicity, anionic strength and enantioselectivity, which increases the separation efficiency. For this study, DL-Ala, DL-Asn, DL-Ser and DL-Pro were selected because their D-forms are candidates for the biomarkers of CKD. This method shows the presence of 0.4–4.8% for D-Asn, 1.5–16.6% for D-Ser, 0.3−11.6% for D-Ala and 0.3−7.1% for D-Pro, and the correlation between the %D values and the eGFR values were observed. The eGFR (estimated Glomerular Filtration Rate) is the best test to measure the level of kidney function and determine the stage of kidney disease. Indeed, these results showed that the more advanced the stage of the disease, the higher the percentage of amino acids in their D-forms. The new 3D-HPLC system has proved to show high selectivity for the chiral analysis of amino acids contained in complex biological matrices as a biomarker for the determination of the disease at different stages [98].

3.1.1.6. Macrocycle Antibiotic-based Chiral Stationary Phases

The complex structure of macrocyclic antibiotics allows for different types of analyte interaction, such as hydrophobic, π–π, dipole-dipole, hydrogen bond, electrostatic, ionic and Van der Waals interactions, due to the high number of stereogenic centres in their structures. Macrocyclic antibiotic-based chiral stationary phases are classed into four groups: ansamycins, polypeptides, glycopeptides and aminoglycosides [84,99]. Recently, a new type of macrocycle antibiotic-based chiral stationary phase was described. Indeed, for this work, two new CSPs, named UHPC-FPP-Titan-Tzwitt and UHPC-SPP-Halo-Tzwitt, were prepared by covalently bonding the glycopeptide teicoplanin chiral selector on zwitterionic columns with different particle diameters (fully- and superficially-porous particles), for the enantioseparation of 31 proteinogenic and non-proteinogenic Fmoc-DL-amino acids by liquid chromatography. The performance in terms of separation obtained by both zwitterionic-teicoplanin columns was compared to those obtained from a commercial teicoplanin-based column (Teicoshell) (Table 6). For 26 pairs, the UHPC-SPP-Halo-Tzwitt column offered higher resolutions of separation ranging from Rs (DL-Asn-(Ttr)): 1.15 to Rs (DL-Lys-(Boc)): 10.90. Complementarily, the UHPC-FPP-Titan-Tzwitt column offered the best enantioresolution for Gln and His. Only Cys and Trp-(Boc) enantiomers

were better separated with the traditional Teicoshell column. However, DL-Pro and DL-Asp were not separated by any of the three columns. These results prove the influence of the zwitterionic column in combination with a teicoplanin chiral stationary phase. Indeed, both UHPC-FPP-Titan-Tzwitt and UHPC-SPP-Halo-Tzwitt columns offered significantly better resolutions than a commercial teicoplanin-based column. The particle diameters also seem to influence the enantioseparation. Indeed, the morphology differences of fully- and superficially-porous particles (FPP and SPP, respectively) affect the enantioselectivity and the resolution, in favour of the SPP column [100].

Table 6. Comparison of Fmoc-amino acid separation performances using the UHPC-FPP-Titan-Tzwitt, UHPC-SPP-Halo-Tzwitt and Teicoshell columns on MeOH-rich mobile-phase condition.

Samples		**Column**						
		UHPC-FPP-Titan-Tzwitt		**UHPC-SPP-Halo-Tzwitt**		**Teicoshell**		**Ref.**
Derivatization	**DL-amino acids**	**Rs**	***α***	**Rs**	***α***	**Rs**	***α***	
Fmoc	Ala	8.06	1.69	9.35	2.02	4.44	-	[100]
	Arg	4.74	2.09	5.26	2.44	2.01	1.32	
	Arg-(Pbf)	6.03	1.72	7.47	2.03	3.35	-	
	Asn	2.41	1.13	2.93	1.17	-	-	
	Asn-(Trt)	-	-	1.15	1.07	-	-	
	Asp	-	-	-	1.02	-	-	
	Asp-(OtBu)	1.95	1.11	2.09	1.16	-	-	
	Cys	1.14	1.14	1.68	1.16	1.82	-	
	Cys-(Trt)	1.59	1.10	2.54	1.16	-	-	
	Gln	7.99	1.30	4.71	1.39	1.58	-	
	Gln-(Trt)	3.23	1.22	4.46	1.32	1.96	-	
	Glu	1.42	1.06	2.76	1.11	2.21	-	
	Glu-(OtBu)	4.28	1.49	5.99	1.64	3.34	-	
	His	3.21	1.30	3.14	1.36	1.44	1.46	
	His-(Trt)	4.45	1.35	5.61	1.53	3.05	-	
	Ile	5.33	1.43	6.43	1.70	2.18	-	
	Leu	6.37	1.48	7.83	1.72	2.78	-	
	Lys	3.49	1.66	4.26	2.06	2.98	1.31	
	Lys-(Boc)	8.93	1.82	10.90	2.25	4.55	-	
	Met	7.33	1.54	8.94	1.82	4.15	-	

Phe	4.56	1.30	5.45	1.43	2.23	-
Pro	-	-	-	-	-	-
Ser	2.92	1.15	3.92	1.21	2.21	-
Ser-(tBu)	4.30	1.27	5.54	1.41	2.35	-
Thr	2.92	1.18	3.74	1.27	1.65	-
Thr-(tBu)	1.16	1.107	1.75	1.11	-	-
Trp	3.07	1.19	3.91	1.29	1.97	-
Trp-(Boc)	-	-	-	-	2.93	-
Tyr	3.86	1.26	4.63	1.38	2.37	-
Tyr-(tBu)	4.78	1.60	7.56	1.82	4.16	-
Val	4.89	1.33	6.38	1.52	2.36	-

3.1.2. Supercritical Fluid Chromatography

One- and two-dimensional supercritical fluid chromatography (SFC) are new techniques developed which may be applied in chiral separation [101,102]. In this part, we will consider the polysaccharide- and crown ether-based chiral stationary phases recently used for the chiral separation of amino acids.

3.1.2.1. Polysaccharide-Based Chiral Stationary Phases

Recently, Jakubec et al. used nine different chiral columns based on celluloses and amyloses covalently immobilized (Chiralpak IB-3, YMC CHIRAL ART Cellulose SB, Chiralpak IA-3 and YMC CHIRAL ART Amylose SA) or coated (Trefoil CEL-1, Chiralpak OD-3, Lux-CEL1, AMY-1 and Chiralpak AD-3) to study their enantioseparation efficiency under different separation conditions by supercritical fluid chromatography. The best performing column was a Chiralpak OD-3 with methanol + 0.1% trifluoroacetic acid and 0.1% diethylamine combined additive separation conditions. The combination of alcohol and alkylamine as a co-solvent and a mobile-phase additive seemed beneficial for enantioseparation [103]. Indeed, Lipka et al. optimized their concentration (20–40% ethanol + 0.3–3.0% trimethylamine) according to the amino acid enantiomers to be separated using an immobilized cellulose-based CSP (i-Cellulose-5). This column allowed separating some enantiomers of underivatized common and uncommon amino acids (DL-Leu, DL-Ile, DL-Nle, DL-Phe, DL-Ala, DL-Val, DL-Nva, DL-Ser, DL-Car and DL-His) with resolutions ranging from 1.51 (DL-Ser) to 2.24 (DL-Val). This study is promising for pure enantiomers. However, separation conditions could be improved for the separation of a complex amino acid mixture [104].

3.1.2.2. Crown Ether-Based Chiral Stationary Phases

Miller et al. have developed a method for the separation of DL-amino acids enantiomers with a Crownpak CR-I (+) column by supercritical fluid chromatography coupled to mass spectrometry. After optimization of the nature of the mobile phase, this column offered a resolution ranking from Rs (His): 1.99 to Rs (Leu): 9.26 and a short retention times of less than three minutes. With these conditions, all D-amino acid eluted faster to its L-amino acid enantiomer. This method of analysis makes it possible to

quickly separate the pure standards of DL-amino acids. However, it can be improved to separate a complex mixture of all of them [105].

3.1.3. Gas Chromatography

Gas chromatography is the most-established separation technique for the enantioseparation of amino acid enantiomers. Nevertheless, Schurig et al. reported some chiral stationary phases for the enantioseparation of derivatized α-amino acids [106]. Recently, only the cyclofructan-based chiral stationary phases was used for the separation of amino acid enantiomers.

3.1.3.1. Cyclofructan-Based Chiral Stationary Phases

Xie et al. have summarised the progress on cyclofructan derivatives-based chiral stationary phases for enantioseparation by gas chromatography. Five chiral stationary phases were used for the enantioseparation of DL-amino acids. Two of them—4,6-Di-O-pentyl-3-O-trifluoroacetyl cycloinulohexaose-based CSP (CF-CSP4) and 4,6-di-O-pentyl-3-O-propionyl cycloinulohexaose-based CSP (CF-CSP5)—were very promising and allowed resolutions ranging from 1.5 to 2.6 for four commons and three uncommon DL-amino acids (Ala, Ile, Leu, Val, allo-Ile, Nle and Nva) to be obtained [107].

3.1.4. Capillary Electrophoresis

During 2017 and 2018, Yu et al. summarized several types of chiral selector used for the enantioseparation by capillary electrophoresis. In CE, the chiral selector is added to the buffer solution as a pseudophase, in the same way as the mobile phase in LC. Chiral separation is made possible through non-covalent interactions between the chiral selector and the enantiomers, such as dipole–dipole, hydrogen bond, electrostatic and steric interactions. These chiral selectors can be classed into two categories: low and high molecular mass molecules. Low molecular mass molecules contain ionic liquids and complexes with a central ion. High molecular mass molecules contain the macromolecules such as oligosaccharides (cyclodextrins, polysaccharides), amino acids and nucleic acid-based polymers, and supramolecules such as bile salt micelles. In capillary electrophoresis, chiral selectors can also use as a dual ligand. Dual ligands were created in a combination of immobilized ligands on the surface of the capillary, as a coating, and the addition of free ligands in the buffer solution. In this part, we will further discuss as a chiral selector, two high molecular mass molecules (cyclodextrins and crown ethers) and one low molecular mass molecule (ligand exchange) used recently [108].

3.1.4.1. Cyclodextrin as a Chiral Selector

Native α, β, γ-cyclodextrins are made up of six, seven and eight glucose units, respectively, and their derivatives can be used as a chiral selector. The hydrophobic cavity dimensions depend on the nature of the cyclodextrin and can form non-covalent interactions with the chiral analytes [109]. These three standard cyclodextrins and their neutrally (mono-, di- and trimethylated and hydroxypropylated) and negatively (sulphated, sulfobutyletherated, succinylated, carboxymethylated, heptakis(6-O-sulfo), hexakis-, heptakis-, octakis(2,3-di-O-methyl-6-O-sulfo) and hep-takis(2,3-di-O-acetyl-6-O-sulfo) modifications) charged cyclodextrin derivatives were used for the chiral separation of six LL- and DD-dipeptide enantiomers. The type of cyclodextrin concerning cavity size and degree of substitution influenced the enantiomer migration order (EMO). In addition, the composition of the background electrolyte (BGE) as nature and pH influenced the analysis time. In acidic separation conditions the analysis time was shorter, and the resolution was worsened. In these two studies, neutrally and negatively β-CDs were shown to be more effective chiral selectors than the α- and γ-CD derivatives for the enantioseparation of Ala-Phe and the peptide analogs [110,111]. The γ-CD was also used in combination with an ionic liquid (L-Carnitine methyl ester bis(trifluoromethane)sulfonimide, L-

CarC1NTf2) as a chiral selector for the enantioseparation of Fmoc-DL-Hcy and Fmoc-DL-Cys. Enantiomeric resolutions obtained for the separation of these pure enantiomers were 6.1 and 6.4, respectively. For the simultaneous separation of these four amino acids derivatized, this method allowed a resolution as high as 4.1 to be obtained for Rs (Hcy): 5.9 and Rs (Cys). This study has not yet been applied to mixtures of biological or more complex compounds [112].

Microchip electrophoresis (MCE) mode can also be used for the enantioseparation of DL-amino acids. Recently, Zhang et al. have used a chiral nematic mesoporous silica (CNMS) as the chiral stationary phase in combination with hydroxypropyl-β-cyclodextrin (HP-β-CD) as the chiral selector for the separation of 10 DL-amino acid pairs. This combination allowed the enantioresolution to be improved to a resolution average of 1.13, while with the use of only HP-β-CD, the resolution average for these 10 DL-amino acid enantiomers was 0.68 [113].

3.1.4.2. Crown Ether as a Chiral Selector

The most commonly used crown ethers are [18] crown-6 and its derivatives. Their cavity structure limits chiral separation to primary amines, including amino acids. However, the separation must be performed under pure acidic conditions so that these amines are fully protonated and other cationic species present in the BGE do not interact with the crown ether [109].

Lee et al. have developed a method for the enantioseparation of underivatized free amino acids using (+)- and (−)-18C6H4 as a chiral selector [114]. This method allows separating all DL-amino acids, except DL-Pro and DL-Asn, with a resolution ranking from Rs(Cys): 0.5 to Rs(Ser): 21.0. Concerning the migration order for amino acids, the L-amino acids migrated faster than their analog D-amino acid counterparts except for Ser, Thr and Met. This method shows promise for the separation of amino acid enantiomers. However, some improvements can be made for the resolution of separation of DL-Cys (Rs: 0.5), DL-Ala (Rs: 0.7) and DL-Glu (Rs: 0.8).

3.1.4.3. Ligand Exchange as a Chiral Selector

Chiral ligand exchange capillary electrochromatography (CLE-CEC) using a central ion has been recently used for enantioseparation of derivatized or underivatized DL-amino acids. First, Zn (II) was used as a central ion for the enantioseparation of Dns-DL-amino acid enantiomers. For this amino acid separation, the authors have used four chiral DL-oligopeptides based on lysine residues as ligands to form [((Gly)-Lysx) Zn (II)(AA)] complexes. The migration order of Dns-DL-amino acid analytes has been influenced by the nature of the chiral ligand employed. Indeed, when the Gly-D-Lys lig-and was used, the Dns-L-amino acid had a faster retention time than the Dns-D-amino acid analyte. Chiral dipeptide ligands, Gly-L-Lys and Gly-D-Lys, allow for a better separation of Dns-DL-amino acid enantiomers with a resolution greater than 2, than other ligands ((L-Lys)2-OH) and ((L-Lys)4-OH) with a resolution below 1. This new method shows the great potential of dipeptides as ligands in the CLE-CEC separation of DL-amino acid enantiomers and can be developed for the separation of complex biological samples [115]. In another study, Feng et al. used different synthetic copolymer dual ligands as chiral ligands to form [(P(MAn-St-MAX) Zn (II)(AA)] complexes. The use of the free chiral ligands and immobilized ligands alone do not allow for separation of the DL-amino acids enantiomers, while the dual ligands have offered a good separation (Figure 7). With the dual ligand system, resolutions ranged from 1.67 to 3.45 for Ala, Gln, Ile, Tyr, Asn, Met and Ser. For other amino acids, including Asp, Trp, Leu, Phe and Thr, their resolutions were lower and are ranked from 0.55 to 1.19. DL-Pro and DL-Val were not separated. Among the four synthesized copolymer dual chiral ligands, the best results were obtained with the poly maleic anhydride-co-styrene-co-N-methacryloyl-L-histidine methyl ester [P(MAn-St-MAH)] [116].

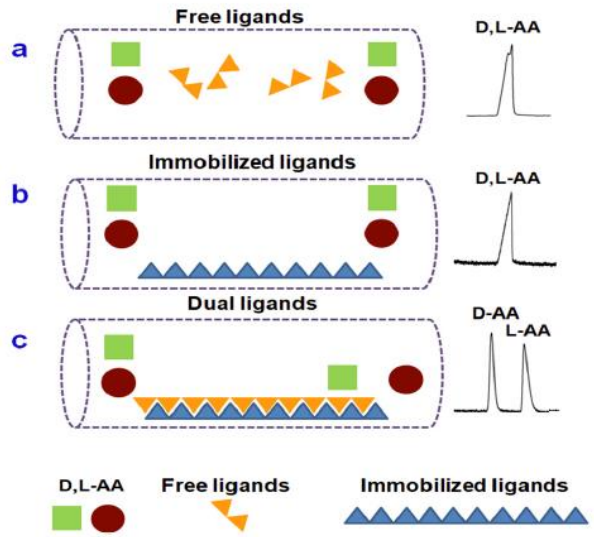


Figure 7. Illustration of enantioseparation mechanism of CLE-CEC by using (a) free ligand alone, (b) immobilized ligand alone and (c) dual ligands. Reproduced with permission from the authors of [116].

Cu (II) was also used as a central ion in complexation with L-His to create the [Cu (II)-L-His] complex with β-cyclodextrin (β-CD) as a dual chiral selector for the enantioseparation of DL-Trp, DL-Tyr and DL-Phe by capillary electrophoresis. For this study, several concentrations of Cu (II), L-His and β-CD were tested at different pH and voltages. Each pair of enantiomers was separated in 20 minutes with a resolution ranging from 3.6 to 6.1. For DL-Phe and DL-Trp, the EMO was similar to the L-enantiomer and eluted faster than its counterpart. Both peaks from the separation of D- and L-Tyr were not identified [117].

3.1.4.4. Micellar Electrokinetic Chromatography (MEKC)

Micellar electrokinetic chromatography is another capillary electrophoresis mode using a high concentration of micelles. Recently, Evans et al. have analyzed 36 primary amines including D- and L-amino acids derivatized with 2,3-naphthalenedicarboxaldehyde and separated by MEKC using sodium dodecyl sulfate and sodium deoxycholate as surfactants. Then, the authors applied this method for the enantioseparation of DL-Ser at low nanomolar concentrations containing in the endocrine portion of the pancreas, the islets of Langerhans. In the study, the successful separation of D-Ser to its counterpart has an important interest due to its agonist activity for the ionotropic N-methyl-D-aspartate receptors [118].

3.1.5. Comparison of Different Techniques

Among chromatographic and electrophoresis techniques previously described, the most efficient of them, in terms of chiral separation performances (resolution and separation factor), were summarized in Table 7. The crown ether used as chiral stationary phase or added to BGE was preferably used for the enantioseparation of complex mixtures of amino acids. Furthermore, in liquid chromatography, zwitterionic-exchange- and macrocycle antibiotic-based chiral stationary phase also offered an excellent separation for a similar complex mixture of DL-amino acids. Although gas chromatography is rarely used for the chiral separation of amino acids, the cyclofructan-based chiral stationary phase (CF-CSP5) provided sufficient separation performance.

Regarding the enantioseparation of small peptides, the crown-ether (Table 4) ion-exchange and zwitterionic (Table 5) chiral stationary phases were used in liquid chromatography. The results proved that the Crownpak CR-I (+) column offered a better average resolution (Rs: 5.12).

Table 7. Summary of the most efficient techniques for enantioseparation of underivatized DL-amino acids.

Techniques	CSP	Columns	Analytes	Rs and/or α average	Ref.
Liquid chromatography	Crown-ether	Cronwpak CR-I(+)	21 DL-amino acids	Rs: 5.12 (min. 1.49, max. 8.90)	[80]
	Zwitterrionic	Chiralpak ZWIX(+)	21 DL-amino acids	α: 2.33 (min. 1.26, max. 5.31)	[93]
	Macrocycle antibiotic	UHPC-SPP-Halo-Tzwitt	28 DL-amino acids	Rs: 5.01 (min. 1.15, max. 10.90) α: 1.52 (min. 1.02, max. 2.44)	[100]
Supercritical fluid chromatography	Crown-ether	Crownpak CR-I(+)	18 DL-amino acids	Rs: 5.27 (min. 1.99, max. 9.26)	[105]
Gas chromatography	Cyclofructan	CF-CSP5	7 DL-amino acids	Rs: 1.81 (min. 1.50, max. 2.60) α: 1.04 (min. 1.03, max. 1.06)	[107]
Capillary electrophoresis	Crown-ether	BGE containing 18C6H4	18 DL-amino acids	Rs: 2.97 (min. 0.70, max. 21.00) α: 1.03 (min. 1.01, max. 1.22)	[114]

3.2. Indirect Method

The indirect method consists of using the derivatization reaction on racemic compounds with a pure chiral reagent, thus resulting in the formation of a pair of diastereomers which can be separated by a chiral or achiral column. The derivatization reaction must be at room temperature to avoid racemization. This method is usually used to facilitate the isolation of the analyte from the biological matrix. In this review, we will summarize the derivatization reagents used recently for the enantioseparation of amino acids and newly synthesized derivatization reagents. Among derivatization reagents that have been used recently, most of them react with the N-terminus of amino acids, and they were commercial ((+)- or (−)-1-(9-fluorenyl)ethyl chloroformate) ((+)- or (−)-FLEC) and N-(4-Nitrophenoxycarbonyl)-L-phenylalanine 2-methoxyethyl ester (S-NIFE)) or synthetic ((R)- or (S)-4-nitrophenyl-N-[2-(diethylamino)-6,6-dimethyl-[1,1-biphenyl]-2-yl] carbamate hydrochloride ((R)- or (S)-BiAC) and N-[1-oxo-5-(triphenylphosphonium) pentyl]-(R)-1,3-thiazolidinyl-4-N-hydroxysuccinimide ester bromide salt (OTPTHE)).

The more popular commercial derivatization reagents were the (+)- and (−)-FLEC. Indeed, Moldovan et al. used these reagents in their both undifferentiated forms as chiral derivatization reagents for the enantioseparation of 19 amino acid pair of diastereomers by RP-HPLC coupled to mass spectrometry with diphenyl and biphenyl stationary phases [119]. The average resolution for each amino acid pairs was 2.9 and 2.3 for the diphenyl and biphenyl stationary phases respectively, under optimal separation conditions. Indeed, with diphenyl stationary phase, 14 of 19 amino acid pair of enantiomers (except Ala, Ser, Pro, Thr and Asp) obtained a good resolution (Rs > 1.5) at pH 4.9, and 13 of 19 amino acid pair of diatereoisomers (except Ala, Ser, Pro, Thr, Asp and Glu) at pH 4.2. Both columns allow DL-amino acids to be separated with the same elution order with D-derivative first respectively at pH 4.9 and 4.2. However, the authors observed a reversal in the elution order with L-form eluted first at pH 2.5 and 7. These two methods allow separating pure FLEC-DL-amino acid enantiomers in less than 30 minutes. This short time does not allow separating a complex mixture of amino acids, because the retention times are very near. Furthermore, (+)-FLEC was used by Pérez-Míguez et al. where a TIMS-TOFMS method was developed for the analysis of 21 common and uncommon (+)-FLEC-DL-amino acid enantiomers [120]. With this new method, 17 were separated separately at different voltage ramps. Then, the enantioseparation of several (+)-FLEC-DL-amino acids in one run was tested. For this work, three groups of four (+)-FLEC-DL-amino acids were used with one voltage ramp applied per group. The first group consisted of Orn, Lys, His and Tyr with voltage ramps ranging from 135 to 170 V; Asn, Ser, Pipe and Gln at 100–130V for the second group; and Met, Phe, SeMet and Arg at 109-135 V for the last group. For each group, some amino acids have co-migrated. Their distinction was possible thanks to their respective m/z (Figure 8). This method allowed Ser and Asp separation with a better resolution than the previous study.

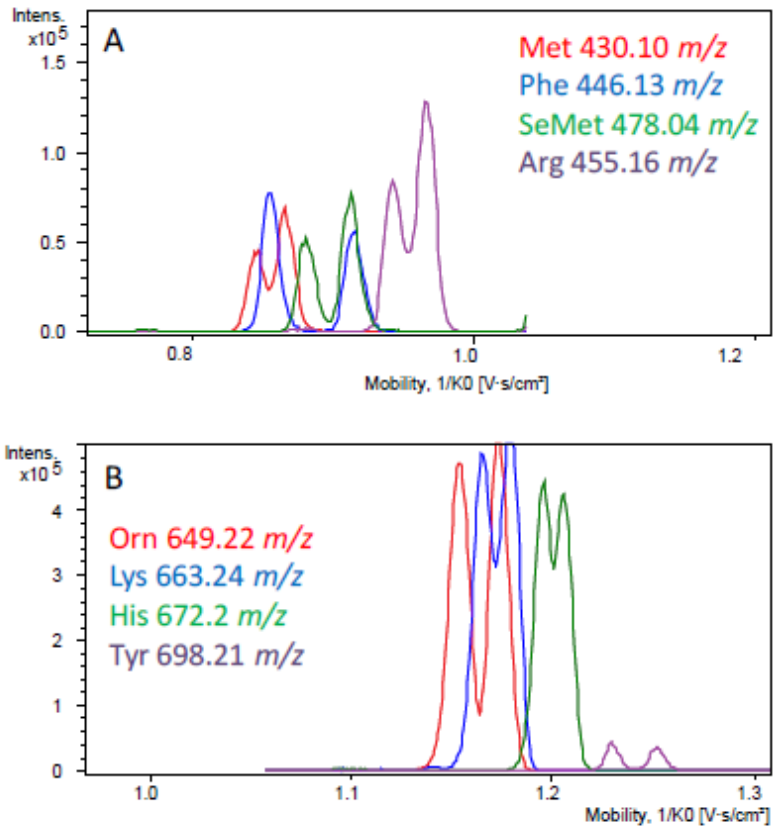


Figure 8. TIMS-TOFMS of FLEC-AA mixtures using a voltage ramp of (A) 109–135V and (B) 135–170V in 510 ms. Reproduced with permission from the authors of [120].

The commercial S-NIFE derivatized reagent, as far as it is concerned, was used by Danielsen et al. which have developed a method for the determination of L- and D-amino acid ratios contained in complex protein matrices using LC-MS/MS analysis. Protein and standard amino acids were hydrolyzed by DCl following two methods: in glass capillaries and vacuum hydrolysis tubes. In this case, deuterated hydrolysis has been used to eliminate the L- and D-amino acid racemization ratio due to the natural conversion. These two hydrolysis methods allow determining the rate of 0.30% and 0.25% D-amino acids, respectively, contained in α-lactalbumin protein. Their exact positions on the protein sequence were not determined [66].

Other derivatization reagents can be synthesized to react with the N-terminus of amino acids. Historically, the axially chiral hydrophobic derivatizing reagent was developed for the analysis of amine and alcohol diastereoisomers by normal-phase liquid chromatography [121,122]. However, for the axial derivatization for hydrophilic compounds, such as amino acids, a new derivatization reagent was necessary to be developed. Recently, Harada et al. have developed biaryl axially derivatives, (R)-4-nitrophenyl N-[2'-(diethylamino)-6,6'-dimethyl-[1,1'-biphenyl]-2-yl] carbamate hydrochloride ((R)-BiAC), as a new axially chiral derivatizing reagent for the enantioseparation of proteinogenic amino acids by RP-HPLC-MS/MS. This new chiral derivatizing reagent allowed all DL-amino acids to be separated within 11.5 minutes with a resolution greater than 1.9, except for complex mixtures, such as threonine mixture (DL-Thr and DL-allo-Thr), and leucine-isoleucine mixture (DL-Leu, DL-Ile and DL-allo-Ile) [123]. The same authors applied this new derivatizing reagent ((R)-BiAC) on DL-amino acids in human urine for their simultaneous determination by liquid chromatography electrospray ionization

tandem mass spectrometry. After optimization of separation conditions, 36 DL-amino acids and glycine were separated in 20 minutes with good resolution. These results were verified using the opposite chiral derivatizing reagent (S)-BiAC. Using (S)-BiAC causes the inversion of the migration order of D- and L-amino acids. The concentration of each amino acid was determined by LC-MS. The (S)-BiAC/(R)-BiAC ratio of amino acid concentration was higher (>84%), except for D-allo-Ile (67%) and for D-Asp, D-Pro, D-Trp and L-Asp whose peaks were below the limit of quantification (LOQ) with both chiral derivatizing reagents. This method is promising for simultaneous analysis of DL-amino acids in biological samples (Figure 9). In this case, the quantification of D-amino acids in human urine can be used as a biomarker to screen for chronic kidney disease [124].

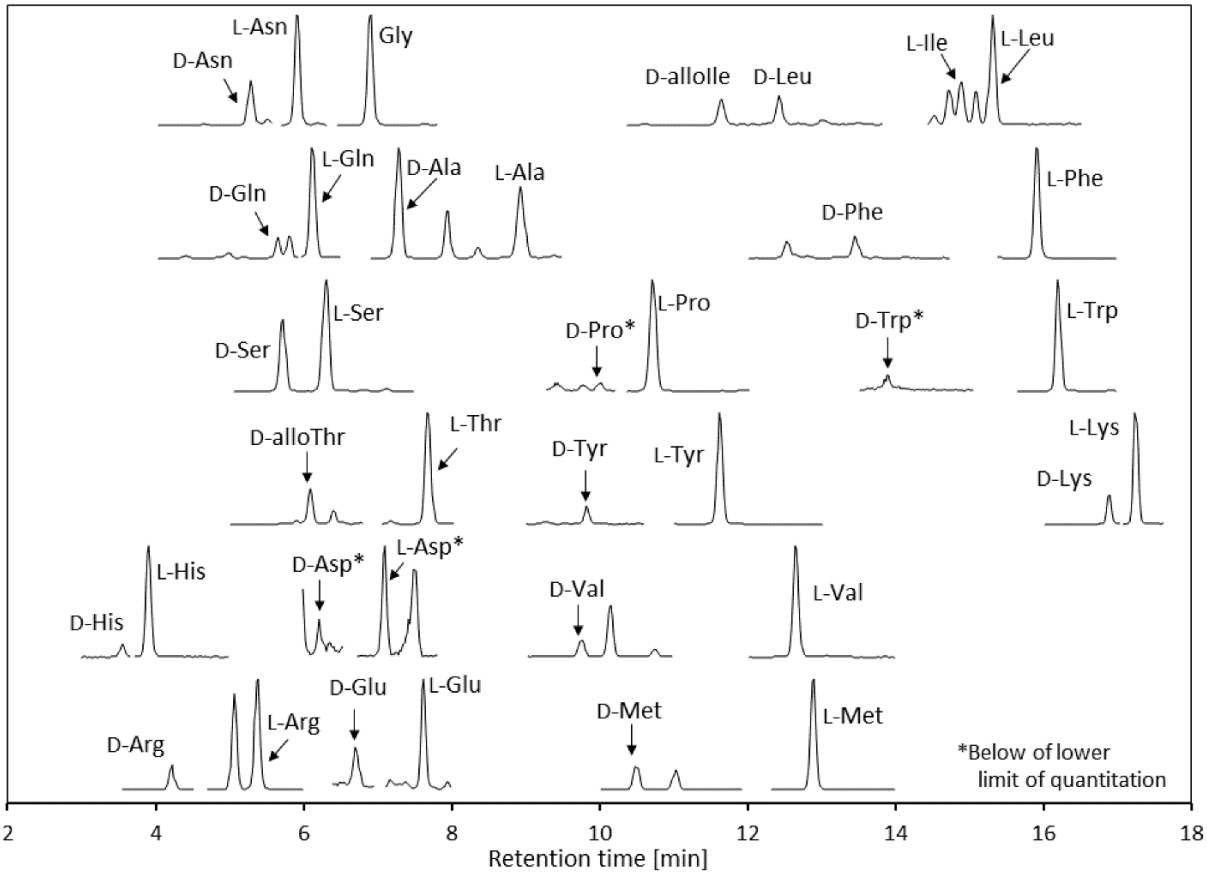


Figure 9. Chromatograms of D- and L-amino acids and glycine in human urine sample. Reproduced with permission from the authors of [124].

Another group has synthesized a new derivatization reagent N-[1-oxo-5-(triphenylphosphonium) pentyl] -(R)-1,3-thiazolidinyl-4-N-hydroxysuccinimide ester bromide salt (OTPTHE) for the enantioseparation of DL-Ser from human plasma by liquid chromatography coupled to mass spectrometry. This reagent first was tested for the separation of 13 DL-amino acids in 22 min, with the resolution from Rs (Thr): 1.62 to Rs (Pro): 2.51. This derivatization method was applied to analyzed DL-Ser from the plasma of 10 healthy volunteers with a similar resolution. The authors do not show the results for the other amino acids contained in the human plasma [125].

Derivatization mainly takes place on the terminal amine of the amino acid, which makes it possible to study all amino acids simultaneously. However, it is possible to derivatize the amino acid with another functional group to allow for more selective analysis. Thiols naturally occur in their reduced or oxidized forms. The tripeptide, glutathione, is found in both of these forms and is primarily studied as a biomarker for the assessment of oxidative stress. Russo et al. compared two cell-permeable

commercial derivatizing agents—N-ethyl maleimide (NEM) and (R)-(+)-N-(1-phenylethyl) maleimide (NPEM)—for derivatization of biological thiols using a biphenyl reversed-phase liquid chromatography–high-resolution mass spectrometry (LCHRMS). Four biological thiols were used for the evaluation: cysteine (Cys), homocysteine (Hcy), N-acetylcysteine (NAC) and glutathione (GSH). The derivatization efficiency of Cys and their derivatives by NEM and NPEM was 97% and 100%, respectively. NPEM also allowed for the enhancement of ESI ionization. However, it was more prone to experience side reactions, such as derivatization of amines and opening of the cycle, and it is more sensible to pH variation, compared to NEM. For cell-permeable thiol-protecting reagent used on mild derivatizing conditions to minimize the impact on the cellular structure and avoid artificial changes in metabolite concentrations, NEM seems to be better than NPEM [126].

4. Concluding Remarks and Prospects

Homochirality is omnipresent in proteins as much by the D- and L-amino acid enantiomers as by the D- and L-oligosaccharide enantiomers as post-translational modifications. The natural conversion of L/D-amino acids in proteins takes place in a nonuniform manner. Thus, it is necessary to examine the chirality of each amino acid in the proteinogenic sequence. For this determination, sequence-dependent and sequence-independent strategies are possible. The sequence-dependent strategy can be applied with two different methods. First, the protein is digested by an enzyme and each peptide obtained was identified on the protein sequence by mass spectrometry analysis. Then, these peptides were hydrolyzed on amino acid and then react with a standard protected amino acid, such as Boc-L-Cys or Boc-L-Leu as an example. This method allows the LL- and DL-dipeptides formed, which have different retention times, to be distinguished. The second consists of digesting the protein by a specific enzyme. Various enzymes are enantioselective such as: endoproteinase Asn-N, L-isoaspartyl methyltransferase, D-aspartyl endoproteinase and glutamyl endoproteinase Glu-C, which recognise only L-α-Asp, L-β-Asp, D-α-Asp and L-Glu, respectively. The sequence-independent strategy, as far as it is concerned, consists of hydrolyzing the protein under acidic conditions. The use of deuterated or tritiated hydrochloric acid in heavy water exchanges the hydrogen on the alpha carbon on free amino acid with a deuterium or tritium atom and limits the natural racemization. Although liquid chromatography is widely used for the determination of D-amino acids, other separation techniques are also used. Some chromatographic and electrophoretic techniques using commercial or synthetic chiral stationary phases can be used for the enantioseparation of free amino acids and peptides. Among these chiral stationary phases, the most recently used was based on polysaccharides, cyclodextrins, crown-ethers, Brush-types, ion- and zwitterion-exchange, macrocycle antibiotics, cyclofructans and ligand exchange. The comparative study of these different chiral stationary phases showed, in terms of separation performance, that the Crownpak CR-I (+) column offered the best separation of DL-amino acid and small peptide enantiomers. Another enantioseparation technique is based on a derivatization reaction with a pure chiral compound on the free amino acids. Although the majority of commercial ((+)- or (−)-FLEC and S-NIFE) or synthetic ((R)- or (S)-BiAC and OTPTHE) derivatization reagents reaction on the N-terminus of amino acids, some chiral reagents were developed to react with another amino acid functional group for more selective analysis, like NEM and NPEM reagents on cysteine thiol functions. For future work, given the many possibilities, new chiral stationary phases and derivatization reagents can be developed and applied to different proteins with important biological functions. This will make it possible to discover other D-amino acids and their position on the proteinogenic sequence and their occurrence on human health and disease.

Funding: This research was funded by Czech Science Foundation, grant number 20-03899S.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Qi, X.; Tester, R.F. Fructose, Galactose and Glucose—In Health and Disease. Nutr. ESPEN 2019, 33, 18–28, doi:10.1016/j.clnesp.2019.07.004.

2. Qi, X.; Tester, R.F. Lactose, Maltose, and Sucrose in Health and Disease. Nutr. Food Res. 2020, 64, 1901082–1901091, doi:10.1002/mnfr.201901082.

3. Sajadimajd, S.; Bahrami, G.; Daglia, M.; Nabavi, S.M.; Naseri, R.; Farzaei, M.H. Plant-Derived Supplementary Carbohydrates, Polysaccharides and Oligosaccharides in Management of Diabetes Mellitus: A Comprehensive Review. Food Rev. Int. 2019, 35, 563–586, doi:10.1080/87559129.2019.1584818.

4. Reily, C.; Stewart, T.J.; Renfrow, M.B.; Novak, J. Glycosylation in Health and Disease. Nat. Rev. Nephrol. 2019, 15, 346–366, doi:10.1038/s41581-019-0129-4.

5. He, B.L.; Lloyd, D.K. Chiral Methods. In Specification of Drug Substances and Products, 2nd ed.; Riley, C.M., Rosanske, T.W., Reid, G., Eds.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 425–458. ISBN 9780081028247.

6. Annavarapu, S.; Nanda, V. Mirrors in the PDB: Left-Handed α-Turns Guide Design with D-Amino Acids. BMC Struct. Biol. 2009, 9, 61–74, doi:10.1186/1472-6807-9-61.

7. Grieco, P.; Carotenuto, A.; Auriemma, L.; Saviello, M.R.; Campiglia, P.; Gomez-Monterrey, I.M.; Marcellini, L.; Luca, V.; Barra, D.; Novellino, E.; et al. The effect of d-amino acid substitution on the selectivity of temporin L towards target cells: Identification of a potent anti-Candida peptide. Biochim. Biophys. Acta Biomembr. 2013, 1828, 652–660, doi:10.1016/j.bbamem.2012.08.027.

8. Gößler-Schöfberger, R.; Hesser, G.; Reif, M.M.; Friedmann, J.; Duscher, B.; Toca-Herrera, J.L.; Oostenbrink, C.; Jilek, A. A stereochemical switch in the aDrs model system, a candidate for a functional amyloid. Arch. Biochem. Biophys. 2012, 522, 100–106, doi: 10.1016/j.abb.2012.04.006.

9. Fujii, N.; Fujii, N.; Kida, M.; Kinouchi, T. Influence of Lβ-, Dα- and Dβ-Asp Isomers of the Asp-76 Residue on the Properties of AA-Crystallin 70–88 Peptide. Amino Acids 2010, 39, 1393–1399, doi:10.1007/s00726-010-0597-0.

10. Fujii, N.; Saito, T. Homochirality and Life. Chem. Rec. 2004, 4, 267–278, doi:10.1002/tcr.20020.

11. Fujii, N. D-Amino Acid in Elderly Tissues. Biol. Pharm. Bull. 2005, 28, 1585–1589, doi:10.1248/bpb.28.1585.

12. Fujii, N.; Kawaguchi, T.; Sasaki, H.; Fujii, N. Simultaneous Stereoinversion and Isomerization at the Asp-4 Residue in βB2-Crystallin from the Aged Human Eye Lenses. Biochemistry 2011, 50, 8628–8635, doi:10.1021/bi200983g.

13. Fujji, N.; Takata, T.; Fujii, N.; Aki, K.; Sakaue, H. D-Amino Acids in Protein: The Mirror of Life as a Molecular Index of Aging. Biochim. Biophys. Acta Proteins Proteom. 2018, 1866, 840–847, doi:10.1016/j.bbapap.2018.03.001.

14. Miyamoto, T.; Homma, H. Detection and Quantification of d-Amino Acid Residues in Peptides and Proteins Using Acid Hydrolysis. Biochim. Biophys. Acta Proteins Proteom. 2018, 1866, 775–782, doi:10.1016/j.bbapap.2017.12.010.

15. Ha, S.; Kinouchi, T.; Fujii, N. Age-Related Isomerization of Asp in Human Immunoglobulin G Kappa Chain. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140410–140417, doi:10.1016/j.bbapap.2020.140410.

16. Bastings, J.J.A.J.; van Eijk, H.M.; Olde Damink, S.W.; Rensen, S.S. D-Amino Acids in Health and Disease: A Focus on Cancer. Nutients 2019, 11, 2205–2223, doi:10.3390/nu11092205.

17. Kuwada, M.; Teramoto, T.; Kumagaye, K.Y.; Nakajima, K.; Watanabe, T.; Kawai, T.; Kawahami, Y.; Niidome, T.; Sawada, K.; Nishizawa, Y.; et al. ω-Agatoxin-TK Containing D-Serine at Position 46, but Not Synthetic ω-[L-Ser]Agatoxin-TK, Exerts Blockade of P-Type Calcium Channels in Cerebellar Purkinje Neurons. Mol. Pharmacol. 1994, 46, 587–593.

18. Richter, G.; Egger, R.; Kreil, G. D-Alanine in the Frog Skin Peptide Dermorphin Is Derived from L-Alanine in the Precursor. Science 1987, 238, 200–202, doi:10.1126/science.3659910.

19. Pavithrra, G.; Rajasekaran, R. Gramicidin Peptide to Combat Antibiotic Resistance: A Review. Int. J. Pept. Res. Ther. 2020, 26, 191–199, doi:10.1007/s10989-019-09828-0.

20. Fujii, N.; Satoh, K.; Harada, K.; Ishibashi, Y. Simultaneous Stereoinversion and Isomerization at Specific Aspartic Acid Residues in OrA-Crystallin from Human Lens. J. Biochem. 1994, 116, 663–669, doi: 10.1093/oxfordjournals.jbchem.a124577.

21. Hooi, M.Y.S.; Truscott, R.J.W. Racemisation and Human Cataract. d-Ser, d-Asp/Asn and d-Thr Are Higher in the Lifelong Proteins of Cataract Lenses than in Age-Matched Normal Lenses. Age 2011, 33, 131–141, doi:10.1007/s11357-010-9171-7.

22. Hooi, M.Y.S.; Raftery, M.J.; Truscott, R.J.W. Accelerated Aging of Asp 58 in AA Crystallin and Human Cataract Formation. Exp. Eye Res. 2013, 106, 34–39, doi:10.1016/j.exer.2012.10.013.

23. Wu, H.-T.; Julian, R.R. Two-Dimensional Identification and Localization of Isomers in Crystallin Peptides Using TWIM-MS. Analyst 2020, 145, 5232–5341, doi:10.1016/j.trac.2021.116287.

24. Fujii, N.; Takata, T.; Kim, I.; Morishima, K.; Inoue, R.; Magami, K.; Matsubara, T.; Sugiyama, M.; Koide, T. Asp Isomerization Increases Aggregation of α-Crystallin and Decreases Its Chaperone Activity in Human Lens of Various Ages. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140446–140457, doi:10.1016/j.bbapap.2020.140446.

25. Masuda, W.; Nouso, C.; Kitamura, C.; Terashita, M.; Noguchi, T. D-Aspartic Acid in Bovine Dentine Non-Collagenous Pho-phoprotein. Arch. Oral Biol. 2002, 47, 757–762, doi:10.1016/S0003-9969(02)00064-X.

26. Powell, J.T.; Vine, N.; Crossman, M. On the Accumulation of D-Aspartate in Elastin and Other Proteins of the Ageing Aorta. Atherosclerosis 1992, 97, 201–208, doi:10.1016/0021-9150(92)90132-Z.

27. Ritz-Timme, S.; Laumeier, I.; Collins, M.J. Aspartic Acid Racemization: Evidence for Marked Longevity of Elastin in Human Skin. Br. J. Dermatol. 2003, 149, 951–959, doi:10.1111/j.1365-2133.2003.05618.x.

28. Ishigo, S.; Negishi, E.; Miyoshi, Y.; Onigahara, H.; Mita, M.; Miyamoto, T.; Masaki, H.; Homma, H.; Ueda, T.; Hamase, K. Establishment of a Two-Dimensional HPLC-MS/MS Method Combined with

DCL/D2O Hydrolysis for the Determination of Trace Amounts of D-Amino Acid Residues in Proteins. Chromatography 2015, 36, 45–50, doi:10.15583/jpchrom.2015.017.

29. Friedrich, M.G.; Hancock, S.E.; Raftery, M.J.; Truscott, R.J.W. Isoaspartic Acid Is Present at Specific Sites in Myelin Basic Protein from Multiple Sclerosis Patients: Could This Represent a Trigger for Disease Onset? Acta Neuropathol. Commun. 2016, 4, 83–95, doi:10.1186/s40478-016-0348-x.

30. Fujii, N.; Ishibashi, Y.; Satoh, K.; Fujino, M.; Harada, K. Simultaneous racemization and isomerization at specific aspartic acid residues in αB-crystallin from the aged human lens. Biochim. Biophys. Acta 1994, 1204, 157–163, doi:10.1016/0167-4838(94)90003-5.

31. Livnat, I.; Tai, H.-G.; Jansson, E.T.; Bai, L.; Romanova, E.V.; Chen, T.-T.; Yu, K.; Chen, S.-A.; Zhang, Y.; Wang, Z.-Y.; et al. A D-Amino Acid-Containing Neuropeptide Discovery Funnel. Anal. Chem. 2016, 88, 11868–11876, doi:10.1021/acs.analchem.6b03658.

32. Young, G.W.; Hoofring, S.A.; Mamula, M.J.; Doyle, H.A.; Bunick, G.J.; Hu, Y.; Aswad, D.W. Protein L-Isoaspartyl Methyl-transferase Catalyzes In Vivo Racemization of Aspartate-25 in Mammalian Histone H2B. J. Biol. Chem. 2005, 280, 26094–26098, doi:10.1074/jbc.M503624200.

33. Ritz, S.; Turzynski, A.; Schütz, H.W.; Hollmann, A.; Rochholz, G. Identification of Osteocalcin as a Permanent Aging Con-stituent of the Bone Matrix: Basis for an Accurate Age at Death Determination. Forensic Sci. Int. 1996, 77, 13–26, doi:10.1016/0379-0738(95)01834-4.

34. Kawamura, I.; Mijiddorj, B.; Kayano, Y.; Matsuo, Y.; Ozawa, Y.; Ueda, K.; Sato, H. Separation of D-Amino Acid-Containing Peptide Phenylseptin Using 3,3′-Phenyl-1,1′-Binaphthyl-18-Crown-6-Ether Columns. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140429–140436, doi:10.1016/j.bbapap.2020.140429.

35. Roher, A.E.; Lowenson, J.D.; Clarke, S.; Wolkow, C.; Wang, R.; Cotter, R.J.; Reardon, I.M.; Zürcher-Neely, H.A.; Heinrikson, R.L.; Ball, M.J.; et al. Structural Alterations in the Peptide Backbone of Beta-Amyloid Core Protein May Account for Its Depo-sition and Stability in Alzheimer's Disease. J. Biol. Chem. 1993, 268, 3072–3083, doi:10.1016/S0021-9258(18)53661-9.

36. Kaneko, I.; Yamada, N.; Sakuraba, Y.; Kamenosono, M.; Tutumi, S. Suppression of Mitochondrial Succinate Dehydrogenase, a Primary Target of β-Amyloid, and Its Derivative Racemized at Ser Residue. J. Neurochem. 1995, 65, 2585–2593, doi:10.1046/j.1471-4159.1995.65062585.x.

37. Zhang, J.; Yip, H.; Katta, V. Identification of Isomerization and Racemization of Aspartate in the Asp–Asp Motifs of a Ther-apeutic Protein. Anal. Biochem. 2011, 410, 234–243, doi:10.1016/j.ab.2010.11.040.

38. Mijiddorj, B.; Kaneda, S.; Sato, H.; Kitahashi, Y.; Javkhlantugs, N.; Naito, A.; Ueda, K.; Kawamura, I. The Role of d-Allo-Isoleucine in the Deposition of the Anti-Leishmania Peptide Bombinin H4 as Revealed by 31 P Solid-State NMR, VCD Spectroscopy, and MD Simulation. Biochim. Biophys. Acta Proteins Proteom. 2018, 1866, 789–798, doi:10.1016/j.bbapap.2018.01.005.

39. Gause, G.F. Gramicidin S review of recent work. Lancet 1946, 248, 46–47, doi:10.1016/S0140-6736(46)90004-9.

40. Yan, Y.; Wei, H.; Fu, Y.; Jusuf, S.; Zeng, M.; Ludwig, R.; Krystek, S.R., Jr.; Chen, G.; Tao, L.; Das, T.K. Isomerization and Oxidation in the Complementarity-Determining Regions of a Monoclonal Antibody: A Study of the Modification–Structure–Function Correlations by Hydrogen–Deuterium Exchange Mass Spectrometry. Anal. Chem. 2016, 88, 2041–2050, doi:10.1021/acs.analchem.5b02800.

41. Soyez, D.; Laverdure, A.-M.; Kallen, J.; van Herp, F. Demonstration of a Cell-Specific Isomerization of Invertebrate Neuro-peptides. Neuroscience 1997, 82, 935–942, doi:10.1016/S0306-4522(97)00254-6.

42. Jimenéz, E.C.; Olivera, B.M.; Gray, W.R.; Cruz, L.J. Contryphan Is a D-Tryptophan-Containing Conus Peptide. J. Biol. Chem. 1996, 271, 28002–28005, doi:10.1074/jbc.271.45.28002.

43. Erspamer, V.; Melchiorri, P.; Falconieri-Erspamer, G.; Negri, L.; Corsi, R.; Severini, C.; Barra, D.; Simmaco, M.; Kreil, G. Del-torphins: A Family of Naturally Occurring Peptides with High Affinity and Selectivity for 6 Opioid Binding Sites. Proc. Natl. Acad. Sci. USA 1989, 86, 5188–5192, doi:10.1073/pnas.86.13.5188.

44. Mor, A.; Delfour, A.; Sagan, S.; Amiche, M.; Pradelles, P.; Rossier, J.; Nicolas, P. Isolation of Dermenkephalin from Amphibian Skin, a High-Affinity (δ-Selective Opioid Heptapeptide Containing a D-Amino Acid Residue. FEBS Lett. 1989, 255, 269–274, doi:10.1016/0014-5793(89)81104-4.

45. Cloos, P.A.C.; Fledelius, C. Collagen Fragments in Urine Derived from Bone Resorption Are Highly Racemized and Isomerized: A Biological Clock of Protein Aging with Clinical Potential. Biochem. J. 2000, 345, 473–480, doi:10.1042/bj3450473.

46. Ohta, N.; Kubota, I.; Takao, T.; Shimonishi, Y.; Yasuda-Kamatani, Y.; Minakata, H.; Nomoto, K.; Muneoka, Y.; Kobayashi, M. Fulicin, a Novel Neuropeptide Containing a D-Amino Acid Residue Isolated from the Ganglia of Achatina Fulica. Biochem. Biophys. Res. Commun. 1991, 178, 486–493, doi:10.1016/0006-291X(91)90133-R.

47. Kamatani, Y.; Minakata, H.; Kenny, P.T.M.; Iwashita, T.; Watanabe, K.; Funase, K.; Sun, X.P.; Yongsiri, A.; Kim, K.H.; No-vales-Li, P.; et al. Achatin-I, an endogenous neuroexcitatoty tetrapeptide from Achatina fulica ferussac containing a D-amino acid residue. Biochem. Biophys. Res. Commun. 1989, 160, 1015–1020, doi:10.1016/S0006-291X(89)80103-2.

48. Lee, C.J.; Qiu, T.A.; Sweedler, J.V. D-Alanine: Distribution, Origin, Physiological Relevance, and Implications in Disease. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140482–140501, doi:10.1016/j.bbapap.2020.140482.

49. Ayon, N.J. Features, roles and chiral analyses of proteinogenic amino acids. AIMS Mol. Sci. 2020, 7, 229–268, doi:10.3934/molsci.2020011.

50. Fisher, G.H.; Garcia, N.M.; Payan, I.L.; Cadilla-Perezrios, R.; Sheremata, W.A.; Man, E.H. D-aspartic acid in purified myelin and myelin basic protein. Biochem. Biophys. Res. Commun. 1986, 135, 683–687, doi:10.1016/0006-291X(86)90047-1.

51. Lee, J.M.; Petrucelli, L.; Fisher, G.; Ramdath, S.; Castillo, J.; Di Fiore, M.; D'Aniello, A. Evidence for D-Aspartyl-β-Amyloid Secretase Activity in Human Brain. J. Neuropathol. Exp. Neurol. 2002, 61, 125–131, doi:10.1093/jnen/61.2.125.

52. Hashimoto, A.; Oka, T. Free D-Aspartate and d-Serine in the Mammalian Brain and Periphery. Prog. Neurobiol. 1997, 52, 325–353, doi:10.1016/S0301-0082(97)00019-1.

53. Homma, H. Biochemistry of D-Aspartate in Mammalian Cells. Amino Acids 2007, 32, 3–11, doi:10.1007/s00726-006-0354-6.

54. Topo, E.; Soricelli, A.; D'Aniello, A.; Ronsini, S.; D'Aniello, G. The Role and Molecular Mechanism of D-Aspartic Acid in the Release and Synthesis of LH and Testosterone in Humans and Rats. Reprod. Biol. Endocrinol. 2009, 7, 120–131, doi:10.1186/1477-7827-7-120.

55. Chung, J.S.; Zmora, N.; Katayama, H.; Tsutsui, N. Crustacean Hyperglycemic Hormone (CHH) Neuropeptidesfamily: Func-tions, Titer, and Binding to Target Tissues. Gen. Comp. Endocrinol. 2010, 166, 447–454, doi: 10.1016/j.ygcen.2009.12.011.

56. Strauch, R.C.; Svedin, E.; Dilkes, B.; Chapple, C.; Li, X. Discovery of a Novel Amino Acid Racemase through Exploration of Natural Variation in Arabidopsis Thaliana. Proc. Natl. Acad. Sci. USA 2015, 112, 11726–11731, doi:10.1073/pnas.1503272112.

57. Miyamoto, T.; Moriya, T.; Homma, H.; Oshim, T. Enzymatic Properties and Physiological Function of Glutamate Racemase from Thermus thermophilus. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140461–140467, doi:10.1016/j.bbapap.2020.140461.

58. Ollivaux, C.; Soyez, D.; Toullec, J.-Y. Biogenesis of D -Amino Acid Containing Peptides/Proteins: Where, When and How? J. Pept. Sci. 2014, 20, 595–612, doi:10.1002/psc.2637.

59. Fujii, N.; Momose, Y.; Ishii, N.; Takita, M.; Akaboshi, M.; Kodama, M. The Mechanisms of Simultaneous Stereoinversion, Racemization, and Isomerization at Specific Aspartyl Residues of Aged Lens Proteins. Mech. Ageing. Dev. 1999, 107, 347–358, doi:10.1016/S0047-6374(98)00129-8.

60. Fujii, N.; Sakaue, H.; Sasaki, H.; Fujii, N. A Rapid, Comprehensive Liquid Chromatography-Mass Spectrometry (LC-MS)-based Survey of the Asp Isomers in Crystallins from Human Cataract Lenses. J. Biol. Chem. 2012, 287, 39992–40002, doi:10.1074/jbc.M112.399972.

61. Mitchell, A.R.; Kent, S.B.H.; Chu, I.C.; Merrifield, R.B. Quantitative Determination of D- and L-Amino Acids by Reaction with Tert-Butyloxycarbonyl-L-Leucine N-Hydroxysuccinimide Ester and Chromatographic Separation as L,D and L,L Dipeptides. Anal. Chem. 1978, 50, 637–640, doi:10.1021/ac50026a025.

62. Yan, L.; Ke, Y.; Kan, Y.; Lin, D.; Yang, J.; He, Y.; Wu, L. New Insight into Enzymatic Hydrolysis of Peptides with Site-Specific Amino Acid d-Isomerization. Bioorg. Chem. 2020, 10, 104389–104399, doi:10.1016/j.bioorg.2020.104389.

63. Du, S.; Readel, E.R.; Wey, M.; Armstrong, D.W. Complete Identification of All 20 Relevant Epimeric Peptides in β-Amyloid: A New HPLC-MS Based Analytical Strategy for Alzheimer's Research. Chem. Commun. 2020, 56, 1537–1540, doi:10.1039/c9cc09080k.

64. Takata, T.; Ha, S.; Koide, T.; Fujii, N. Site-Specific Rapid Deamidation and Isomerization in Human Lens AA-crystallin In Vitro. Protein Sci. 2020, 29, 941–951, doi:10.1002/pro.3821.

65. Kaiser, K.; Benner, T. Hydrolysis-Induced Racemization of Amino Acids: Hydrolysis-Induced Amino Acid Racemization. Limnol. Oceanogr. Methods 2005, 3, 318–325, doi:10.4319/lom.2005.3.318.

66. Danielsen, M.; Nebel, C.; Dalsgaard, T.K. Simultaneous Determination of L- and D-Amino Acids in Proteins: A Sensitive Method Using Hydrolysis in Deuterated Acid and Liquid Chromatography–Tandem Mass Spectrometry Analysis. Foods 2020, 9, 309–323, doi:10.3390/foods9030309.

67. Manning, J.M. Determination of D- and L-Amino Acid Residues in Peptides. Use of Tritiated Hydrochloric Acid to Correct for Racemization during Acid Hydrolysis. J. Am. Chem. Soc. 1970, 92, 7449–7454, doi:10.1021/ja00728a033.

68. Davankov, V. The Nature of Chiral Recognition: Is It a Three-Point Interaction? Chirality 1997, 9, 99–102, doi:10.1002/(SICI)1520-636X(1997)9:2<99::AID-CHIR3>3.0.CO;2-B.

69. Yu, L.; Wang, S.; Zeng, S. Chiral Mobile-Phase Additives in HPLC Enantioseparations. In Chiral Separations. Methods and Protocols, 3rd ed.; Scriba, G.K.E., Ed.; Methods in Molecular Biology; Humana: Totowa, NJ, USA; Springer: New York, NY, USA, 2019; pp. 81–91. ISBN 9781493994380.

70. Chankvetadze, B. Recent Trends in Preparation, Investigation and Application of Polysaccharide-Based Chiral Stationary Phases for Separation of Enantiomers in High-Performance Liquid Chromatography. Trends Anal. Chem. 2020, 122, 115729–115724, doi:10.1016/j.trac.2019.115709.

71. Mejía-Carmona, K.; da Silva Burato, J.S.; Borsatto, J.V.B.; de Toffoli, A.L.; Lanças, F.M. Miniaturization of liquid chromatog-raphy coupled to mass spectrometry: 1. Current trends on miniaturized LC columns. Trends Anal. Chem. 2020, 122, 115735–115750, doi:10.1016/j.trac.2019.115735.

72. Zhao, Y.; Zhu, X.; Jiang, W.; Liu, H.; Sun, B. Chiral Recognition for Chromatography and Membrane-Based Separations: Recent Developments and Future Prospects. Molecules 2021, 26, 1145–1175, doi:10.3390/molecules26041145.

73. Okamoto, Y.; Kawashima, M.; Hatada, K. Useful Chiral Packing Materials for High-Performance Liquid Chromatographic Resolution of Enantiomers: Phenylcarbamates of Polysaccharides Coated on Silica Gel. J. Am. Chem. Soc. 1984, 106, 5357–5359, doi:10.1021/ja00330a057.

74. Lin, Z.; Tai, H.-C.; Zhu, C.; Fabiano, A.; Borges-Muñoz, A.; Ye, Y.K.; He, B.L. Evaluation of a Polysaccharide-Based Chiral Reversed-Phase Liquid Chromatography Screen Strategy in Pharmaceutical Analysis. J. Chromatogr. A 2021, 1645, 462085–462094, doi:10.1016/j.chroma.2021.462085.

75. Chankvetadze, B. Polysaccharide-Based Chiral Stationary Phases for Enantioseparations by High-Performance Liquid Chromatography: An Overview. In Chiral Separations, 3rd ed.; Scriba, G.K.E., Ed.; Methods in Molecular Biology; Humana: New York, NY, USA; Springer: New York, NY, USA, 2019; pp. 93–126. ISBN 9781493994380.

76. Mitchell, C.R.; Armstrong, D.W. Cyclodextrin-Based Chiral Stationary Phases for Liquid Chromatography. In Chiral Separation, 1st ed.; Gübitz, G., Schmid, M.G., Eds.; Methods in Molecular Biology; Humana Press, 2004; Volume 243, pp. 61–112. ISBN 9781592596485.

77. Dai, Y.; Wang, S.; Tang, W.; Ng, S.-C. Cyclodextrin-Based Chiral Stationary Phases for High-Performance Liquid Chroma-tography. In Modified Cyclodextrins for Chiral Separation, 1st ed.; Tang, W., Ng, S.-C., Sun, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 67–101. ISBN 9783642376481.

78. Li, X.; Wang, Y. HPLC Enantioseparation on Cyclodextrin-Based Chiral Stationary Phases. In Chiral Separations, 3rd ed.; Gerhard, K.E. Scriba, Ed.; Methods in Molecular Biology; Humana: Totowa, NJ, USA; Springer: New York, NY, USA, 2019; pp. 159–169. ISBN 978-1-4939-9438-0.

79. Shuang, Y.; Liao, Y.; Zhang, T.; Li, L. Preparation and Evaluation of an Ethylenediamine Dicarboxyethyl Diamido-Bridged Bis(β-Cyclodextrin)-Bonded Chiral Stationary Phase for High Performance Liquid Chromatography. J. Chromatogr. A 2020, 1619, 460937–460947, doi:10.1016/j.chroma.2020.460937.

80. Yoshikawa, K.; Furuno, M.; Tanaka, N.; Fukusaki, E. Fast Enantiomeric Separation of Amino Acids Using Liquid Chroma-tography/Mass Spectrometry on a Chiral Crown Ether Stationary Phase. J. Biosci. Bioeng. 2020, 130, 437–442, doi:10.1016/j.jbiosc.2020.05.007.

81. Upmanis, T.; Kažoka, H.; Arsenyan, P. A Study of Tetrapeptide Enantiomeric Separation on Crown Ether Based Chiral Sta-tionary Phases. J. Chromatogr. A 2020, 1622, 461152–441161, doi:10.1016/j.chroma.2020.461152.

82. Nakano, Y.; Taniguchi, M.; Fukusaki, E. High-Sensitive Liquid Chromatography-Tandem Mass Spectrometry-Based Chiral Metabolic Profiling Focusing on Amino Acids and Related Metabolites. J. Biosci. Bioeng. 2019, 127, 520–527, doi:10.1016/j.jbiosc.2018.10.003.

83. Carrão, D.B.; Perovani, I.S.; de Albuquerque, N.C.P.; de Oliveira, A.R.M. Enantioseparation of Pesticides: A Critical Review. TrAC Trends Anal. Chem. 2020, 122, 115719–115734, doi:10.1016/j.trac.2019.115719.

84. Teixeira, J.; Tiritan, M.E.; Pinto, M.M.M.; Fernandes, C. Chiral Stationary Phases for Liquid Chromatography: Recent Devel-opments. Molecules 2019, 24, 865, doi:10.3390/molecules24050865.

85. Fernandes, C.; Phyo, Y.Z.; Silva, A.S.; Tiritan, M.E.; Kijjoa, A.; Pinto, M.M.M. Chiral Stationary Phases Based on Small Mol-ecules: An Update of the Last 17 Years. Sep. Purif. Rev. 2017, 47, 89–123, doi:10.1080/15422119.2017.1326939.

86. Kohout, M.; Hovorka, Š.; Herciková, J.; Wilk, M.; Sysel, P.; Izák, P.; Bartůněk, V.; von Baeckmann, C.; Pícha, J.; Frühauf, P. Evaluation of Silica from Different Vendors as the Solid Support of Anion-exchange Chiral Stationary Phases by Means of Preferential Sorption and Liquid Chromatography. J. Sep. Sci. 2019, 42, 3653–3661, doi:10.1002/jssc.201900731.

87. Hsiao, S.-H.; Ishii, C.; Furusho, A.; Hsieh, C.-L.; Shimizu, Y.; Akito, T.; Mita, M.; Okamura, T.; Konno, R.; Ide, T.; et al. De-termination of Phenylalanine Enantiomers in the Plasma and Urine of Mammals and D-Amino Acid Oxidase Deficient Rodents Using Two-Dimensional High-Performance Liquid Chromatography. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140540–140547, doi:10.1016/j.bbapap.2020.140540.

88. Bäurer, S.; Ferri, M.; Carotti, A.; Neubauer, S.; Sardella, R.; Lämmerhofer, M. Mixed-Mode Chromatography Characteristics of Chiralpak ZWIX(+) and ZWIX(−) and Elucidation of Their Chromatographic Orthogonality for LC × LC Application. Anal. Chim. Acta 2020, 1093, 168–179, doi:10.1016/j.aca.2019.09.068.

89. Geibel, C.; Dittrich, K.; Woiwode, U.; Kohout, M.; Zhang, T.; Lindner, W.; Lämmerhofer, M. Evaluation of Superficially Porous Particle Based Zwitterionic Chiral Ion Exchangers against Fully Porous Particle Benchmarks for Enantioselective Ultra-High Performance Liquid Chromatography. J. Chromatogr. A 2019, 1603, 130–140, doi:10.1016/j.chroma.2019.06.026.

90. Zhu, L.; Zhu, L.; Sun, X.; Wu, Y.; Wang, H.; Cheng, L.; Shen, J.; Ke, Y. Novel Chiral Stationary Phases Based on 3,5-dimethyl Phenylcarbamoylated B-cyclodextrin Combining Cinchona Alkaloid Moiety. Chirality 2020, 32, 1080–1090, doi:10.1002/chir.23237.

91. Bajtai, A.; Ilisz, I.; Howan, D.H.O.; Tóth, G.K.; Scriba, G.K.E.; Lindner, W.; Péter, A. Enantioselective Resolution of Biologically Active Dipeptide Analogs by High-Performance Liquid Chromatography Applying Cinchona Alkaloid-Based Ion-Exchanger Chiral Stationary Phases. J. Chromatogr. A 2020, 1611, 460574–460586, doi:10.1016/j.chroma.2019.460574.

92. Horak, J.; Lämmerhofer, M. Stereoselective Separation of Underivatized and 6-Aminoquinolyl-N-Hydroxysuccinimidyl Car-bamate Derivatized Amino Acids Using Zwitterionic Quinine and Quinidine Type Stationary Phases by Liquid Chromatog-raphy–High Resolution Mass Spectrometry. J. Chromatogr. A 2019, 1596, 69–78, doi:10.1016/j.chroma.2019.02.060.

93. Horak, J.; Lämmerhofer, M. Derivatize, Racemize, and Analyze—An Easy and Simple Procedure for Chiral Amino Acid Standard Preparation for Enantioselective Metabolomics. Anal. Chem. 2019, 91, 7679–7689, doi:10.1021/acs.analchem.9b00666.

94. Pucciarini, L.; González-Ruiz, V.; Zangari, J.; Martinou, J.-C.; Natalini, B.; Sardella, R.; Rudaz, R. Development and Validation of a Chiral UHPLC-MS Method for the Analysis of Cysteine Enantiomers in Biological Samples. J. Pharm. Biomed. Anal. 2020, 177, 112841–112849, doi: 10.1016/j.jpba.2019.112841.

95. Seki, T.; Sato, M.; Konno, A.; Hirai, H.; Kurauchi, Y.; Hisatsune, A.; Katsuki, H. D-Cysteine Promotes Dendritic Development in Primary Cultured Cerebellar Purkinje Cells via Hydrogen Sulfide Production. Mol. Cell. Neurosci. 2018, 93, 36–47, doi: 10.1016/j.mcn.2018.10.002.

96. Kimura, R.; Tsujimura, H.; Tsuchiya, M.; Soga, S.; Ota, N.; Tanaka, A.; Kim, H. Development of a Cognitive Function Marker Based on D-Amino Acid Proportions Using New Chiral Tandem LC-MS/MS Systems. Sci. Rep. 2020, 10, 804–817, doi:10.1038/s41598-020-57878-y.

97. Kimura, T.; Hesaka, A.; Yoshitaka, I. Utility of D-Serine Monitoring in Kidney Disease. Biochim. Biophys. Acta Proteins Proteom. 2020, 1868, 140449–140454, doi:10.1016/j.bbapap.2020.140449.

98. Furusho, A.; Koga, R.; Akita, T.; Mita, M.; Kimura, T.; Hamase, K. Three-Dimensional High-Performance Liquid Chromato-graphic Determination of Asn, Ser, Ala, and Pro Enantiomers in the Plasma of Patients with Chronic Kidney Disease. Anal. Chem. 2019, 18, 11569–11575, doi:10.1021/acs.analchem.9b01615.

99. Ilisz, I.; Orosz, T.; Péter, A. High-Performance Liquid Chromatography Enantioseparations Using Macrocyclic Glycopep-tide-Based Chiral Stationary Phases: An Overview. In Chiral Separations, 3rd ed.; Scriba, G.K.E., Ed.; Methods in Molecular Biology; Humana: New York, NY, USA; Springer: New York, NY, USA, 2019; pp. 201–237. ISBN 9781493994380.

100. Mazzoccanti, G.; Manetto, S.; Ricci, A.; Cabri, W.; Orlandin, A.; Catani, M.; Felletti, S.; Cavazzini, A.; Ye, M.; Ritchie, H.; et al. High–Throughput Enantioseparation of Nα–Fluorenylmethoxycarbonyl Proteinogenic Amino Acids through Fast Chiral Chromatography on Zwitterionic-Teicoplanin Stationary Phases. J. Chromatogr. A 2020, 1624, 461235–461246, doi:10.1016/j.chroma.2020.461235.

101. West, C. Recent Trends in Chiral Supercritical Fluid Chromatography. Trends Anal. Chem. 2019, 120, 115648–115657, doi:10.1016/j.trac.2019.115648.

102. Kaplitz, A.; Mostafa, M.E.; Calvez, S.A.; Edwards, J.L.; Grinias, J.P. Two-dimensional Separation Techniques Using Super-critical Fluid Chromatography. J. Sep. Sci. 2021, 44, 426–437, doi:10.1002/jssc.202000823.

103. Jakubec, P.; Douša, M.; Nováková, L. Supercritical Fluid Chromatography in Chiral Separations: Evaluation of Equivalency of Polysaccharide Stationary Phases. J. Sep. Sci. 2020, 43, 2675–2689, doi:10.1002/jssc.202000085.

104. Lipka, E.; Dascalu, A.-E.; Messara, Y.; Tsutsqiridze, E.; Farkas, T.; Chankvetadze, B. Separation of Enantiomers of Native Amino Acids with Polysaccharide-Based Chiral Columns in Supercritical Fluid Chromatography. J. Chromatogr. A 2019, 1585, 207–212, doi:10.1016/j.chroma.2018.11.049.

105. Miller, L.; Yue, L. Chiral Separation of Underivatized Amino Acids in Supercritical Fluid Chromatography with Chiral Crown Ether Derived Column. Chirality 2020, 32, 981–989, doi:10.1002/chir.23204.

106. Schurig, V. Gas Chromatographic Enantioseparation of Derivatized α-Amino Acids on Chiral Stationary Phases—Past and Present. J. Chromatogr. B 2011, 879, 3122–3140, doi:10.1016/j.jchromb.2011.04.005.

107. Xie, S.-M.; Chen, X.-X.; Zhang, J.-H.; Yuan, L.-M. Gas Chromatographic Separation of Enantiomers on Novel Chiral Stationary Phases. Trends Anal. Chem. 2020, 124, 115808–115828, doi:10.1016/j.trac.2020.115808.

108. Yu, R.B.; Quirino, J.P. Chiral Selectors in Capillary Electrophoresis: Trends during 2017–2018. Molecules 2019, 24, 1135, doi:10.3390/molecules24061135.

109. Koster, N.; Clark, C.P.; Kohler, I. Past, Present, and Future Developments in Enantioselective Analysis Using Capillary Elec-tromigration Techniques. Electrophoresis 2021, 42, 38–57, doi:10.1002/elps.202000151.

110. Konjaria, M.-L.; Scriba, G.K.E. Enantioseparation of Analogs of the Dipeptide Alanyl-Phenylalanine by Capillary Electropho-resis Using Neutral Cyclodextrins as Chiral Selectors. J. Chromatogr. A 2020, 1623, 461158–461165, doi:10.1016/j.chroma.2020.461158.

111. Konjaria, M.-L.; Scriba, G.K.E. Enantioseparation of Alanyl-Phenylalanine Analogs by Capillary Electrophoresis Using Nega-tively Charged Cyclodextrins as Chiral Selectors. J. Chromatogr. A 2020, 1632, 461585–461474, doi:10.1016/j.chroma.2020.461585.

112. Greño, M.; Castro-Puyana, M.; Marina, M.L. Enantiomeric Separation of Homocysteine and Cysteine by Electrokinetic Chromatography Using Mixtures of γ-Cyclodextrin and Carnitine-Based Ionic Liquids. Microchem. J. 2020, 157, 105070–105078, doi:10.1016/j.microc.2020.105070.

113. Zhang, Y.; Hu, X.; Wang, Q.; He, P. Investigation of Hydroxypropyl-β-Cyclodextrin-Based Synergistic System with Chiral Nematic Mesoporous Silica as Chiral Stationary Phase for Enantiomeric Separation in Microchip Electrophoresis. Talanta 2020, 218, 121121–121128, doi:10.1016/j.talanta.2020.121121.

114. Lee, S.; Kim, S.-J.; Bang, E.; Na, Y.-C. Chiral Separation of Intact Amino Acids by Capillary Electrophoresis-Mass Spectrometry Employing a Partial Filling Technique with a Crown Ether Carboxylic Acid. J. Chromatogr. A 2019, 1586, 128–138, doi:10.1016/j.chroma.2018.12.001.

115. Liu, L.; Bao, P.; Qiao, J.; Zhang, H.; Qi, L. Chiral Ligand Exchange Capillary Electrophoresis with L-Dipeptides as Chiral Ligands for Separation of Dns-D,L-Amino Acids. Talanta 2020, 217, 121069–121075, doi:10.1016/j.talanta.2020.121069.

116. Feng, W.; Qiao, J.; Li, D.; Qi, L. Chiral Ligand Exchange Capillary Electrochromatography with Dual Ligands for Enanti-oseparation of D, L-Amino Acids. Talanta 2019, 194, 430–436, doi:10.1016/j.talanta.2018.10.059.

117. Xu, Z.; Guan, J.; Shao, H.; Fan, S.; Li, X.; Shi, S.; Yan, F. Combined Use of Cu(II)-L-Histidine Complex and β-Cyclodextrin for the Enantioseparation of Three Amino Acids by CE and a Study of the Synergistic Effect. J. Chromatogr. Sci. 2020, 58, 969–975, doi:10.1093/chromsci/bmaa058.

118. Evans, K.; Wang, X.; Roper, M.G. Chiral micellar electrokinetic chromatographic separation for determination of L- and D-primary amines released from murine islets of Langerhans. Anal. Methods 2019, 11, 1276–1283, doi:10.1039/C8AY02471E.

119. Moldovan, R.-C.; Bodoki, E.; Servais, A.-C.; Crommen, J.; Oprean, R.; Fillet, M. Selectivity Evaluation of Phenyl Based Sta-tionary Phases for the Analysis of Amino Acid Diastereomers by Liquid

Chromatography Coupled with Mass Spectrometry. J. Chromatogr. A 2019, 1590, 80–87, doi:10.1016/j.chroma.2018.12.068.

120. Pérez-Míguez, R.; Bruyneel, B.; Castro-Puyana, M.; Marina, M.L.; Somsen, G.W.; Domínguez-Vega, E. Chiral Discrimination of DL-Amino Acids by Trapped Ion Mobility Spectrometry after Derivatization with (+)-1-(9-Fluorenyl)Ethyl Chloroformate. Anal. Chem. 2019, 91, 3277–3285, doi:10.1021/acs.analchem.8b03661.

121. Goto, J.; Goto, N.; Nambara, T. New Type of Derivatisation Reagents for Liquid Chromatographic Resolution of Enantiomeric Hydroxyl Compounds. Chem. Pharm. Bull. 1982, 30, 4597–4599, doi:10.1248/cpb.30.4597.

122. Miyano, S.; Okada, S.-I.; Hotta, H.; Takeda, M.; Kabuto, C.; Hashimoto, H. Optical Resolution of 2′-Methoxy-1,1′-Binaphthyl-2-Carboxylic Acid, and Application to Chiral Derivatizing Agent for HPLC Separation of Enan-tiomeric Alcohols and Amines. Bull. Chem. Soc. Jpn. 1989, 62, 1528–2533, doi:10.1246/bcsj.62.1528.

123. Harada, M.; Karakawa, S.; Yamada, N.; Miyano, H.; Shimbo, K. Biaryl Axially Chiral Derivatizing Agent for Simultaneous Separation and Sensitive Detection of Proteinogenic Amino Acid Enantiomers Using Liquid Chromatography–Tandem Mass Spectrometry. J. Chromatogr. A 2019, 1593, 91–101, doi:10.1016/j.chroma.2019.01.075.

124. Harada, M.; Karakawa, S.; Miyano, H.; Shimbo, K. Simultaneous Analysis of d,l-Amino Acids in Human Urine Using a Chi-rality-Switchable Biaryl Axial Tag and Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry. Symmetry 2020, 12, 913–927, doi:10.3390/sym12060913.

125. Han, Y.; Jin, M.-N.; Xu, C.-Y.; Qian, Q.; Nan, J.; Jin, T.; Min, J.Z. Evaluation of Chiral Separation Efficiency of a Novel OTPTHE Derivatization Reagent: Applications to Liquid-chromatographic Determination of DL-serine in Human Plasma. Chirality 2019, 31, 1043–1052, doi:10.1002/chir.23133.

126. Russo, M.S.T.; Napylov, A.; Paquet, A.; Vuckovic, D. Comparison of N-Ethyl Maleimide and N-(1-Phenylethyl) Maleimide for Derivatization of Biological Thiols Using Liquid Chromatography-Mass Spectrometry. Anal. Bioanal. Chem. 2020, 412, 1639–1652, doi:10.1007/s00216-020-02398-x.

Chapter 4

Analysis of aging proteins

4.1 Aging proteins

Aging can cause different changes in protein sequences, such as loss of proteolytic capacity [36–38], increased surface hydrophobicity [36], the appearance of post-translational modifications such as oxidation [36, 39, 40], phosphorylation [39], methylation [39, 41], deamination [40, 42], and acylation (particularly acetylation [39], carbonylation [40], carboxymethylation [40]) [36], and racemization of amino acids [43–48]. 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, 49–52]. These racemases can proceed to free amino acid isomerization before or during peptide elongation [53]. Second, *via* a non-enzymatic process, the amino acids isomerization is possible by a succinimidyl intermediate to form an intramolecular cyclization [54]. 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 [44]. 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 [55, 56]. To prevent this amino acids racemization, a DCl/D_2O condition is privileged [7, 57–59]. 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 [57].

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 [36]. 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 [36]. 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 [40, 42]. 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 [36]. 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 [36]. 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 [60]. 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) [61]. 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 [61, 62].

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) [61]. The most common collagen is type I with two heterotrimer chains α 1 and 2 (Figure 4.1) that represents 90% of total collagens [62]. 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 [63]. In addition, the presence of D-amino acids in aging collagen [64–67] 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 [68, 69].

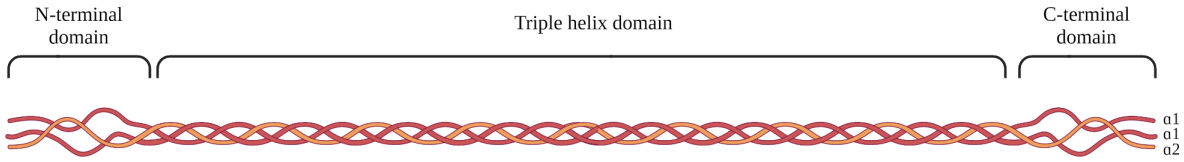


FIGURE 4.1: Representation of the collagen type I structure
Created with BioRender.com

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 [61].

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.

The chiral proteomic analysis applied to aging collagens by LC-MS: Amino acid racemization, post-translational modifications, and sequence degradations during the aging process
Marine Morvan, Ivan Mikšík
Analytica Chimica Acta, **2023**, 1262, 341260.

Supplementary data is available online

Article

The chiral proteomic analysis applied to aging collagens by LC-MS: amino acid racemization, post-translational modifications, and sequence degradations during the aging process

Marine Morvan[1,2,*], Ivan Mikšík[2]

[1] Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20 Prague, Czech Republic

[2] Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 532 10 Pardubice, Czech Republic

* corresponding author: marine.morvan@fgu.cas.cz (M. M.)

Citation: Morvan, M.; Mikšík, I., 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. https://doi.org/10.1016/j.aca.2023.341260

Abstract: Collagen is the most abundant protein in the animal and human bodies, and it is not exempt from this aging phenomenon. Some age-related changes may appear on collagen sequences, such as increased surface hydrophobicity, the appearance of post-translational modifications, and amino acids racemization. This study has shown that the protein hydrolysis under deuterium conditions is privileged to limit the natural racemization during the hydrolysis. Indeed, under the deuterium condition, the homochirality of recent collagens is preserved whose amino acids are found in their L-form. However, in aging collagen, a natural amino acid racemization was observed. These results confirmed that the % D-amino acids are progressive according to age. The collagen sequence is degraded over time, and a fifth of the sequence information is lost during aging. Post-translational modifications (PTMs) in aging collagens can be a hypothesis to explain the modification of the hydrophobicity of the protein with the decrease of hydrophilic groups and the increase of hydrophobic groups. Finally, the exact positions of D-amino acids and PTMs have been correlated and elucidated.

Keywords: aging, chiral separation, collagen, D-amino acids, liquid chromatography, mass spectrometry.

1. Introduction

Aging is a natural and uncontrolled phenomenon that the animal body, including human, undergoes. *In vivo*, molecules are not exempt from this aging process. Indeed, some age-related protein changes may appear like a loss of proteolytic capacity [1]–[3], increased surface hydrophobicity [1], appearance of post-translational modifications such as oxidation [1], [4], [5], phosphorylation [4], methylation [4], [6], deamination [5], [7], and acylation (particularly acetylation [4], carbonylation [5], carboxymethylation [5]) [1], and racemization of amino acids [8]–[13].

Recent animal and human proteins are found in their L-amino acid forms. However, during the natural aging process, an amino acid racemization can take place in proteins *via* amino acids racemases (enzymatic process) or/and *via* a succinimidyl intermediate (non-enzymatic process) [14]. This amino acid racemization affects the three-dimensional protein conformation and can induce some aggregation [15]–[18], disorder [16], malfunctions [16], misfolding [16], and toxicity [16]. In addition, the amount of D-amino acids in proteins is progressive according to age. In 1975, D-Asp enrichment in human tooth enamel proteins *via* natural racemization was measured as a content of approximately 0,1 % per year [19]. Furthermore, some D-amino acids were related to aging human proteins and can be linked to aging diseases. It is the case of D-Asp in elastin linked to arteriosclerosis [20], [21], D-Asp in β-amyloid linked to Alzheimer's disease [22], D-Asp, D-Asn, D-Ser, and D-Thr in α-crystallin linked to cataract [9]. The complete list and the exact position of these D-amino acids in till now described protein sequences are recently summarized [14].

D-amino acids are also found in aging and/or ancient proteins and can be used for the age-estimation. Indeed, different studies used the quantification of D-Asp from bone [23], lens [24], and tooth [25], [26] for age-estimation, using a HCl/H_2O protein hydrolysis. However, in a hydrogenated environment, a natural racemization of L-amino acids to its enantiomer can take place. Different kinetics were observed according to the nature of amino acids [27], [28]. To prevent this amino acids racemization during the acidic hydrolysis, a DCl/D_2O condition is privileged [14], [24], [29], [30]. Indeed, in a deuterium environment, the hydrogen on the alpha carbon was exchanged with a deuterium atom and decreases considerably the racemization. Yasunaga *et al.* compared the racemization of D-Asp for age-estimation under HCl/H_2O and DCl/D_2O hydrolysis conditions. Errors in age estimations were decreased by at least one-half compared with that of the HCl/H_2O hydrolysis method [24].

Recently, chiral chromatographic and electrophoretic separation methods for the enantioseparation of (un-)derivatized D- and L-amino acids were described [14]. Mass spectrometry optimization can also enhance the discrimination of derivatized D- and L-amino acids, specifically using ion-mobility mass spectrometry [31], [32].

Collagen is the most abundant protein in animal bodies, including humans, and is the major structural component of the extracellular matrix present in tissues and organs [33], [34]. To date, 28 different types of collagens were described [33]. However, the most common of them is collagen type I which is found in 90 % [34]. Its three-dimensional structure is a triple-helix form made up of polypeptide chains α 1 and 2 (COL1A1 and COL1A2). The appearance of D-amino acids in the collagen protein reverses the rotation of the triple helix and decreases the helix-helix interactions until the overall destabilization of the three-dimensional structure [35], [36]. However, D-amino acids can also be involved in the cross-linking of assembly molecular material. Indeed, in aging collagen, the presence of D-Ala and D-Glu destabilized the structure while D-Lys stabilized it [2].

For this work, the amino acids racemization in collagen samples (type I alpha 1 and 2) from different organisms (bovine and rat) and at different ages (recent, 4-months and 4-years) were studied. Artificial aging was applied to standard bovine collagens to mimic natural aging. A comparison of HCl/H_2O and DCl/D_2O effects on collagen hydrolysis was studied. The amount of D-amino acids in bovine collagen after artificial aging was compared to that obtained in rat collagen all along the natural aging at the same age. After enzymatic treatment, arising peptides at different ages were compared. Common peptides between recent and aging collagen were identified as peptides from enzymatic digestions, and non-common peptides come from aging degradations. The evolution of post-translational modifications on the collagen sequences was also studied according to age. Finally, the exact position of D-amino acids and amino acids modified by post-translational modifications were elucidated.

2. Materials and methods

2.1. Chemicals

Acetic acid (AcOH, ≥99,8%), acetic acid-d4 (DOAc, ≥99,9 atom %D), acetonitrile (ACN, ≥99,9%), ammonium hydrogen carbonate (NH_4HCO_3, ≥99%), chloroform (99,8%), α-chymotrypsin from bovine pancrease, type II (≥40 units/mg protein), collagen from bovine achilles tendon, type I (bovine collagens), deuterium chloride solution (DCl, ≥99,9 atom %D), deuterium oxide (D_2O, 99.9 atom %D), dimethyl sulfoxide (DMSO, ≥99,9%), ethyl acetate (EtOAc, ≥99,7%), formic acid (FA, ≥98%), heptafluorobutyric acid (HFBA, ≥99,5%), methanol (MeOH), Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide (FDAA, Marfey's reagent, ≥99%), pepsin from porcine stomach mucosa (≥2,500 units/mg protein (E1%/280)), phosphate buffer solution (1M, pH 7,4), proteinase K from *Tritirachium album* (≥30 units/mg protein), sodium bicarbonate ($NaHCO_3$, ≥99,5%), tris(hydroxymethyl)aminomethane (Tris, ≥99,8%), trypsin from porcine pancreas (13,000-20,000 BAEE units/mg protein) were obtained from Sigma-Aldrich (St. Louis, MO, USA). Acetone (≥99,5%) and sodium chloride (NaCl) were from Penta (Chrudim, Czech

Republic). Ultrapure water was obtained using a Milli-Q system from Millipore (Bedford, MA, USA)

Hereditary hypertriglyceridemic (HTG) rats were originally selected from Wistar rats as previously described [37]. HTG rats aged 4-months (bred in the Institute of Physiology, Academy of Sciences of the Czech Republic, Prague) were studied. Age-matched normotensive Wistar-derived Lewis rats were used as controls. Rats were housed under controlled conditions (temperature 23±1 °C, 12 h : 12 h light-dark cycle). They were fed a standard rat chow (Velaz, ST-l) containing 0.4% sodium chloride. Water and food were available *ad libitum*.

Amino acid: D-α-alanine (D-Ala, ≥98%), D-arginine (D-Arg, ≥98%), L-arginine (L-Arg, ≥98%), D-aspartic acid (D-Asp, 99%), D-cysteine hydrochloride monohydrate (D-Cys, ≥98%), L-cysteine (L-Cys, 97%), D-glutamic acid (D-Glu, ≥99%), glycine (Gly, ≥99%), D-histidine monohydrochloride monohydrate (D-His, ≥98%), DL-5-hydroxylysine hydrochloride (DL-Hyl, ≥98%), trans-4-hydroxy-D-proline (D-Hyp, 97%), DL-isoleucine (DL-Ile, 99%), D-leucine (D-Leu, 99%), D-lysine (D-Lys, ≥98%), DL-methionine sulfoxide (DL-Met sulfoxide, ≥98.5%), L-methionine sulfoxide (L-Met sulfoxide, ≥98%), D-phenylalanine (D-Phe, ≥98%), D-proline (D-Pro, ≥99%), L-proline (L-Pro, ≥99%), DL-serine (DL-Ser, ≥98%), D-serine (D-Ser, ≥98%), D-threonine (D-Thr, ≥98%), L-threonine (L-Thr, ≥98%), D-tryptophan (D-Trp, ≥98%), L-tryptophan (L-Trp, ≥98%), D-tyrosine (D-Tyr, 99%), L-tyrosine (L-Tyr, ≥98%), and D-valine (D-Val, ≥98%) were obtained from Sigma-Aldrich (St. Louis, MO, USA). DL-α-alanine (DL-Ala), L-glutamic acid (L-Glu), DL-leucine (DL-Leu), DL-methionine (DL-Met), and DL-β-phenyl-α-alanine (DL-Phe) were obtained from Reachem (Mississauga, Canada). L-aspartic acid (L-Asp, 98%), L-histidine (L-His, 98%), L-methionine (L-Met, 98%) and L-valine (L-Val, 98%) were obtained from Roana (Budapest, Hungary). L-isoleucine (L-Ile, *allo*-isoleucine free), L-lysine monohydrochloride (L-Lys) were obtained from Calbiochem (San Diego, USA). L-hydroxyproline NH.CH2.CH(OH).CH2 (L-Hyp, ≥98,5%) was obtained from BDH (BDH Chemicals Ltd., Poole, UK).

2.2. Sample preparations

2.2.1. Standard and proteinogenic amino acids preparation

Standard amino acids preparation method: standard amino acids were dissolved in D_2O at 50 mM.

Amino acids derivatization method: To standard and proteinogenic amino acid solutions, 100 µL of $NaHCO_3$ (1 M) in D_2O was added, and 200 µL of 38.7 mM Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide (FDAA) in acetone. Solutions were vortexed and incubated at 40°C for 60 min. Reactions were quenched by the addition of 50 µL of DCl (6 M) in D_2O. Compounds of interest are extracted by ethyl acetate (0.5 mL). A saturated sodium chloride solution (300 µL) was added in the aqueous phase to increase the difference in density between the aqueous and organic phases and facilitate the extraction. Undesired hydrogen fluoride formed during the derivatization reaction stays in the aqueous phase and is not injected to limit the HPLC and LC column degradations.

From extracted organic phases, 20 µL of these solutions in 20 µL of DMSO were analyzed by HPLC at 340 nm or LC-MS.

2.2.2. Recent and aging collagens preparation

2.2.2.1. Artificial aging of bovine collagens

Bovine collagen samples (n=6) in phosphate buffer (100 mM, pH 7,4) at 2 $mg.mL^{-1}$ in glass tubes, were placed outside under natural meteorological conditions for 4-months and 4-years.

2.2.2.2. Extraction of natural aging HTG rat collagens

Tendons from the tails of HTG rats (n=2) were scraped with a scalpel, then rinsed three times with 1M NaCl solution, each for 24h at 4°C. Extracted collagens were rinsed with ultrapure water and lyophilized.

2.2.2.3. Enzymatic digestion

Recent and aging bovine and rat collagens type I alpha 1 and 2 were digested successively by pepsin (0.1 $mg.mL^{-1}$ in 3 % AcOH, 50/1 sample/enzyme, w/w) and trypsin (0.02 $mg.mL^{-1}$ in 50 mM NH_4HCO_3, 50/1 sample/enzyme, w/w). Collagen residues were lyophilized after both enzymatic treatments. Bovine and rat 4-months aging collagens produce an insoluble material after these treatments while recent and 4-years aging collagens do not. These insoluble parts were then digested by proteinase K (0.05 $mg.mL^{-1}$ in 10 mM Tris, 50/1 sample/enzyme, w/w), and resulting insoluble parts were digested by chymotrypsin (0.04 $mg.mL^{-1}$ in 50 mM NH_4HCO_3, 50/1 sample/enzyme, w/w). The totality of bovine and rat 4-months aging collagens was soluble after chymotrypsin enzymatic treatment. Each enzymatic digestion was performed overnight at 37°C in the darkroom.

2.2.2.4. Hydrolysis

Proteins and peptides hydrolysis method: intact proteins and each peptide fraction were hydrolyzed in individual glass tubes placed into 10-mL Pierce Reacti-vials (Kimble Chase, Mexico) that had been purged with nitrogen thrice. After 18 h at 130°C in 400 µL 1:1 DCI/(2H4)acetic acid (DCl/DOAc) on the bottom of the vial, 200 µL of D_2O was added to the proteinogenic amino acid solution obtained in each tube. Then, these proteinogenic amino acids were derivatized following the amino acids derivatization method described in section 2.2.1.

2.3. HPLC parameters

2.3.1. HPLC for amino acids analysis

Analyses were performed in a HPLC system from Agilent (Agilent Technologies, Santa Clara, USA). Mobile phase A constituted to 0.1 % of FA in ultrapure water (v/v) and mobile phase B constituted to 0.1 % of FA in ACN (v/v). The chromatographic column was the Phenomenex Aeris

3.6 µm peptide XB-C18, 250 x 2.1 mm (Phenomenex Inc., Torrance, USA). The mobile phase operated in gradient mode, starting at 2 % mobile phase B and increasing for 30 min to reach 8 % B, 110 min to reach 40 %, and then increasing to 100 % B at 125 min, remaining for 20 min, and returning to 2 % B for 10 min maintaining constant until the end of the analysis. The flow rate was set at 0.250 $mL.min^{-1}$, the column was held at ambient temperature (25 °C), the injection volume was 5 µL, and UV detection was performed at 340 nm.

The Agilent HPLC system was connected to a maXis quadrupole time-of-flight (Q-TOF) mass spectrometer with ultrahigh resolution (Bruker Daltonics, Bremen, Germany).

2.3.2. HPLC for peptides analysis

Analyses were performed on an HPLC system from Agilent (Agilent Technologies, Santa Clara, USA) consisting of a degasser, binary pump, autosampler, thermostatted column compartment and multi-wavelength detector. Mobile phase A constituted to 0.1 % of HFBA in ultrapure water (v/v) and mobile phase B constituted to 0.1 % of HFBA in ACN (v/v). Elution was achieved using a linear gradient when separation was started by running the system isocratically for two minutes with 2 % of mobile phase B, followed by a gradient elution to 35 % B at 40. min, Next gradient was 10 minutes to 100 % B. Finally, the column was eluted with 100 % B for 10 min. Equilibration before the next run was achieved by washing with buffer A for 10 min. The chromatographic column was the Jupiter 4 µm Proteo 90Å column (250 x 2 mm ID, Phenomenex Inc., Torrance, USA). The flow-rate was 0.250 $mL.min^{-1}$, the column temperature was held at 25°C and UV absorbance detection was done at 214 nm, injection volume was 20 µL.

Bovine and rat collagen peptides obtained after enzymatic treatments were separated and collected in several fractions, *e.g.* 31 fractions for collagen sample treated by pepsin and trypsin, 41 fractions after proteinase K treatment, and 37 fractions after chymotrypsin treatment. Chromatograms with fraction delimitations are presented in supplementary materials (Figures S1, S2, and S3).

2.4. NanoLC-MS parameters

The nanoHPLC apparatus used for protein digest analysis was a Proxeon Easy-nLC (Proxeon, Odense, Denmark). It was coupled to an ultrahigh resolution MaXis Q-TOF (quadrupole – time of flight) mass spectrometer (Bruker Daltonics, Bremen, Germany) by nanoelectrosprayer. The nLC-MS/MS instruments were controlled with the software packages HyStar 3.2 and micrOTOF-control 3.0. The data were collected and manipulated with the software packages ProteinScape 3.0 and DataAnalysis 4.0 (Bruker Daltonics).

Five microliters of the peptide mixture were injected into an NS-AC-12dp3-C18 Biosphere C18 column (particle size: 3 µm, pore size: 12 nm, length: 200 mm, inner diameter: 75 µm) with an NS-MP-10 Biosphere C18 precolumn (particle size: 5 µm, pore size: 12 nm, length: 20 mm, inner diameter: 100 µm), both manufactured by NanoSeparations (Nieuwkoop, Holland).

The separation of peptides was achieved via a linear gradient between mobile phase A (ultrapure water) and B (ACN), both containing 0.1 % FA (v/v). Separation was started by running the system with 5 % mobile phase B, followed by a gradient elution to 7 % B at 5 min, and 30 % B at 180 min. The next step was a gradient elution to 50 % B in 10 min and then a gradient to 100 % B in 10 min. Finally, the column was eluted with 100 % B for 20 min. Equilibration between the runs was achieved by washing the column with 5 % mobile phase B for 10 min. The flow rate was 0.200 $\mu L.min^{-1}$ and the column was held at ambient temperature (25 °C).

2.5. MS parameters

Online nano-electrospray ionization (easy nano-ESI) was used in positive mode. The ESI voltage was set to +4.5 kV, scan time: 3 Hz. Operating conditions: drying gas (N_2): 4 $L.min^{-1}$; drying gas temperature: 180 °C; nebulizer pressure: 100 kPa. Experiments were performed by scanning from 50 to 2200 m/z. The reference ion used (internal mass lock, chip cube high mass reference (HP-1221), Agilent Technologies, Waldbronn, Germany) was a monocharged ion of $C_{24}H_{19}F_{36}N_3O_6P_3$ (m/z 1221.9906). Mass spectra corresponding to each signal from the total ion current chromatogram were averaged, enabling an accurate molecular mass determination. All HPLC-MS and nanoLC-MS analyses were done in duplicate.

2.6. Peptide identifications

Peptide data were processed using ProteinScape software (Bruker Daltonics, Bremen, Germany). Collagen type I alpha 1 and 2 chains were identified by correlating tandem mass spectra to the NCBI, IPI, and SwissProt databases, using the MASCOT search engine (http://www.matrixscience.com). The taxonomy was limited to *Bos taurus* and *Rattus norvegius* due to the provenance of the samples. SemiTrypsin was chosen as the enzyme parameter. Three missed cleavages were allowed, and an initial peptide mass tolerance of ±15.0 ppm was used for MS and ±0.03 Da for MS/MS analysis. Arginine, asparagine, and glutamine were assumed to be deamidated, methionine to be dioxidized, lysine, methionine, and proline to be oxidized, lysine, serine and threonine to be formylated, lysine to be acetylated, carbamylated, carboxylated and carboxymethylated, arginine, aspartic acid, cysteine, histidine, lysine, serine, threonine, and tyrosine were allowed to be phosphorylated. All these possible modifications were set to be variable. Monoisotopic peptide charge was set at 1+, 2+, and 3+. The Peptide Decoy option was selected during the data search process to remove false-positive results. Only significant hits (accepted as identified on ProteinScape) were selected.

3. Results and discussion

3.1. Effect of hydrolysis conditions

Figure 1 shows the amino acid racemization of recent bovine collagens under HCl/H_2O (Figure 1A) and DCl/D_2O (Figure 1B) hydrolysis conditions. Amino acid racemization takes place only under HCl/H_2O hydrolysis conditions for all amino acids except Ser and Thr which were not

racemized. Methionine, phenylalanine, tryptophan, and tyrosine were not detected. For this work, a DCl/D_2O hydrolysis is privileged to limit the racemization during the hydrolysis and not distort the results when determining the % D-amino acids.

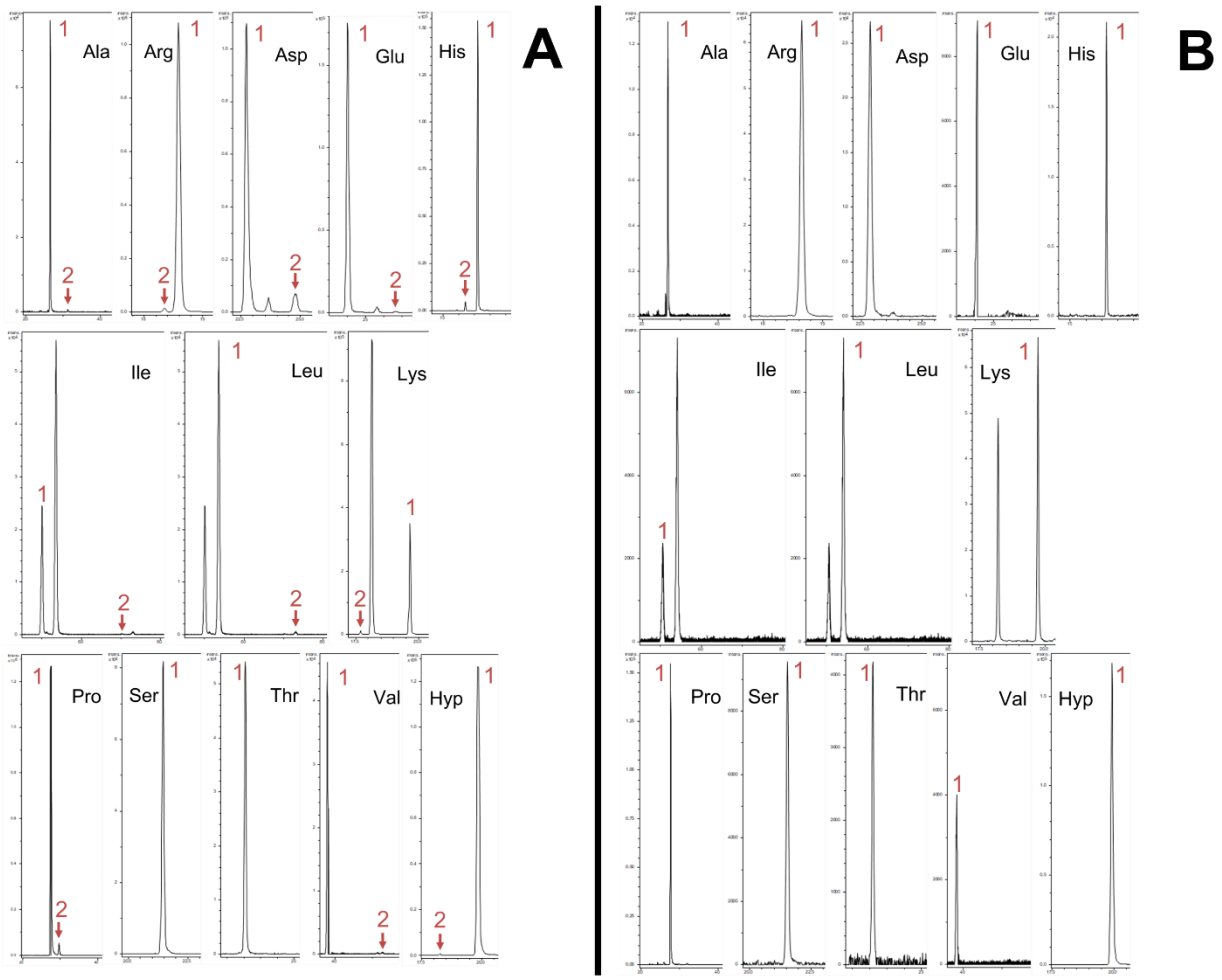


Figure 1: Effect of hydrolysis conditions on proteinogenic amino acid racemization from standard bovine collagens. Hydrolysis conditions: HCl/H_2O (A) and DCl/D_2O (B). Peak 1 identifies amino acids in their L-form and Peak 2 in their D-form. Exact masses of derivatized amino acids: 342,104420 (Ala), 427,168420 (Arg), 386,094250 (Asp), 400,109900 (Glu), 408,126220 (His), 384,151370 (Ile/Leu), 651,211740 (Lys), 368,120070 (Pro), 358,099340 (Ser), 372,114990 (Thr), 370,135720 (Val) and 384,114990 (Hyp).

3.2. D-amino acids in aging collagens

This chiral analysis method, described in section 2.3.1 and newly developed for this work using standard amino acids, was then used to confirm the presence of D-amino acids in bovine collagen after *in vitro* artificial aging, at different ages. In addition, this analytical method was also used for the detection of D-amino acids in rat collagen with *in vivo* natural aging. Preliminary study was

performed to confirm the presence of amino acid only in their L-form in recent bovine collagen (Table S5). During the aging process, the artificial amino acid racemization, applied to bovine aging collagens, mimes the natural amino acid racemization. Most of the different amino acids were racemized over time. It is the case of D-Arg, D-Asp, D-Leu, D-Pro, D-Ser, D-Thr, and D-Val most frequently racemized in 4-years than in 4-months collagen samples. In addition, the percentage of Arg, His, Thr, and Val in their D-form was higher in the older sample (Figures 2 and 3, Figures S4 and S5, Table S1 – S5).

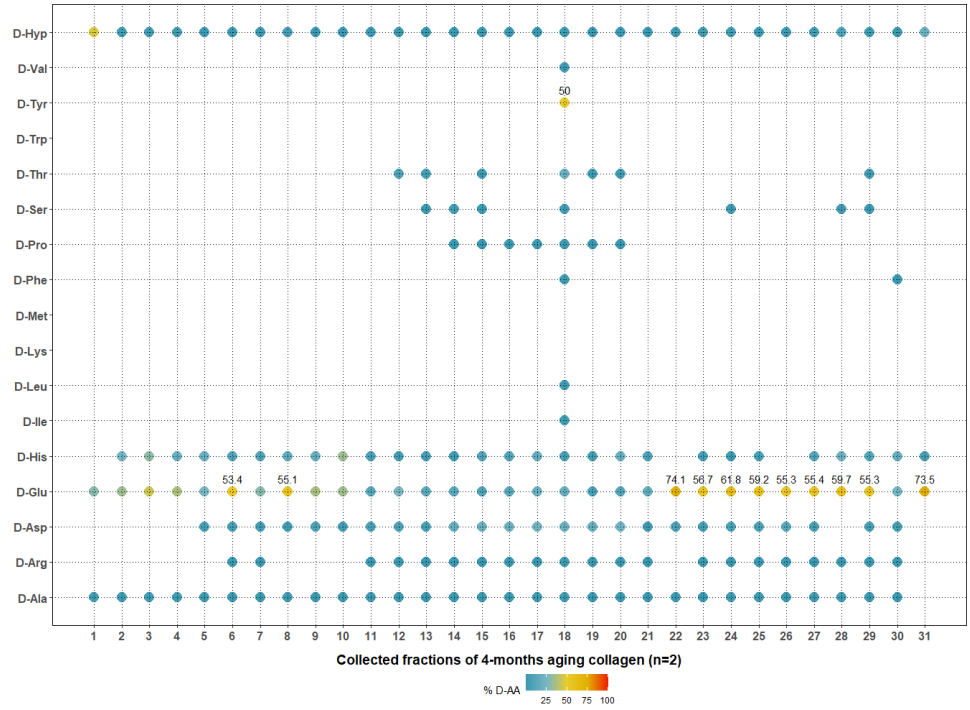


Figure 2: Representation of % D-amino acids in 4-months aging bovine collagen fractions (n=2) after pepsin and trypsin enzymatic treatment. Percentages of D-amino acids superior to 50 % are indicated on the graph. All percentages of D-amino acids are summarized in Table S1. The total % RSD of %D-amino acids was 4.37 %.

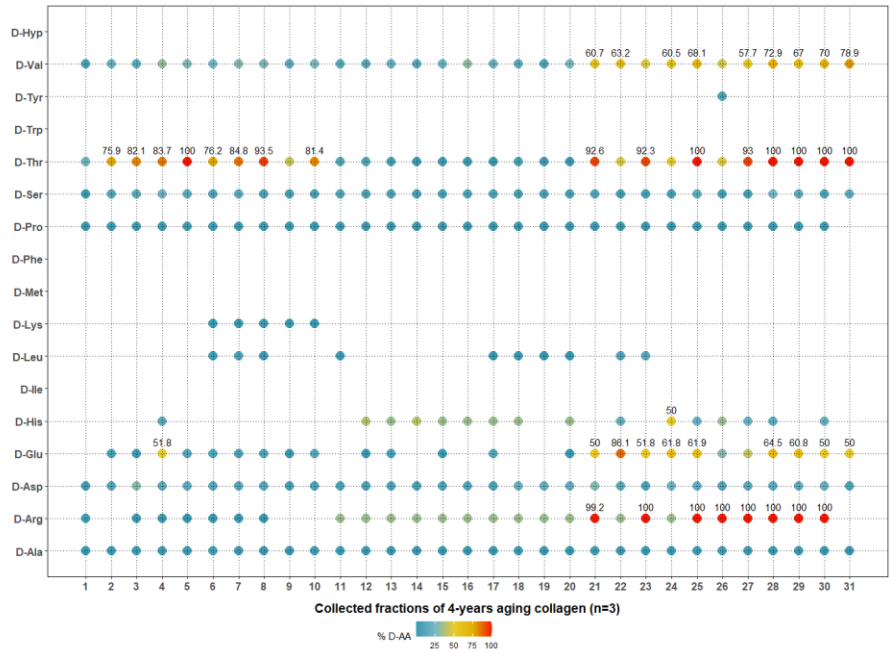


Figure 3: Representation of % D-amino acids in 4-years aging bovine collagen fractions (n=3) after pepsin and trypsin enzymatic treatment. Percentages of D-amino acids superior to 50 % are indicated on the graph. All percentages of D-amino acids are summarized in Table S2. The total % RSD of %D-amino acids was 10.97 %.

Most interesting amino acid fractions of bovine collagens were selected and compared to the same amino acid fractions of rat collagens at the same age. The selection of fractions was made based on the high percentage of amino acid racemization, and the diversity of the nature of amino acids racemized. The comparison of the percentage of D-amino acids is shown in Figures S6 and S7 and table S2 and S6. Into selected fractions of soluble part of collagen after pepsin and trypsin enzymatic treatment, some D-amino acids, *e.g.* D-Arg, D-Ile, D-Leu, D-Phe, D-Pro, and D-Tyr, were present in bovine collagen after artificial aging and not in rat collagen with natural aging. Other amino acids were less frequently present in rat collagen with natural aging, such as D-Asp, D-His, D-Ser, and D-Hyp. Conversely, D-Thr and D-Val were more frequently present in rat collagen with natural *in vivo* aging. Concerning the % D-amino acids, D-Glu, D-His, D-Thr, and D-Val were most racemized in their D-form in rat collagen with natural aging. The same trend was observed in selected fractions of the insoluble part of bovine and rat collagens after pepsin, trypsin, and proteinase K enzymatic treatment (Figure 4 and Table S3 and S7). Some D-amino acids, *e.g.* D-Arg, D-Lys, D-Phe, and D-Pro, were present in bovine collagen and not in rat collagen. D-Asp, D-His, D-Ser, and D-Hyp were less frequently present in rats. Conversely, D-Thr was more frequently present in rat collagen. Concerning the % D-amino acids, D-Glu and D-Thr were most racemized in their D-form in rat collagen with the natural aging, and D-Val was most racemized in bovine collagen after the artificial aging.

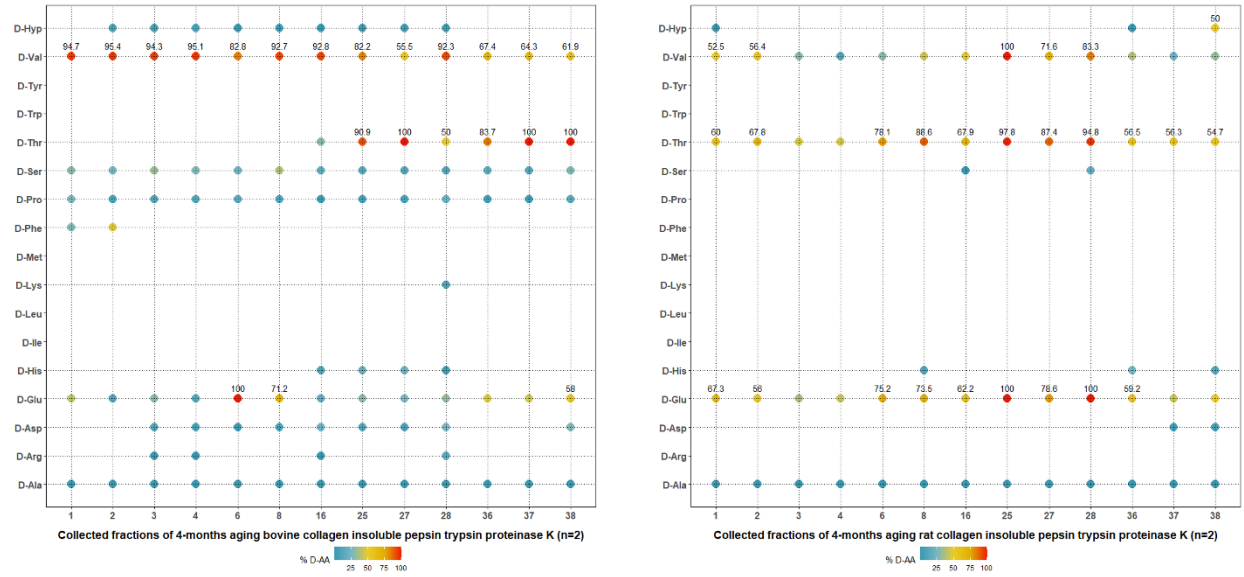


Figure 4: Representation of % D-amino acids in selected fractions of bovine and rat aging collagen after pepsin, trypsin, and proteinase K enzymatic treatment. Percentages of D-amino acids superior to 50 % are indicated on the graph. All percentages of D-amino acids are summarized in Table S3 and S7. The total % RSD of %D-amino acids from rat collagen was 7.07 %.

3.3. Peptide analysis

Peptides in each fraction from recent and aging collagens were analyzed by nanoLC-MS (sections 2.4 and 2.5) and identified (section 2.6)

3.3.1. Position of D-amino acids

The identification of peptides in each collected fraction (for the method see section 2.6) allowed us to determine their position on the bovine and rat collagen sequences (type I alpha 1 and 2). Regrouping the racemization rate of each amino acid in each fraction (section 3.2) and the peptide identification in these same fractions, the exact position of amino acids totally racemized in their D-form have been determined on the bovine and rat collagen sequences (Figure 5, 6, S8 and S9, Table 1 and S8).

In bovine collagens, 21 specific sites of Thr were identified as totally racemized in the D-form for 4-months aging sample and 18 specific sites for 4-years aging sample in the COL1A1 collagen sequence. Their exact positions have shown that 70 % of them were common at both ages. For COL1A2, 14 and 15 Thr specific sites were found in the 4-months and 4-years aging samples respectively, and 45 % were common at both ages. These specific sites of D-Thr totally racemized represent 48 % (4-months) and 41 % (4-years) of total Thr on the bovine COL1A1 sequence. For bovine COL1A2, these D-Thr specific sites represent 32 % (4-months) and 34 % (4-years) of total Thr on the protein sequence. D-Val was exclusively totally racemized in 4-months aging collagens and represent 19 % of total Val on bovine COL1A1 and 4 % on bovine COL1A2 sequences. D-

Arg was entirely racemized in 4-years aging collagens and represented 75 % of total Arg on bovine COL1A1 and 63 % on bovine COL1A2 sequences.

For amino acids totally racemized in fractions of 4-month aging rat collagens selected according to most promising bovine fractions at the same age, 12 specific sites for D-Glu (17 %) and 6 for D-Val (14 %) were identified in COL1A1. In COL1A2, 10 specific sites for D-Glu (16 %) and 6 for D-Val (12 %) were identified.

Table 1: Exact position of D-amino acids totally racemized in their D-form on the bovine collagen sequences at different ages.

		COL1A1	COL1A1			COL1A2	COL1A2
AA	Position	4-months	4-years	AA	Position	4-months	4-years
R	131		✓	R	31		✓
	186		✓		38		✓
	219		✓		41		✓
	239		✓		130		✓
	243		✓		154		✓
	252		✓		163		✓
	267		✓		178		✓
	303		✓		232		✓
	309		✓		280		✓
	311		✓		325		✓
	321		✓		340		✓
	360		✓		379		✓
	369		✓		397		✓
	414		✓		411		✓
	468		✓		421		✓
	471		✓		430		✓
	486		✓		438		✓
	510		✓		448		✓
	527		✓		474		✓
	563		✓		484		✓
	573		✓		571		✓
	597		✓		586		✓
	630		✓		607		✓
	675		✓		643		✓
	684		✓		661		✓
	696		✓		669		✓
	732		✓		673		✓
	744		✓		691		✓
	762		✓		706		✓
	795		✓		739		✓
	801		✓		775		✓
	864		✓		792		✓
	881		✓		828		✓

	917		✓		877		✓
	957		✓		904		✓
	966		✓		924		✓
	969		✓		946		✓
	993		✓		976		✓
	1013		✓		994		✓
	1025		✓		1003		✓
	1035		✓		1015		✓
	1065		✓		1051		✓
	1083		✓		1078		✓
	1092		✓		1116		✓
	1104		✓		1203		✓
	1140		✓		1216		✓
	1167		✓				
	1169		✓				
	1213		✓				
	1435		✓				
T	117	✓		T	29		✓
	327	✓	✓		168	✓	✓
	336	✓	✓		207		✓
	444		✓		211		✓
	453	✓	✓		337	✓	✓
	540	✓	✓		382	✓	✓
	552	✓	✓		424	✓	
	765	✓			561	✓	
	792	✓	✓		601	✓	✓
	873	✓	✓		658	✓	
	912	✓	✓		741	✓	✓
	947	✓	✓		754	✓	
	1038	✓			784	✓	✓
	1068	✓	✓		793	✓	
	1097		✓		832	✓	✓
	1170	✓	✓		852	✓	✓
	1232	✓	✓		858	✓	✓
	1233	✓	✓		1319		✓
	1366	✓			1325		✓
	1380	✓			1331		✓
	1430	✓	✓				
	1431	✓	✓				
	1433	✓	✓				
V	599	✓		V	156	✓	
	606	✓			157	✓	
	686	✓					
	921	✓					
	1056	✓					

Figure 5: Exact position of D-amino acids totally racemized in their D-form on the bovine collagen COL1A1 sequence at 4-months (left) and 4-years (right).

Figure 6: Exact position of D-amino acids totally racemized in their D-form on the bovine collagen COL1A1 sequence at 4-months (left) and on the rat collagen COL1A1 sequence at 4-months (right).

3.3.2. Peptides from enzymatic digestion and aging degradation

After successively enzymatic digestions by pepsin and trypsin, all peptides and residues from recent and aging bovine collagens (COL1A1 and COL1A2) were soluble and identified by nanoLC-MS (section 2.4). Peptides from the insoluble part of collagen protein after pepsin and trypsin treatment at 4-months of aging were not added in the study. Arising peptides in aging soluble collagen samples common to arising peptides in recent collagen were identified as peptides from enzymatic digestions. Other peptides were identified as peptides from aging degradation (Figure 7). For aging bovine collagens type I alpha 1 and 2 (COL1A1 and COL1A2), the number of peptides from aging degradation is higher than arising peptides from enzymatic digestions. Indeed at 4-months, 46 % (COL1A1) and 48 % (COL1A2) of peptides come from aging degradations for both collagen types, and 55 % (COL1A1) and 46 % (COL1A2) at 4-years. The total of arising peptides from enzymatic digestion and aging degradation did not reach 100 % for both collagen chains and at both ages. Approximately, a fifth of the sequence information was lost during aging.

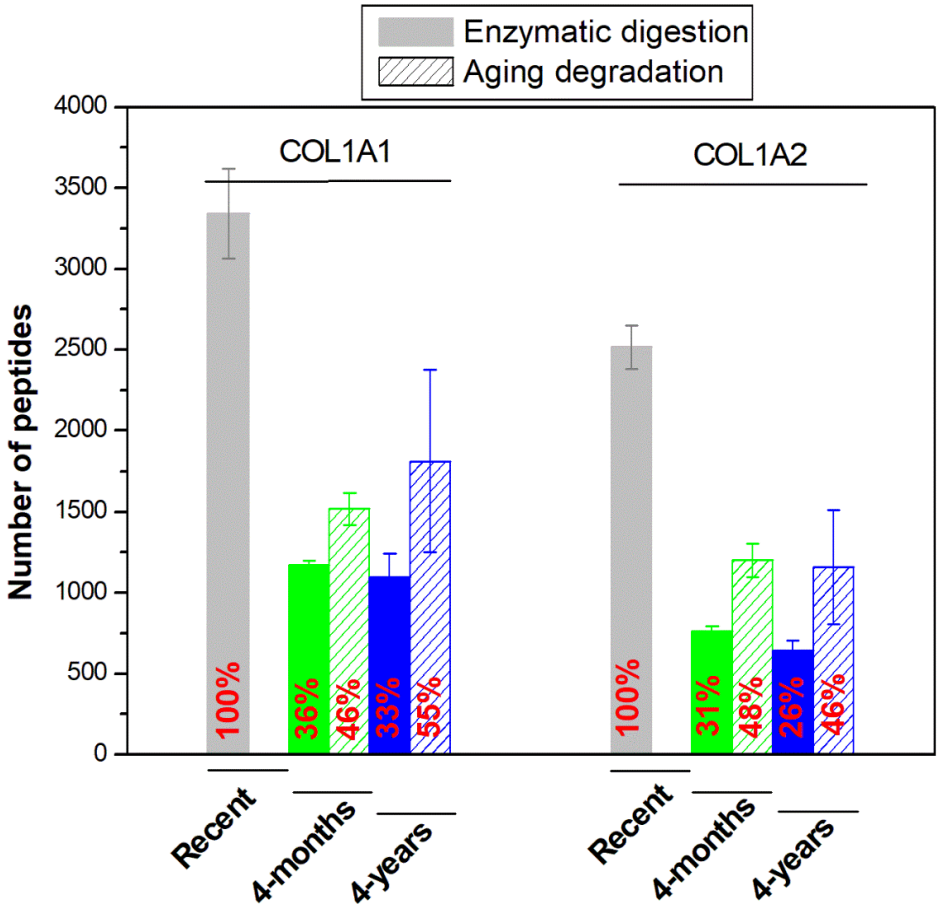


Figure 7: Number of arising peptides from enzymatic digestion and aging degradation in recent, 4-month, 4-year bovine artificial aging collagens.

3.3.3. Post-translational modifications during the aging process

Hydrophilic (*e.g.* oxidation, dioxidation, deamidation, phospho, sulfo) and hydrophobic (*e.g.* acetyl, carbamyl, carboxy, carboxymethyl, carboxyethyl, formyl, methyl) post-translational modifications were studied on the bovine and rat collagen sequences (COL1A1 and COL1A2). Differences according to the organism, age, and artificial/natural aging were observed. The number of modifications and the corresponding percentage according to the total number of possible modified sites are summarized in Table 2 and S9-11 for each sample at different ages.

Based on previous results (see section 3.3.2), about a fifth of bovine sequence information was lost during aging. The post-translational modifications studies show that the percentage of hydrophilic modifications (deamidation, phospho and sulfo) in COL1A1 decreases considerably. Exception for oxidation modifications on lysines, methionines, and prolines (exact positions in Figures S10 – S12, Tables S9, and S10), which increase in 4-months samples, before decreasing in older samples in correlation with the degradation and the loss of sequence information. On the

other hand, the polypeptide chain alpha 2 seems more subject to hydrophobic (acetyl, carbamyl, carboxymethyl, formyl, and methyl) modifications than the COL1A1, during the aging process. Exception for oxidation and phospho modifications which increase (Table 2). In sum, the decrease of hydrophilic groups and the increase of hydrophobic groups during aging may be a hypothesis to explain the evolution of the insolubility of collagen along life, to add to the protein degradations. The comparison between artificial (bovine) and natural (rat) aging on post-translational modifications on collagen sequences is summarized in Table 3. The percentages of modifications were slightly higher under artificial aging.

Table 2: Total number and percentage of post-translational modifications on bovine collagen COL1A1 and COL1A2 sequences at different artificial aging.

COL1A1							
Modifications	Amino acids	Recent (n=3)	% recovery	4-months (n=2)	% recovery	4-years (n=3)	% recovery
Oxidation	KPM	79	22,8%	98	28,2%	65	18,7%
Dioxidation	M	0	0%	0	0%	0	0%
Deamidation	NQR	5	3,4%	4	2,7%	3	2,0%
Phospho	CDKHRSTY	27	8,0%	20	5,9%	9	2,7%
Acetyl	K	0	0%	1	1,8%	0	0%
Carbamyl	K	0	0%	0	0%	0	0%
Carboxy	K	0	0%	0	0%	0	0%
Carboxymethyl	K	0	0%	0	0%	0	0%
Carboxyethyl	K	0	0%	0	0%	0	0%
Formyl	KST	9	5,3%	12	7,1%	9	5,3%
Methyl	DE	4	3,0%	3	2,2%	2	1,5%
Sulfo	STY	4	3,2%	4	3,2%	1	0,8%
COL1A2							
Modifications	Amino acids	Recent (n=3)	% recovery	4-months (n=2)	% recovery	4-years (n=3)	% recovery
Oxidation	KPM	40	13,8%	56	19,4%	40	13,8%
Dioxidation	M	0	0%	0	0%	0	0%
Deamidation	NQR	4	2,6%	3	1,9%	4	2,6%
Phospho	CDKHRSTY	9	2,9%	16	5,2%	17	5,6%
Acetyl	K	0	0%	0	0%	1	2,0%
Carbamyl	K	0	0%	1	2,0%	0	0%
Carboxy	K	0	0%	0	0%	0	0%
Carboxymethyl	K	0	0%	0	0%	1	2,0%
Carboxyethyl	K	0	0%	0	0%	0	0%
Formyl	KST	4	2,5%	6	3,8%	7	4,5%
Methyl	DE	2	1,9%	3	2,9%	3	2,9%
Sulfo	STY	1	0,8%	1	0,8%	0	0%

Table 3: Number and percentage of post-translational modifications on bovine and rat collagens COL1A1 and COL1A2 sequences at 4-months after pepsin and trypsin enzymatic treatment.

COL1A1					
Modifications	Amino acids	4-months bovine (n=2)	% recovery	4-months rat (n=2)	% recovery
Oxidation	KPM	89	25,5%	57	16,4%
Dioxidation	M	0	0%	0	0%
Deamidation	NQR	3	2%	2	1,4%
Phospho	CDKHRSTY	18	5,6%	4	1,2%
Acetyl	K	1	1,8%	0	0%
Carbamyl	K	0	0%	0	0%
Carboxy	K	0	0%	0	0%
Carboxymethyl	K	0	0%	0	0%
Carboxyethyl	K	0	0%	0	0%
Formyl	KST	8	5,0%	3	1,8%
Methyl	DE	3	2,1%	1	0,7%
Sulfo	STY	2	1,7%	0	0%
COL1A2					
Modifications	Amino acids	4-months bovine (n=2)	% recovery	4-months rat (n=2)	% recovery
Oxidation	KPM	39	13,2%	31	10,7%
Dioxidation	M	0	0%	0	0%
Deamidation	NQR	2	1,3%	2	1,3%
Phospho	CDKHRSTY	12	4,0%	5	1,6%
Acetyl	K	0	0%	0	0%
Carbamyl	K	0	0%	0	0%
Carboxy	K	0	0%	0	0%
Carboxymethyl	K	0	0%	0	0%
Carboxyethyl	K	0	0%	0	0%
Formyl	KST	4	2,7%	4	2,5%
Methyl	DE	3	2,8%	1	1,0%
Sulfo	STY	1	0,9%	1	0,8%

3.3.4. Comparison of exact positions of D-amino acids and posttranslational modifications

Both amino acid racemization and post-translational modifications can appear, simultaneously or not, during aging. Figures 8 and S14-18 regroup the exact position of D-amino acids totally racemized and the exact position of post-translational modifications.

Figure 8: 4-years bovine collagen COL1A1 sequence modifications: post-translational modifications (red), amino acids in their D-form totally racemized (blue), and both modifications on the same amino acid (green).

Conclusion

Artificial aging was successfully applied to standard collagen to mimic the natural aging in biological samples. Several collagen samples from different organisms (bovine and rat), were studied at different ages (recent, 4-months, 4-years). DCl/D_2O hydrolysis was privileged to HCl/H_2O to limit the natural racemization of amino acids during the protein hydrolysis. This new

chiral analysis method allowed us to determine the percentage of D-amino acids in collagen according to age. Results presented show a % D-amino acids – age correlation. Peptide analysis shows that the amount of peptides from aging degradation is higher than arising peptides from enzymatic digestions. The collagen protein sequence is increasingly degraded over time. The post-translational modifications study showed a decrease of hydrophilic groups and an increase of hydrophobic groups during aging. These sequence modifications may be a hypothesis to explain the evolution of the insolubility of collagen throughout life. The percentage of post-translational modifications was not significantly different between artificial and natural aging in both organisms. The combination of these results allowed us to determine the exact positions of D-amino acids and PTMs.

CRediT authorship contribution statement

Marine Morvan: Conceptualization, Methodology, Validation, Formal analysis, Inestigation, Writing – original draft, Writing – Review & Editing, Visualisation. **Ivan Mikšík**: Conceptualization, Methodology, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Declaration of competing interest

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

Acknowledgements

This work was supported by the Czech Science Foundation, grant number 20-03899S.

References

[1] R. L. Levine, E. R. Stadtman, Oxidative modification of proteins during aging, Experimental Gerontology 36 (2001) 1495-1502.
[2] G. Krishnamoorthy, R. Selvakumar, T. P. Sastry, A. B. Mandal, M. Doble, Effect of d-amino acids on collagen fibrillar assembly and stability: Experimental and modelling studies, Biochemical Engineering Journal 75 (2013) 92–100.
[3] J. W. Silzel, G. Ben-Nissan, J. Tang, M. Sharon, R. R. Julian, Influence of Asp Isomerization on Trypsin and Trypsin-like Proteolysis, Anal. Chem. 94 (2022) 15288–15296.
[4] A. L. Santos, A. B. Lindner, Protein Posttranslational Modifications: Roles in Aging and Age-Related Disease, Oxidative Medicine and Cellular Longevity 2017 (2017) 1–19.

[5] I. Mikšík, P. Sedláková, S. Pataridis, F. Bortolotti, R. Gottardo, Proteins and their modifications in a medieval mummy: Proteomic Analysis of Mummy, Protein Science 25 (2016) 2037–2044.
[6] R. J. W. Truscott, J. Mizdrak, M. G. Friedrich, M. Y. Hooi, B. Lyons, J. F. Jamie, M. J. Davies, P. A. Wilmarth, L. L. David, Is protein methylation in the human lens a result of non-enzymatic methylation by S-adenosylmethionine?, Experimental Eye Research 99 (2012) 48–54.
[7] S. Sh. Atavliyeva, P. V. Tarlykov, Paleoproteomics studies of ancient caprinae: a review 89 (2021) 4-14.
[8] N. Fujii, T. Kawaguchi, H. Sasaki, N. Fujii, Simultaneous Stereoinversion and Isomerization at the Asp-4 Residue in βB2-Crystallin from the Aged Human Eye Lenses, Biochemistry 50 (2011) 8628–8635.
[9] M. Y. S. Hooi, R. J. W. Truscott, Racemisation and human cataract. d-Ser, d-Asp/Asn and d-Thr are higher in the lifelong proteins of cataract lenses than in age-matched normal lenses, AGE 33 (2011) 131–141.
[10] M. Y. S. Hooi, M. J. Raftery, R. J. W. Truscott, Accelerated aging of Asp 58 in αA crystallin and human cataract formation, Experimental Eye Research 106 (2013) 34–39.
[11] B. Lyons, A. H. Kwan, J. Jamie, R. J. W. Truscott, Age-dependent modification of proteins: N-terminal racemization, FEBS J 280 (2013) 1980–1990.
[12] N. Fujii, T. Takata, N. Fujii, K. Aki, H. Sakaue, D-Amino acids in protein: The mirror of life as a molecular index of aging, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1866 (2018) 840–847.
[13] S. Ha, T. Kinouchi, N. Fujii, Age-related isomerization of Asp in human immunoglobulin G kappa chain, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1868 (2020) 140410.
[14] M. Morvan and I. Mikšík, Recent Advances in Chiral Analysis of Proteins and Peptides, Separations 8 (2021) 112.
[15] N. Fujii, T. Takata, I. Kim, K. Morishima, R. Inoue, K. Magami, T. Matsubara, M. Sugiyama, T. Koide, Asp isomerization increases aggregation of α-crystallin and decreases its chaperone activity in human lens of various ages, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1868 (2020) 140446.
[16] V. V. Dyakin, T. M. Wisniewski, A. Lajtha, Racemization in Post-Translational Modifications Relevance to Protein Aging, Aggregation and Neurodegeneration: Tip of the Iceberg, Symmetry 13 (2021) 455.
[17] D. Vicente-Zurdo, S. Rodríguez-Blázquez, E. Gómez-Mejía, N. Rosales-Conrado, M. E. León-González, Y. Madrid, Neuroprotective activity of selenium nanoparticles against the effect of amino acid enantiomers in Alzheimer's disease, Anal. Bioanal. Chem. 414 (2022) 7573–7584.
[18] A. A. Ageeva, A. B. Doktorov, N. E. Polyakov, T. V. Leshina, Chiral Linked Systems as a Model for Understanding D-Amino Acids Influence on the Structure and Properties of Amyloid Peptides, IJMS 23 (2022) 3060.
[19] P. M. Helfman, J. L. Bada, Aspartic acid racemization in tooth enamel from living humans, Proc. Natl. Acad. Sci. U.S.A. 72 (1975) 2891–2894.
[20] J. T. Powell, N. Vine, M. Crossman, On the accumulation of d-aspartate in elastin and other proteins of the ageing aorta, Atherosclerosis 97 (1992) 201–208.

[21] S. Ritz-Timme, I. Laumeier, M. J. Collins, Aspartic acid racemization: evidence for marked longevity of elastin in human skin, Br. J. Dermatol. 149 (2003) 951–959.
[22] J. M. Lee, L. Petrucelli, G. Fisher, S. Ramdath, J. Castillo, M. M. Di Fiore, A. D'Aniello, Evidence for D-Aspartyl-β-Amyloid Secretase Activity in Human Brain, J. Neuropathol. Exp. Neurol. 61 (2002) 125–131.
[23] S. Ritz, A. Turzynski, H. W. Schütz, A. Hollmann, G. Rochholz, Identification of osteocalcin as a permanent aging constituent of the bone matrix: basis for an accurate age at death determination, Forensic Science International 77 (1996) 13–26.
[24] G. Yasunaga, S. Inoue, T. Bando, T. Hakamada, Y. Fujise, Aspartic acid enantiomer quantification using ultraperformance liquid chromatography–tandem mass spectroscopy combined with deuterium-chloride hydrolysis to improve age estimation in Antarctic minke whale Balaenoptera bonaerensis, Marine Mammal Science (2022) 1-19.
[25] N. S. Mahlke, S. Renhart, D. Talaa, A. Reckert, S. Ritz-Timme, Molecular clocks in ancient proteins: Do they reflect the age at death even after millennia?, Int. J. Legal. Med. 135 (2021) 1225–1233.
[26] G. T. Matteussi, V. Jacometti, A. Franco, R. H. A. da Silva, Age estimation in humans through the analysis of aspartic acid racemization from teeth: A scoping review of methods, outcomes, and open research questions, Forensic Science International 331 (2022) 111154.
[27] H. Frank, W. Woiwode, G. Nicholson, E. Bayer, Determination of the rate of acidic catalyzed racemization of protein amino acids, Liebigs Ann. Chem. 1981(1981) 354-365.
[28] J. Csapá, Z. Csapó-Kiss, L. Wágner, T. Tálos, T. G. Martin, S. Folestad, A. Tivesten, S. Némethy, Hydrolysis of proteins performed at high temperatures and for short times with reduced racemization, in order to determine the enantiomers of d- and l-amino acids, Analytica Chimica Acta 339 (1997) 99–107.
[29] T. Miyamoto, H. Homma, Detection and quantification of d -amino acid residues in peptides and proteins using acid hydrolysis, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1866 (2018) 775–782.
[30] M. Danielsen, C. Nebel, T. K. Dalsgaard, Simultaneous Determination of L- and D-Amino Acids in Proteins: A Sensitive Method Using Hydrolysis in Deuterated Acid and Liquid Chromatography–Tandem Mass Spectrometry Analysis, Foods 9 (2020) 309.
[31] R. Pérez-Míguez, B. Bruyneel, M. Castro-Puyana, M. L. Marina, G. W. Somsen, E. Domínguez-Vega, Chiral Discrimination of DL-Amino Acids by Trapped Ion Mobility Spectrometry after Derivatization with (+)-1-(9-Fluorenyl)ethyl Chloroformate, Anal. Chem. 91 (2019) 3277–3285.
[32] T. Furuhashi, W. Weckwerth, Isomer analysis by mass spectrometry in clinical science, TrAC Trends in Analytical Chemistry 153 (2023) 116907.
[33] S. Ricard-Blum, The Collagen Family, Cold Spring Harbor Perspectives in Biology 3 (2011) a004978–a004978.
[34] S. Köster, H. M. Evans, J. Y. Wong, T. Pfohl, An In Situ Study of Collagen Self-Assembly Processes, Biomacromolecules 9 (2008) 199–207.
[35] N. K. Shah, B. Brodsky, A. Kirkpatrick, J. A. M. Ramshaw, Structural consequences ofD-amino acids in collagen triple-helical peptides, Biopolymers 49 (1999) 297–302.
[36] V. Punitha, S. Sundar Raman, R. Parthasarathi, V. Subramanian, J. Raghava Rao, Balachandran Unni Nair, and T. Ramasami, Molecular Dynamics Investigations on the Effect of D Amino Acid Substitution in a Triple-Helix Structure and the Stability of Collagen, J. Phys. Chem. B 113 (2009) 8983–8992.

[37] A. Vrána, L. Kazdová, The hereditary hypertriglyceridemic nonobese rat: an experimental model of human hypertriglyceridemia, Transplant. Proc. 22 (1990) 2579.

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 [70, 71]. 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 [72, 73]. This estimate can be performed using three approaches based on sexual dimorphism: osteoarchaeology, genomics, and proteomics [74]. 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.

In anthropology, sexual dimorphism is the morphological differences in bones between males and females, and is useful for estimating the sex of skeletal remains. Sex estimation methods, based on sexual dimorphism, are composed of three different techniques, *i.e.* visual, metric, and geometric-morphometric methods [75]. However, the full development of sexual dimorphism is present only in adulthood, which reduces the number of samples eligible for sex estimation. In addition, the complete adult skeleton is preferred for sex estimation. Indeed, if the skeleton is fragmented or from a sub-adult, the sex estimation is more difficult and less reliable than with a complete adult skeleton, and some individuals can be misclassified. In addition, genetic factors and sex hormones influence the dimorphism of the skeleton in all populations.

First, the visual method based on three indicators in the pubic bone (*os pubis*), *i.e.* the medial aspect of the ischiopubic ramus, the subpubic concavity, and the ventral arch [76], is a highly subjective method depending on the variability of sexual dimorphism within and between populations. Only estimates that have a reference set of multiple populations and are supported by software designed for sex estimation can provide usable results. With the visual method, the accuracy of the sex estimation varies from 96% to 100% [77] with an estimation error of < 10% being acceptable [78].
Second, the metric method consists in measuring the specific sizes and shapes of the cranial and post-cranial skeleton (Figure 5.1). Cranial analysis regroups 18 different metric measurements and associated landmarks (examples in Figure 5.2A) [79, 80], and 10 different angles [81]. On the other hand, the part of the post-cranial skeleton most studied is the pelvis with the pelvic inlet, the pubic arch, and the sacrum (*os sacrum*) concavity discrimination (Figure 5.2B). However, the entire pelvis is often damaged or missing in archaeological skeletons [82]. That is why, the analysis of the hip bone (*os coxae*), a part of the pelvis (Figure 5.1), is privileged with the measurement of different variables (Figure 5.2C) [83, 84]. The combination of these measurements allows estimating sex to achieve a precision of 97% and eliminates subjectivity from the visual method [85, 86].
Third, the discriminant function analysis (DFA) based on geometric-morphometric is the most commonly used for sex estimation. The principle of this method is to apply a discriminant score to males and females according to different discriminant observable functions. This individual discriminant score cannot be calculated if the unique discriminant function is missing. A greater alternative is the use of multiple observable functions to calculate the discriminant score. These scores greater than the cut-off point fixed to 0,5 seems to correspond to males, and less to females. The precision of the method varies from 80 to 95%. Indeed, some misclassified individuals have been reported [87, 88]. These misclassifications are due to the overlap of the discriminant score between males and females. This overlap area is called the zone of uncertainty.

To conclude, these osteoarchaeology methods have the advantage of being non-destructive. The average accuracy of the sex estimation ranges from 80% to 100% but is not absolute. In addition, the estimation of sex is possible only on adulthood and complete skeletons with a population reference, which reduces

considerably the number of skeleton candidates for the sex estimation. A complementary method, such as genomics with DNA analysis, which is more efficient and less restrictive, must be developed to estimate sex.

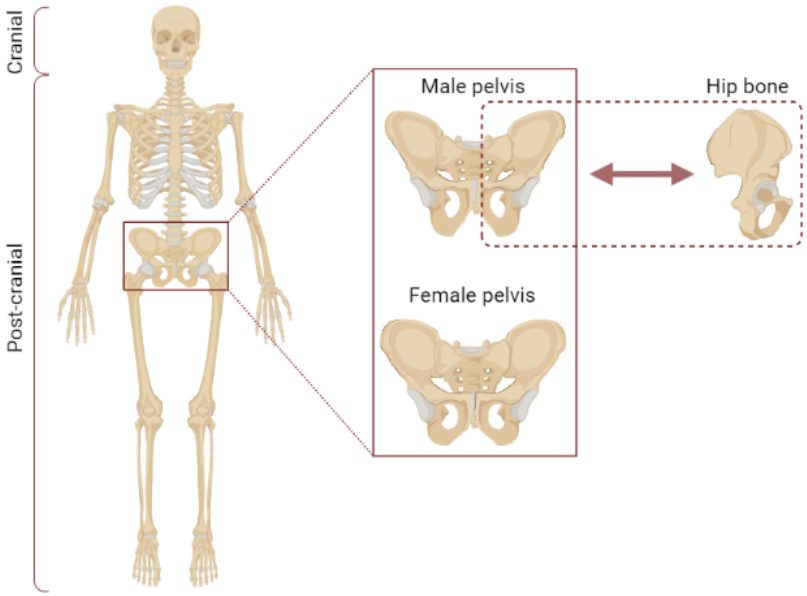


FIGURE 5.1: Representation of human skeleton, pelvis and hip bone
Created with BioRender.com

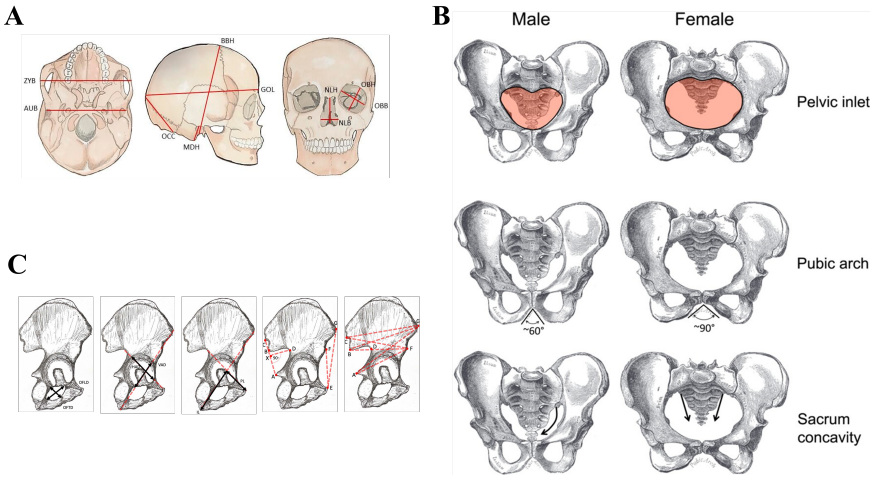


FIGURE 5.2: Metric measurements on the human (A) cranial, (B) pelvis and (C) hip bone
(A) reproduced with permission [80] and (C) reproduced with permission [83]

Molecular biology is an increasingly alternative method used to estimate the sex of skeletal remains when traditional anthropometric analyzes do not accurately estimate sex due to incomplete, fragmented, and/or refer to not adult skeletons [89–91]. This genomics method consists of analyzing DNA or in this case the ancient DNA (aDNA), to discriminate the X and Y chromosomes responsible for biological sex discrimination. For this determination, the combination of two techniques is required *i.e.* PCR (Figure A.1) and gel electrophoresis (Figure A.2). In DNA, the amelogenin gene is present in two sexually distinct forms (AMELX and AMELY genes) located on X and Y sex chromosomes, respectively (Figure 5.3). When only the AMELX gene is present, the sample is estimated as a female, and when both the AMELX and AMELY genes are present, the male sample is estimated. The homology between the nucleotide sequences of AMELX (528 base pairs) and AMELY (579 base pairs) is 88.9 % [92]. An AMELX gene isoform, namely AMELX-3, is composed of 576 base pairs, giving it greater homology with the AMELY gene (Figure 5.4). The PCR technique allows to amplify both forms of the amelogenin gene. The separation and identification of both genes can then be performed by gel electrophoresis on agarose gel. Indeed, according to their nucleotide sequence differences, their lengths are different (528 and 579 base pairs, respectively) [93, 94]. However, the main problems encountered for the aDNA analysis are the contamination of the samples and their strong long-term degradation. Most contamination occurs during the extraction of aDNA from biological materials. To limit this risk and decontaminate the sample, crucial steps deserve special attention, such as drilling or removing surface material, soaking or rinsing in bleach, UV irradiation to cross-link surface DNA, and extraction [95]. Furthermore, due to postmortem decomposition, the aDNA is very damaged. For example, in prehistoric skeletons, less than 0.5% of aDNA is preserved [96]. For less recent aDNA, a large amount of sample, *i.e.* 500 mg from bones or tissues, is necessary for PCR amplification. Although the accuracy of PCR in recent DNA approaches 100% [97], for aDNA its precision is 73. 8% when it was only 66.0% with the osteoarchaeology methods on the same skeletal remains [98].

To conclude, although genomics seems more efficient than the osteoarchaeological method when the sex estimation is ambiguous, only a small amount of aDNA is available for analysis, and it is often contaminated and strongly degraded. A new competitive method, such as proteomics to analyze the protein encoded by

the amelogenin gene, must be developed that is more efficient and consumes fewer samples.

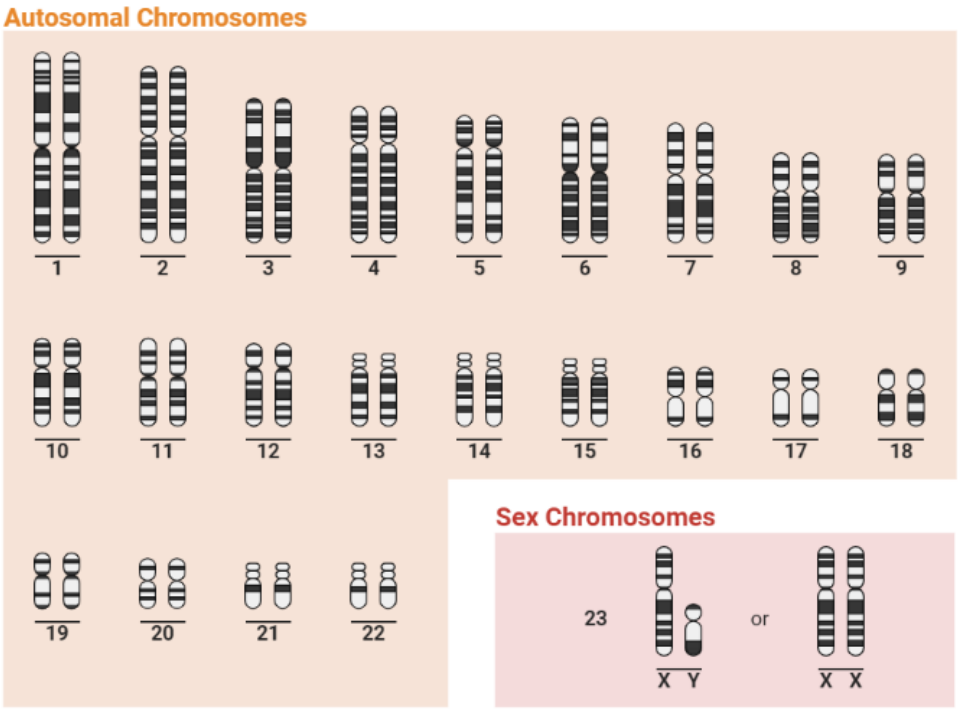


FIGURE 5.3: Representation of human chromosomes
Created with BioRender.com

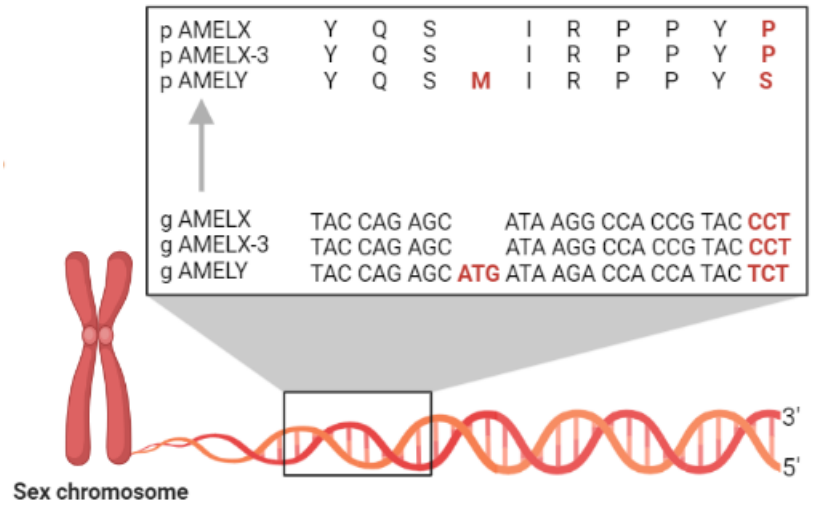


FIGURE 5.4: Examples of differences in amelogenin gene sequences encoded to different amelogenin protein sequences
CCDS: 14146.1 (g AMELX), 14144.1 (g AMELX-3), 14778.1 (g AMELY)
Created with BioRender.com

Proteomics analysis is designed to analyze both forms of amelogenin protein encoded by both forms of the amelogenin gene (Figure 5.4). 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 5.5). 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 5.5). *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 [99]. 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 [100–103]. 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 5.6) [99, 101, 104]. 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 [100, 105–111]. 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 [111]. 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 [112]. 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 [113].

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 5.5: 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)

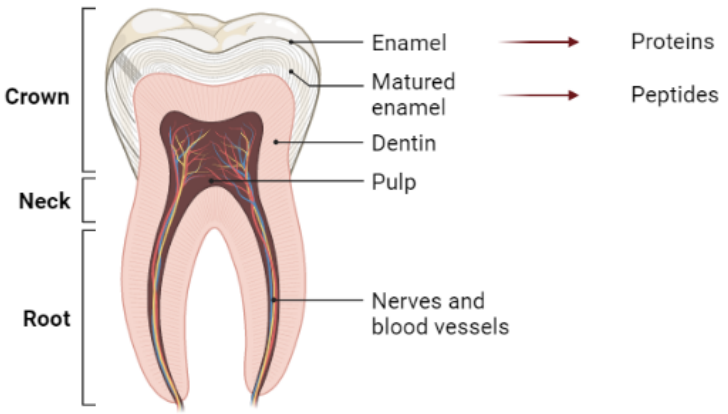


FIGURE 5.6: Composition of tooth

Created with BioRender.com

Peptide analysis of tooth enamel - a sex estimation tool for archeaological, anthropological, or forensic research

Ivan Mikšík, Marine Morvan, Jaroslav Brůžek

Journal of Separation Science, **2023**, 46, 2300183.

Review

Peptide analysis of tooth enamel – A sex estimation tool for archaeological, anthropological, or forensic research

Ivan Mikšík[1], Marine Morvan[1,2], Jaroslav Brůžek[3]

1 Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Pardubice, Czech Republic

2 Laboratory of Translational Metabolism, Institute of Physiology, CAS, Prague, Czech Republic

3 Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Prague, Czech Republic

* Correspondence: IMiksik@seznam.cz (I.M.)

Citation: Mikšík, I.; Morvan, M.; Brůžek, J. Peptide analysis of tooth enamel – A sex estimation tool for archaeological, anthropological, or forensic research, *Journal of Separation Science*, **2023**, *46*, 2300183. https://doi.org/10.1002/jssc.202300183

Abstract: Proteomics has become an attractive method to study human and animal material, biological profile, and origin as an alternative to DNA analysis. It is limited by DNA amplification in ancient samples and its contamination, high cost, and limited preservation of nuclear DNA. Currently, three approaches are available to estimate sex–osteology, genomics, or proteomics, but little is known about the relative reliability of these methods in applied settings. Proteomics provides a new, seemingly simple, and relatively non-expensive way of sex estimation without the risk of contamination. Proteins can be preserved in hard teeth tissue (enamel) for tens of thousands of years. It uses two sexually distinct forms of the protein amelogenin in tooth enamel detectable by liquid chromatography-mass spectrometry; the protein amelogenin Y isoform is present in enamel dental tissue only in males, while amelogenin isoform X can be found in both sexes. From the point of view of archaeological, anthropological, and forensic research and applications, the reduced destruction of the methods used is essential, as well as the minimum requirements for sample size.

Keywords: archeology, sex determination, tooth enamel

1. Introduction

The estimation of sex is fundamental to many archaeological, anthropological, and forensic research. This determination has three possible approaches: osteology, genomics, or proteomics [1]. The first two approaches are traditional methods, but the proteomic one is a relatively new method with the increasing attention of scientists. Recently the use of proteomic methods in paleontology (so-called “Paleoproteomics”) is rapidly growing, and it is expected that these methods can be helpful for many general applications and connect molecular biology, paleontology, archaeology, paleoecology, and history. Warinner et al. said about Paleoproteomics: ‘‘Growing from a handful of studies in the 1990s on individual highly abundant ancient proteins, paleoproteomics today is an expanding field with diverse applications ranging from the taxonomic identification of highly fragmented bones and shells and the phylogenetic resolution of extinct species to the exploration of past cuisines from dental calculus and pottery food crusts and the characterization of past diseases." [2]. From this point of view,

the most frequently analyzed protein is collagen type I, which is a stable and rigid protein, as was demonstrated in a well-known study in 2007 by determining collagen type I at 68-million-year-old bones of Tyrannosaurus rex [3, 4]. However, at the last years, the ability to detect and identification of proteins in ancient (but not only) tissue was significantly increased and allows to use of protein not only for identification in tissue but also for various problems such as interaction with food, interaction animal-human, environmental exchanges but also (from the view of this review) estimation of sex. Also, the significance of proteomic research has increased in archaeology [5]. It is important to note the differences in using the terms sex/gender. The terms sex and gender are often interchanged in conversation, documentation, and scientific literature, although they are not synonymous, and confusion in their usage is increasing. The biological concept of sex fundamentally differs from the social concept of gender, and the two terms are not interchangeable [6]. In this article, we will stick strictly to using the term sex only.

2. Osteological (morphological) estimation of sex

The existence of sexual dimorphism, that is, of differences in bone variables between males and females, is an inevitable prerequisite for any anthropological sex estimation method. According to the approach, the methods are divided into visual, which evaluates the development of characters according to categories, metric methods, which use the dimensions of bones and their size; and geometric-morphometric methods, which explore the interplay and differences between shape and sizes from a metric perspective [7]. Both genetic factors and the action of sex hormones influence the sexual dimorphism of the skeleton. Its final state is modified by external environmental factors (e.g., [8]). Before the development of secondary sexual characteristics, however, there was limited sexual dimorphism in skeletal features, rendering sexing methods unreliable it is not recommended to use them [9].

The full development of sexual dimorphism exists only in adulthood. In adults, the accuracy of the methods is different in every population and age group and depends on which bone is used. The dimensions of the skeleton show considerable variability. For these reasons, the methods must contain more variables that are not correlated with each other. In forensic anthropology and bioarchaeology, most methods for sex estimation rely on statistical models and tools generated through osteometric data collected from identified populations [10]. The most common method of sex estimation is discriminant function analysis (DFA). A discrimination score greater than the cut-off point between the sexes corresponding to a probability of 0.5 indicates a male and a score less than the cut-off point indicates a female. The accuracy of the methods is not absolute, and we always find a certain number of misclassified individuals who match the error. The skull and pelvis are still the preferred skeletal elements for sex estimation, with an accuracy that varies from 80% to 95% [11].

An overview of all possible anthropological methods for determining sex according to the skeleton is given in recent publications (e.g., [12–14]).

A common misunderstanding of DFA results is that the overall accuracy of sex classification can be applied to every individual in the sample. Every bone measurement and discriminant score shows an overlap between female and male distributions. The overlapping area represents the “zone of uncertainty”, where the skeletal variables of females and males are similar and cannot reliably be distinguished from one another [15]. This “zone of uncertainty” varies depending on the size difference between the population where the method was proposed and the population in which we want to use it. We call this the population specificity of morphological methods, and ignoring it causes a dramatic decrease in accuracy and an increase in misclassification [16]. Exceptions are methods that use pelvic

bone dimensions, which describe sexual dimorphism as a whole. Such methods have general validity and can be applied to all anatomically modern people with a high success rate and an error risk of less than 5% [17]. Unfortunately, in archaeological discoveries, the pelvis is often very damaged or completely missing [18]. For these reasons, the approach of primary and secondary sex diagnosis was proposed [19]. Primary sex diagnosis applies reliable methods in individuals with pelvic bone. In the sample obtained in this way, classification methods specific to the given population are proposed, which use extra-pelvic dimensions of the skeleton. These methods are applied to individuals with a missing pelvis [20]. Sex estimation in the "zone of uncertainty", that is, the overlapping region between sexes, should be avoided to reduce misclassification. Sex should be assigned only to those individuals with a posterior probability of being female or male higher than 0.95. Although such an approach limits the practical applicability of DFAs, because some portion of the individuals remains unclassified, it allows a high classification accuracy to be maintained at the individual level [15].

Another way to estimate the sex of a skeleton is to use visual methods and assess the degree of trait development using categories. These are highly subjective, and the variability of sexual dimorphism within and between populations is considerable. Only methods with a reference set from multiple populations and designed software for sex estimation can provide usable results [21–23].

3. Genetic estimation of sex

For a long time, hundreds of years, scientists searched for a system of sex-determining pathways. Aristotle (in 335 BCE) proposed that heat/cold could determine sex. This environmental theory was popular until about 1900 – when sex chromosomes were discovered. However, Aristotle was partly right when in some reptiles, the temperature of the nest can determine the sex of the embryo [24]. The discovery of chromosomal sex determination is credited to Nettie M. Stevens in 1905 with the finding that in most animals' sex is determined chromosomally, with males producing two types of gametes (carrying a Y or X chromosome) and thus being heterogametic, while females produce only one type and they are therefore homogametic [25]. Despite the great variety of sex determination mechanisms in some lineages, there is a surprising consistency. Amongst eutherian mammals, birds and some insects rely on heteromorphic sex chromosomes, which differ from each other in size, morphology, and gene content, to determine the sex of the individual [26].

Molecular biology and genetics can use aDNA at the individual level to identify the biological sex of a skeleton, make phenotypic inferences from an individual's genotype, and identify specific pathogens within an infected individual [27]. Molecular biology techniques are increasingly used to identify the sex of skeletal remains when traditional anthropometric analyses do not successfully identify the sex of remains that are incomplete, fragmented, and/or refer to immature individuals [28–30]. As Raff [27] reports, aDNA research cannot be done in regular molecular biology laboratories. While next-generation methods have made it much easier to distinguish contamination from endogenous DNA based on DNA damage patterns, preventing contamination during the extraction process is still extremely difficult (e.g., [31, 32]). Contamination is such a pervasive problem that the work requires specialised facilities and workers trained in aDNA protocols. Sample decontamination is essential before beginning extraction. They will include one or more of the following steps: drilling or removing surface material, soaking or rinsing in bleach, and UV irradiation to cross-link surface DNA and extraction [33]. Laboratories investigating aDNA must be positively pressurised, with HEPA-filtered air, polymerase chain reaction (PCR) enclosures, separation from post-PCR or modern DNA laboratories, and strict access protocols. They should be staffed by researchers trained in the specialised laboratories and methods necessary for ancient biomolecules [27].

From the analytical point of view, we can also mention sex estimation by capillary gel electrophoresis with amelogenin locus as a marker from 1998 [34]. The analysis of ten male and ten female samples and double peaks verified the male sample, while one peak verified the female sample. As a result, two peaks were obtained—the first peak was for the X and the second for the Y locus.

The most common method of genetic sex estimation relies on differences in the amelogenin gene (e.g., [35,36]). Using cloning techniques, the amelogenin gene was localized to the sex chromosomes. With the advent of PCR, researchers developed robust and relatively simple amplification techniques for determining the presence of X and Y chromosome versions of this gene. This technique exploited the fact that insertion/deletion polymorphisms between the X and Y chromosomes lead to differently sized amplicons, which can easily be visualised by size separation using gel electrophoresis. These methods were used in the first commercially available kits [37]. However, the amelogenin-based method can be problematic due to allelic dropout. It means that if only the copy present on the X chromosome is retained and not the copy on the Y chromosome, the individual will be falsely identified as female rather than male. Another problem is modern contamination from a male source, which could mislabel an ancient individual as male [27]. However, more recent profiling kits include assays for a separate insertion/deletion polymorphism and a short tandem repeat locus found only on the Y chromosome. Additionally, analysis of other short tandem repeats, other insertions, deletions, or single nucleotide polymorphisms on the X or Y chromosome can also be used to determine sex [37–40].

To overcome the before mentioned shortcomings, high-throughput shotgun sequencing has been proposed, which represents a more accurate method of sex assignment, as it simultaneously avoids the problem of allelic dropout by testing many more discriminating loci and allows the detection of contaminating fragments by assessing DNA damage patterns [41]. Skoglund et al. [42] evaluated the feasibility of this approach for applications in population-scale aDNA investigations. They found that approximately 100 000 reads are required for accurate sex estimation so that indexed genomes of up to 13 individuals could be pooled and simultaneously sequenced at low coverage at an estimated cost of < $300 per sample. While considerably more expensive than a simple amelogenin size-based analysis, it does have the advantage of being more accurate. It is likely that as shotgun sequencing costs continue to decrease, this approach will become much more routinely applied to archaeological populations [27]. However, this degree of preservation may be problematic for many archaeological remains, as Mittnik et al. [43] noted. To reduce the required number of mapped human sequences, Mittnik and colleagues proposed an alternative method of sex estimation using high-throughput shotgun sequenced DNA. This method relies on the proportion of reads mapped to the human X chromosome compared to the proportion of reads mapped to each autosomal chromosome. By down-sampling reads from the same high-quality ancient DNA data sets used in Skoglund et al. [41], the proposed method could give confident assignments with as few as 1000 human genomes reads [1].

However, both Skoglund's and Mittnik's approaches [42, 43] are limited and admit some risk of error when the confidence interval is within certain parameters [44]. To determine the genetic sex is compared the ratio for the number of Y-chromosomal 1240k positions with available data relative to the 1240k position on the X-chromosome. Individuals with a ratio greater than 0.35 were considered genetic males and individuals with a ratio less than 0.03 were considered genetic females [45]. The current rapid development of genomics will certainly enable even more sophisticated methods of genetic sex estimation. Current research is yielding very encouraging results in the field of archeology that were previously unavailable (i.e., [46–48]).

However, it cannot be neglected that the and extraction process is invasive and generally involves the destruction of a small (<0.5 g) amount of bone or tissue. The material for extraction should not be used for osteological research, such as non-pathological ribs. Recently, a petrous bone from the inner ear

has been an excellent source of well-preserved, high-quality aDNA and should therefore represent an optimal extraction target [49, 50]. However, even petrous bone is an essential element in anthropology that should not be sacrificed to obtain high-quality aDNA [51]. In addition, petrous also has several applications for scientific analysis beyond sex estimation from ancient DNA. There is a risk of potential bias in using the pars petrosa for ancient DNA analysis [52]. Ethical guidelines for petrous bone sampling need to be expanded further, as the demands for destructive sampling will only increase as the number of researchers and laboratories conducting aDNA research increases.

However, the genomic method looks like a standard and required method for the estimation of sex; there is also a developed proteomic method (see next section). One could ask the question: why? One of the answers is that genetic material is not so resistant to degradation, but what about the precision of sex estimation? An interesting comparison of the three methods (osteological, genomic, and proteomic) was made on 55 individuals between 2440 and 100 cal BP [1]. Agreement between all methods was excellent when DNA shotgun sequencing was about 100 000 total sequences. However, more than half samples were below this threshold, and the conflict of sex estimation increased. On the opposite proteomic signal was not decreasing so significantly. It was concluded that proteomic data could complement osteological and genomic results/data.

4. Proteomics estimation of sex

Nowadays, proteomic methods enable us to determine minor differences in protein composition (qualitative and quantitative) enabled by many influences such as diet, aging, breeding, and sex [53]. Of course, the gender approach is a popular modern society theme, as visualized by the European Commission: “Gendered Innovations. How Gender Analysis Contribute to Research”, and thirty years of research have revealed that sex and gender biases are socially harmful and costly [54].

For archaeological, anthropological, or forensic research, it is necessary to consider tissues resistant to degradation caused by long-term effects, aggressive decomposition phenomena in the soil, and external physical and chemical damage. For these purposes, the ideal tissue is teeth enamel. Enamel is one of the most calcified tissues in mammalian organisms and can protect teeth for tens of thousands of years [55]. A nanocomposite bioceramic shields teeth against multiple chemical and physical (mechanical) efforts to disturb them. Protein amelogenins regulate crystallite formation during enamel development; however, they are specifically degraded during teeth maturation [56]. Amelogenin genes, in humans, are located on X and Y chromosomes (AMELX and AMELY, respectively). Proteins encoded by these genes have a different amino acid sequence (see Figure 1). The result of the presence of these two genes on the X and Y chromosomes is the sex-dependent presence of different proteins: AMELX for females and AMELX and AMELY for males [57]. During enamel maturation, proteins are degraded (proteolytic procedure), so mature tooth enamel is rich in various peptide fragments of constituted proteins [55]. We have to remind that amelogenin is a relatively small molecule highly concentrated in the extracellular matrix [58]. Because amelogenin is a major enamel protein, these peptides mainly originated from amelogenin. For this reason, free peptides from sex-dependent amelogenins (AMELX and AMELY) are a good choice for estimating sex. If we are looking for a universal sex determination pathway for all mammals, we have to pay attention: AMELX gene on the X chromosome X and AMELY on the Y chromosome are presented not in all mammalian families; it is presented in Hominidae, Suidae, and Bovidae, but rodent species have only one AMELX [58]. We must also mention that the tissue used (teeth enamel) is difficult to cross-contaminate.

Figure 1. Structure of Amelogenins and their comparison. Data are from UniProtKB reviewed (Swiss-Prot): Q99217 Amelogenin, X isoform, Q99217-3 Amelogenin, X isoform, Rare isoform 3; Q99218 Amelogenin, Y isoform. Differences in amino acid sequence between X and Y amelogenins are indicated by white characters highlighted in black.

In 1991 Fincham et al. [59] remark that a diagnosis of differences in human enamel proteins can permit the distinction of specimens according to the sex of the individual. Porto et al. [60, 61] discover/describe enamel proteins looking for sex-depending peptides. They used whole-crown etching, enzymatic (trypsin) treatment, and analysis by MALDI-TOF/TOF. They found the amino-terminal amelogenin peptides in the ancient sample (Mummy—ad 800–1100). They also found a peptide (WYQSIRPPYP) specific for the amelogenin X-isoform but no peptide specific for Y isoform. However, the authors did not specify the sex of the samples [60]. Although the MALDI-TOF/TOF method is suitable for protein and peptide analysis, nanoLC-MS/MS is a method of choice due to its higher sensitivity for peptide detection and identification. [62].

Another attempt to analyze human teeth enamel was made by Castiblanco et al. [63] in 2015. They investigated the protein composition of healthy contemporary pulverized teeth without the digestion of an enzyme (trypsin). After extraction and cleaning on C18 StageTips, they analyzed peptides/proteins on nLC/MS-MS using QExactive MS (Thermo Fisher Scientific). Seven proteins were identified when four were specific for enamel (AMELX/AMELY, enamelin, and ameloblastin). Although eight teeth were used, the sex was not mentioned, and the method was not used for its detection.

Stewart et al. [64] first published specific etching of single teeth to identify sex-sensitive amelogenins from enamel in 2016. This procedure (and with some slight modification) is now commonly used for peptide extraction from enamel (for the scheme, see Figure 2). The method consists of two parts [65]: etching and extraction. First teeth are mechanically treated (by dental burr) to clean them from all macroscopic impurities. Cleaning continued by water washing. Next, procedures were continued in the cap of a separate microcentrifuge tube (leaving a convex meniscus protruding above the lip when the tooth was lowering into the cap). The tooth's crown was washed with 3% H_2O_2 for 5 min, twice washed with H_2O, and etched with 10% (v/v) HCl for 2 min. All these solutions were discarded. A second 2-min etch was used for analysis. The next stage was extraction. C18 resin-loaded ZipTip (Millipore, MA, USA) was used for extraction. ZipTip was previously conditioned three times with 100% ACN and then three times with 0.1% (vol/vol) formic acid (each draw discarded). The peptides were bound to the ZipTip (by up and down pipetting by 10 times. The ZipTip was washed six times with 0.1% (vol/vol) formic acid (each wash discarded). Bounded peptides were eluted by 4-µl of 60% ACN/0.1% formic acid. This fraction was lyophilized and dissolved in 12 µl of 2% formic acid in water and analyzed by reversed-phase nanoLC/MS (with gradient similarly generally used for peptide separation, when used solvents were water and ACN with 0.1% TFA). The mass spectrometer was a hybrid linear ion trap orbitrap

(Orbitrap XL, Thermo Scientific). They selected peptides AMELY-(58-64; SM(ox)IRPPY ([M + 2H]2+, m/z 440.2233) peptide and AMELX-(44-52; SIRPPYPSY ([M + 2H]2+, m/z 540.2796) (Figure 3).

Authors [64] also compared the influence of trypsin digestion on protein/peptide identification. However, trypsin increased the variety of peptides, but no significant differences were observed in "sex" proteins (enamel specific proteins) when trypsin-treated and untreated samples were compared.

The same approach was used for sex estimation of the Late Antique (probably 4th–6th century) 'Lovers of Modena' [66]. The authors extracted peptides (after etching of teeth) by HyperSep SpinTips (Thermo Scientific) with C18 phase, and the analytical system was UHPLC-HRMS (Q Exactive MS from Thermo Scientific). Besides SM(ox)IRPPY (m/z 440.2233) peptide, they used at least two other peptides—SMIRPPY (m/z 432.2258) and M(ox)IRPPY (m/z 396.7073). All these peptides agreed in the distinction of specimens according to the sex, i.e. in the determination of AMELY. Surprisingly the 'Lovers of Modena' were males. The same method was successfully used to research sex-related morbidity and mortality for 30 nonadults from the Early Medieval Italian site from the 7th century AD [67]. The authors concluded that this method is used for archaeological (and forensic) research.

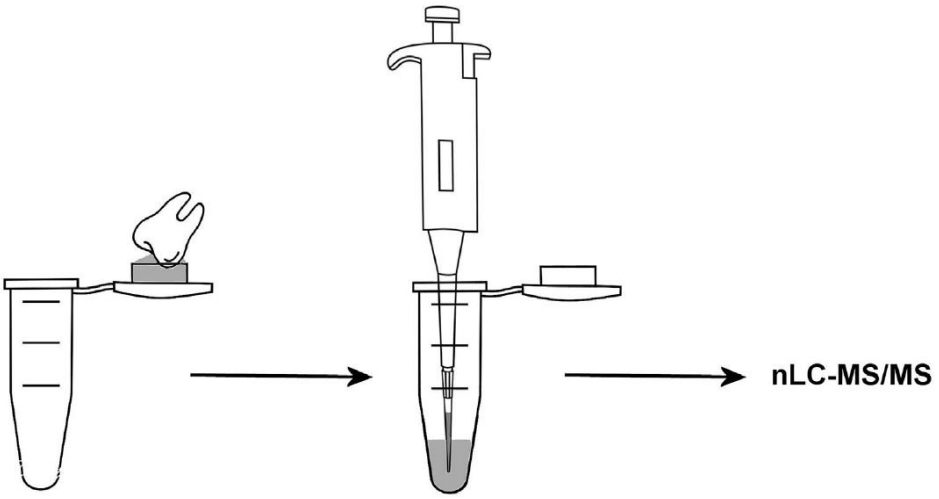


Figure 2. Workflow of the sample preparation (etching method) for amelogenin analysis.

This method was also successfully used for sex estimation from deciduous and permanent teeth from non-adults (incl. perinatal subjects) from archaeological sites in England (1st–2nd centuries AD, and 18th–19th centuries) [68].

Again the same method [66, 67] was used for the sex estimation of the horsemen of the Early Middle age (7th century AD) [69]. Determination of sex by proteomic (enamel) method was in agreement with osteological and archaeological determination when in many cases, only proteomic analysis was the only acceptable method for sex estimation. Authors also used confident identification of AMELY by triple-peptide approach (m/z 396.7073 for M(ox)IRPPY, m/z 432.2258 for SMIRPY, and m/z 440.2233 for SM(ox)IRPPY).

The above-described etching techniques were studied for optimization using three treatment procedures by 1.5M HCl: 3-sequential 4-min, 4 min only, and 10 min incubations [55]. Samples were desalted on SDB-RPS StageTips before analysis on nLC-MS/MS (using Q Exactive Plus Mass spectrometer) system. From the point of view of sex estimation, AMELX, and AMELY peptides authors recommend 10-min etching. The method was validated on a set of 23 archaeological teeth in

comparison to two different methods of sex estimation: morphological and archaeological (based burial rite) methods. All three methods agreed except in one case (when the archaeological method was opposite to morphological and protein methods) [55]. Besides morphological methods, sex determination by this method was also described for sex estimation of prepubertal individuals from Roman Italy (1st−4th c. CE) and Late Roman Gaul (4th−5th c. CE) era [70].

Osteological, proteomic, and isotopic analyses were used to evaluate a 9000-year-old human burial from the Andean highland site. These analyses indicate that this early hunter was a young adult female. This challenge the man-the-hunter hypothesis [71]. For proteomic analysis, a small piece of enamel (20 mg) was cut from teeth, powdered, and demineralized by 1.2 M HCl. After reduction and alkylation, samples were treated with trypsin. The next step was extraction on SepPak C18. Peptides were analyzed by nLC-MS/MS (Thermo ScientificQ-Extractive Plus Orbitrap mass spectrometer). Multiple peptide detection was used for AMELX (and AMELY) identification [71].

Two fast methods (1-3 min) were also developed for sex estimation without the separation step [72]. Both methods are based on the FIA (flow injection analysis) using high-resolution mass spectrometry (Q Exactive Orbitrap MS from Thermo Scientific) or tandem MS (Xevo TQ-S from Waters, Milford). Analyzed peptides were again SM(ox)IRPPY and SIRPPYPSY when specific transition ions were selected. Sample preparation is the same as above. The advantages of these methods are rapid analysis (three, respectively 1 min per sample) and the possibility to use low-resolution mass spectrometers (MS/MS), i.e. relatively low-cost instruments, for sex estimation.

There are also described other amelogenin peptides as diagnostic peptides using a nanoLC system coupled to a Q Exactive orbitrap mass spectrometer [73, 74]. The sample preparation was, in principle, the same as in the method by Stewart et al. [65], but the AMELY peptide was SM(ox)IRPPYS (m/z 486.7393), and AMELX peptide was YEVLTPLKWYQSIRPPYP (m/z 750.7368) (Figure 4). This method was successfully applied for the estimation of the sex of a child (5–6-year-old boy) murdered in the Early Bronze Age from Schleinbach, Austria (c. 1950–1850 BCE) [73] as well as for successful sex classification of 70 (from 75) children under 12 years at death buried at the Early Bronze Age cemeteries in Franzhausen I, Austria (c. 2050−1680 BCE). The coincidence between archaeological and peptide-based results was very high (62 of 63 individuals, 98.4%) when one was of the female sex, based on body position and orientation [74].

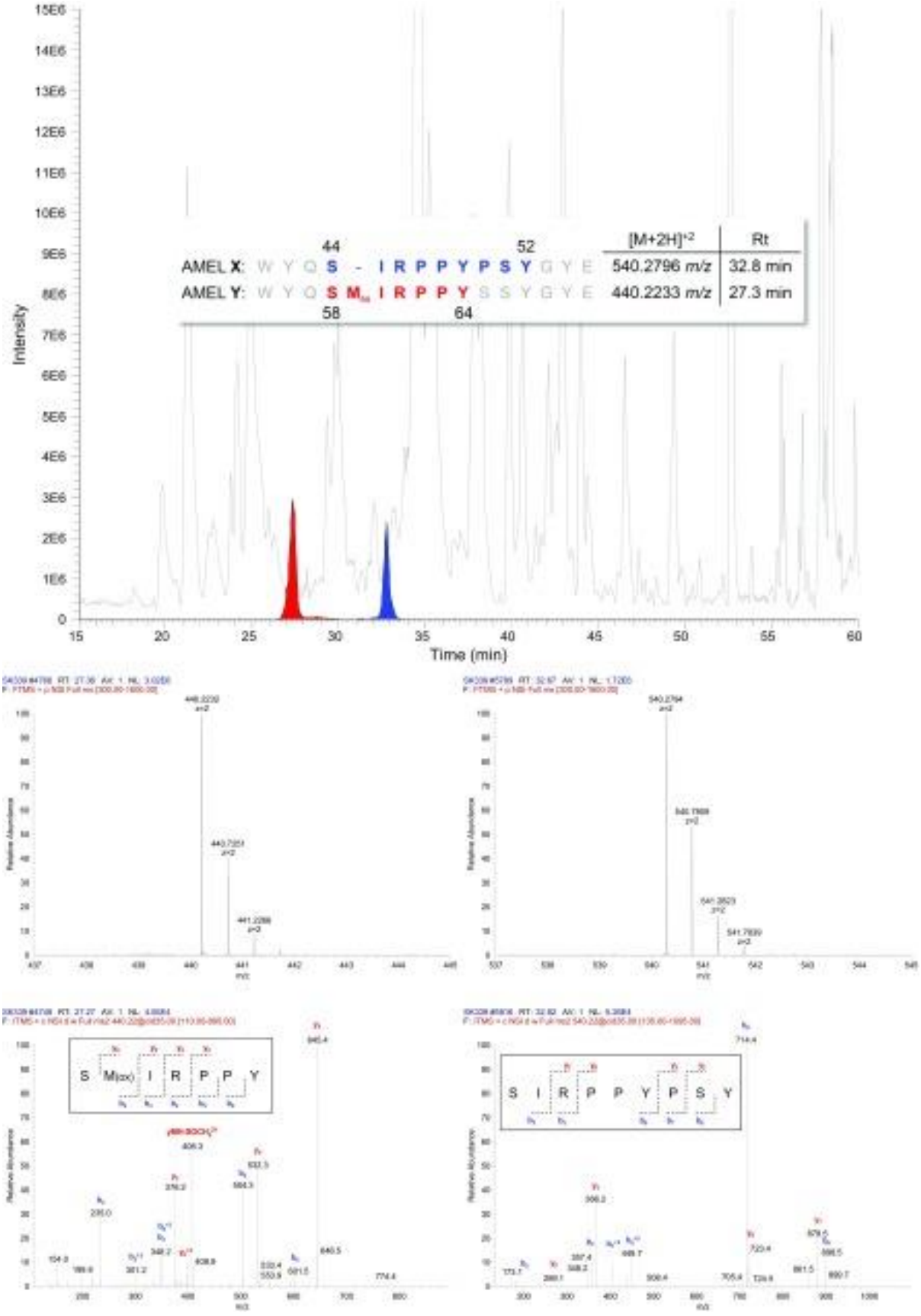


Figure 3. A base peak chromatogram (300−1600 m/z) with two marked peptides of amelogenin: AMELY-[58–64] and AMELX-[44–52]. The reconstructed ion chromatograms (to 4 ppm) for each are shown in

red and blue; full-scan MS and corresponding MS/MS are shown below. Reprinted [65] with permission. Copyright 2017 National Academy of Sciences.

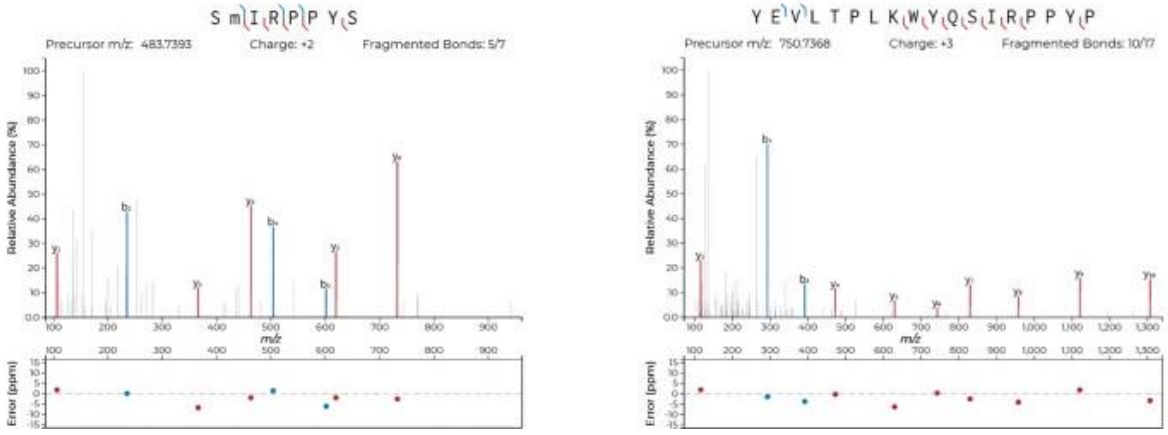


Figure 4. Left: MS/MS fragment spectrum of peptide SM(ox)IRPPYS (m/z 486.7393) (AMELY) and corresponding mass errors in ppm; Right: MS/MS fragment spectrum of peptide YEVLTPLKWYQSIRPPYP (m/z 750.7368).(AMELX) and corresponding mass errors in ppm. Reprinted [73] with permission. Copyright 2020 Springer Nature.

Another approach for analysis of archaeological samples was described for Iron Age individuals (ca. 2000–1000 years B.P.) from the tropical environment (northwest Thailand) [75]. It was possible to identify 212 proteins. Analyses were done at nLC-MS/MS system (OrbitrapVelos, Thermo Electron, Bremen, Germany) using the multiple reaction monitoring methods were used for the identification of two AMELX peptides (TALVLTPLK and WYQSIRPPYPSY(G)), and one AMELY peptide (IALVLTPLK). Regarding sample preparation, the tooth enamel/dentin was crushed, treated with 0.5MHCl, reduced, alkylated, and cleaved by trypsin.

Another protein/peptide extraction was used by Froment et al. [76] to analyze 5000-year-old human teeth. In principle, they powdered whole teeth demineralized in EDTA, and proteins were denatured, lysed, washed, and alkylated. Finally, proteins were treated with trypsin. For analysis of peptides, nanoLC coupled to Orbitrap Fusion MS (Thermo Scientific) was used when using the targeted MS approach and parallel reaction monitoring (PRM). They demonstrated that the PRM method maximizes the sensitivity and reproducibility of “sex” peptides. The selected “sex” peptides were TALVLTPLK (m/z 478.3130) for AMELX and IALVLTPLK (m/z 474.3325) for AMELY (Figure 5).

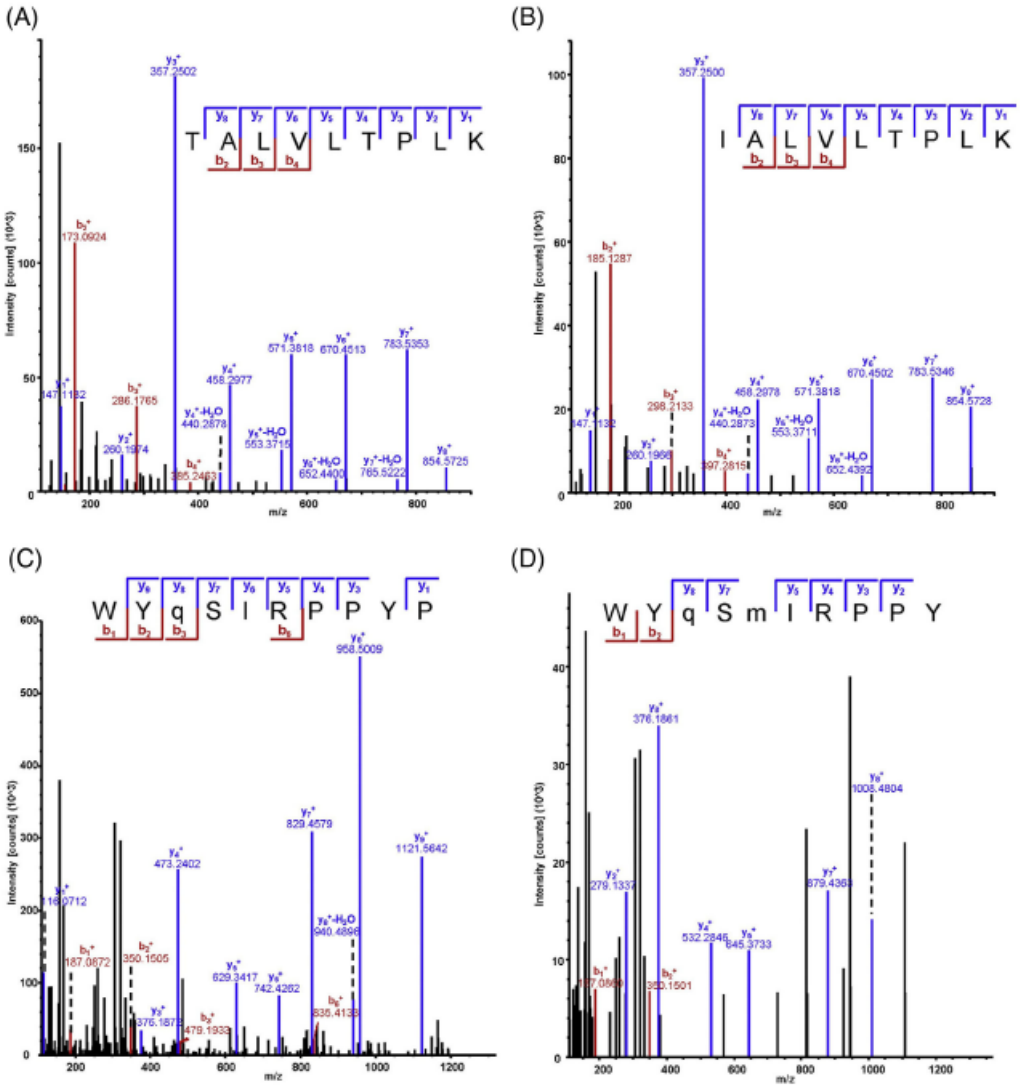


Figure 5. MS/MS spectra of specific peptides: (A) AMELX peptide, TALVLTPLK (precursor m/z 478.3130). (B) AMELY peptide, IALVLTPLK (precursor m/z 484.3325). (C) AMELX peptide, WYqSIRPPYP (precursor m/z 654.3259), and (D) AMELY peptide, WYqSmIRPPY (precursor m/z 679.3165). The series of y- and b-ions are highlighted in blue and red, respectively. q: deamidated glutamine residue; m: oxidized methionine residue. Reprinted [76] with permission. Copyright 2019 Elsevier B.V.

A similar method for sample preparation, that is, demineralized (by 1.2 M HCl) milled teeth were alkylated and treated by trypsin, was used to analyze teeth aged up to 7300 years [77]. ZipTip C18 tips cleaned samples/peptides. These peptides were analyzed by nLC-MS/MS using Q Exactive Plus Orbitrap MS (Thermo Scientific). The authors used bioinformatic methods for the identification of AMELX and AMELY proteins. To elute false negative samples, that is, samples with low male signal, authors used a probability curve of female sex as a function of the logarithm of AMELX using logistic regression.

False female assignments were discussed in this context [77–79]. It is some probability that the method using AMELX and AMELY could produce inaccurate results due to the presence of low frequency of AMELY deletion variants in some populations [78]. However, after analysis of many genomic projects, it was concluded that the probability of false sex estimation is low, and AMELY deletion should not affect the routine estimation of the biomolecular sex [79].

The oldest dental proteome was probably studied in the Early and Middle Pleistocene hominin (Homo antecessor) and Homo erectus tooth [80]. For protein/peptide extraction, authors used three methods: 1) demineralization by HCl without alkylation and enzymatic digestion, 2) pellet after demineralization was reduced, alkylated, and digested by LysC and trypsin, and 3) demineralization by TFA without alkylation and enzymatic digestion. The first and third extraction gave more extensive peptide recovery than the second one. Peptides were analyzed by nLC-MS/MS using Q-Exactive HF or HF-X mass spectrometer (Thermo Fisher Scientific). It was described that an average peptide length decreased with the age of the enamel sample. Enamel-specific proteins were identified as enamelin, ameloblastin, MMP20, and amelogenins (both AMELX and AMELY). At the teeth of Homo antecessor, AMELY-specific peptide sequences (such as SM(ox)IRPPY) were found, and so it was concluded that he was male [80]. We must mention that Homo antecessor is an extinct archaic human species recorded in Spain that lived between 1.2 and 0.8 million years ago during the early Pleistocene.

In the end, we have to mention that acid etching of teeth is, mainly in comparison to genetic methods, a limited destructive method, but still is (limited) destructive and not quantitative (the signal cannot be normalized).

5. Concluding remarks

Sexual diagnosis is often crucial to archaeological, anthropological, and forensic research. Proteomic methods are relatively young, rapidly growing, and have applications in many scientific areas. Nowadays, there are used three methods of sex estimation: osteology, genetics, and proteomics. In the sex estimation method, use protein amelogenin in X- and Y-form (AMELX and AMELY). It was established as a method that overcomes osteological and genomic methods as a more precious, sensitive, and relatively simple and rapid method. The crucial advantage is that this method uses teeth enamel. The enamel is one of the most calcified tissues in the mammalian organism. For this reason, it is relatively resistant to degradation caused by long-term effects, aggressive decomposition phenomena in the soil, and external physical and chemical damage. It was proven that proteins (peptides) in the enamel could be protected for tens or hundreds of thousands of years.

So it can be concluded that the proteomic method for sex estimation using isoforms of amelogenin be successfully used in archaeological, anthropological, and forensic research and overcome other previously used methods.

Conflict of interest statement

The authors declare no conflict of interest.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. Buonasera T, Eerkens J, de Flamingh A, Engbring L, Yip J, Li H, et al. A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation. Sci Rep. 2020;10(1):11897.

2. Warinner C, Korzow Richter K, Collins MJ. Paleoproteomics Paleoproteomics. Chem Rev. 2022;122(16):13401–46.

3. Schweitzer MH, Suo Z, Avci R, Asara JM, Allen MA, Arce FT, et al. Analyses of soft tissue from Tyrannosaurus rex suggest the presence of protein. Sci. 2007;316(5822):277–80.

4. Asara JM, Schweitzer MH, Freimark LM, Phillips M, Cantley LC. Protein sequences from Mastodon and Tyrannosaurus rex revealed by mass spectrometry. Sci. 2007;316(5822):280–5.

5. Hendy J. Ancient protein analysis in archaeology. Sci Adv. 2021;7(3):eabb9314.

6. Garofalo EM, Garvin HM. The confusion between biological sex and gender and potential implications of misinterpretations. In: Klales AR, editor. Sex estimation of the human skeleton. Cambridge, MA: Academic Press; 2020. pp. 35–52.

7. Klales A. Sex estimation of the human skeleton. Cambridge, MA: Academic Press; 2020.

8. Dunsworth HM. Expanding the evolutionary explanations for sex differences in the human skeleton. Evolut Anthropol. 2020;29(3):108–16.

9. Scheuer L, Black S. Skeletal development and ageing. In: Scheuer L, Black S, editors. Developmental juvenile osteology. London: Academic Press; 2000. pp. 4–17.

10. Obertová Z, Stewart A, Cattaneo C. Statistics and probability in forensic anthropology. 2020. 1–407.

11. Brůžek J, Murail P. Methodology and reliability of sex determination from the skeleton. In: Schmitt AEC, Pinheiro J, editors. Forensic anthropology and medicine: complementary sciences from recovery to cause of death. Totowa: Humana Press; 2006. pp. 225–42.

12. Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh RK. A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. Forensic Sci Int. 2016;261:165.e1–8.

13. Klales AR. Current State of Sex Estimation in Forensic Anthropology. Forensic Anthropol. 2021;4:1.

14. Kotěrová A, Rmoutilová R, Brůžek J. Current trends in methods for estimating age and sex from adult human skeleton. Anthropologie. 2022;60(2):225–52.

15. Galeta P, Brůžek J. Sex estimation using continuous variables: problems and principles of sex classification in the zone of uncertainty. 2020. p. 155–82.

16. Kotěrová A, Velemínská J, Dupej J, Brzobohatá H, Pilný A, Brůžek J. Disregarding population specificity: its influence on the sex assessment methods from the tibia. Int J Legal Med. 2017;131(1):251–61.

17. Brůžek J, Santos F, Dutailly B, Murail P, Cunha E. Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. Am J Phys Anthropol. 2017;164(2):440–9.

18. Waldron T. The relative survival of the human skeleton: implications for palaeopathology. Death Decay Reconstruct. 1987:55–64.

19. Murail P, Bruzek J, Braga J. A new approach to sexual diagnosis in past populations. Practical adjustments from Van Vark's procedure. Int J Osteoarchaeol. 1999;9(1):39–53.

20. Santos F. rdss: an R package to facilitate the use of Murail et al.'s (1999) approach of sex estimation in past populations. Int J Osteoarchaeol. 2021;31(3):382–92.

21. Klales A, Garvin H, Gocha TP, Lesciotto KM, Walls M. Examining the reliability of morphological traits for sex estimation implications for the Walker (2008) Klales et al. (2012) methods. Forensic Anthropol. 2020;3(3):139–51

22. Klales AR. MorphoPASSE: morphological pelvis and skull sex estimation program. In: Klales AR, editor. Sex estimation of the human skeleton. Cambridge, MA: Academic Press; 2020. pp. 271–8.

23. Santos F, Guyomarc'h P, Rmoutilova R, Bruzek J. A method of sexing the human os coxae based on logistic regressions and Bruzek's nonmetric traits. Am J Phys Anthropol. 2019;169(3):435–47.

24. Hake L,O'Connor C.Genetic mechanisms of sex determination. Nat Educ. 2008;1(1):435.

25. Ramos L, Antunes A. Decoding sex: Elucidating sex determination and how high-quality genome assemblies are untangling the evolutionary dynamics of sex chromosomes. Genomics. 2022;114(2):110277.

26. Mank JE. Sex-specific morphs: the genetics and evolution of intra-sexual variation. Nat Rev Genet. 2023;24(1):44–52.

27. Raff JA. Ancient DNA and bioarchaeology. A companion to anthropological genetics. 2019. pp. 44–52.

28. Stone AC,Milner GR, Pääbo S, Stoneking M. Sex determination of ancient human skeletons using DNA. Am J Phys Anthropol. 1996;99(2):231–8.

29. Afonso C, Nociarova D, Santos C, Martinez-Labarga C, Mestres I, Duran M, et al. Sex selection in late Iberian infant burials: Integrating evidence from morphological and genetic data. Am J Hum Biol. 2019;31(1):231.

30. Poma A, Cesare P, BonfigliA, Volpe AR, Colafarina S, Vecchiotti G, et al. A qPCR-duplex assay for sex determination in ancient DNA. PLoS One. 2022;17(6):e0269913.

31. Willerslev E, Cooper A. Ancient DNA. Proc Biol Sci. 2005;272(1558):3–16.

32. Yang DY, Watt K. Contamination controls when preparing archaeological remains for ancient DNA analysis. J Archaeol Sci. 2005;32(3):331–6.

33. Kaestle FA, Horsburgh KA. Ancient DNA in anthropology: Methods, applications, and ethics. Am J Phys Anthropol. 2002;119(S35):331–6.

34. Komuro T, Nakamura M, Tsutsumi H, Mukoyama R. Gender determination from dental pulp by using capillary gel electrophoresis of amelogenin locus. J Forensic Odonto Stomatol. 1998;16(2):92–130.

35. Faerman M, Filon D, Kahila G, Greenblatt CL, Smith P, Oppenheim A. Sex identification of archaeological human remains based on amplification of the X and Y amelogenin alleles. Gene. 1995;167(1-2):327–32.

36. Gibbon V, Paximadis M, Štrkalj G, Ruff P, Penny C. Novel methods of molecular sex identification from skeletal tissue using the amelogenin gene. Forensic Sci Int Genet. 2009;3(2): 327–32.

37. Thomas RM. Sex determination using DNA and its impact on biological anthropology. Sex estimation of the human skeleton: history, methods, and emerging techniques. 2020. pp. 343–50.

38. Butler E, Li R. Genetic markers for sex identification in forensic DNA analysis. J Forensic Investig. 2014;2(3).

39. Hagelberg E, Hofreiter M, Keyser C. Ancient DNA: the first three decades. Phil Trans Royal Soc B. 2015;370(1660):20130371.

40. Dash HR, Rawat N, Das S. Alternatives to amelogenin markers for sex determination in humans and their forensic relevance. Mol Biol Rep. 2020;47(3):2347–60.

41. Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, et al. Origins and genetic legacy of neolithic farmers and hunter-gatherers in Europe. Science 2012;336(6080):2347–60.

42. Skoglund P, Storå J, Götherström A, Jakobsson M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. J Archaeol Sci. 2013;40(12):4477–82.

43. Mittnik A, Wang C-C, Svoboda J, Krause J. A Molecular approach to the sexing of the triple burial at the upper paleolithic site of Dolní Věstonice. PLoS One. 2016;11(10):4477.

44. Loreille O, Ratnayake S, Bazinet AL, Stockwell TB, Sommer DD, Rohland N, et al. Biological sexing of a 4000-year-old Egyptian Mummy head to assess the potential of nuclear DNA recovery from the most damaged and limited forensic specimens. Genes 2018;9(3):e0163019.

45. Fowler C, Olalde I, Cummings V, Armit I, Büster L, Cuthbert S, et al. A high-resolution picture of kinship practices in an Early Neolithic tomb. Nature 2022;601(7894):584–7.

46. Sjögren KG, Olalde I, Carver S, Allentoft ME, Knowles T, Kroonen G, et al. Kinship and social organization in Copper Age Europe. A cross-disciplinary analysis of archaeology, DNA, isotopes, and anthropology from two Bell Beaker cemeteries. PLoS One. 2020;15(11):584.

47. Villalba-Mouco V, Oliart C, Rihuete-Herrada C, Childebayeva A, Rohrlach AB, Fregeiro MI, et al. Genomic transformation and social organization during the Copper Age-Bronze Age transition in southern Iberia. Sci Adv. 2021;7(47):e0241278.

48. Reitsema LJ, Mittnik A, Kyle B, Catalano G, Fabbri PF, Kazmi ACS, et al. The diverse genetic origins of aClassical period Greek army. Proc Natl Acad Sci. 2022;119(41):e2205272119.

49. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 2010;463(7282):757–62.

50. Pinhasi R, Fernandes D, Sirak K, NovakM, Connell S, Alpaslan-Roodenberg S, et al. Optimal ancient DNA Yields from the inner ear part of the human petrous bone. PLoS One. 2015;10(6):757.

51. Trinkaus E. The labyrinth of human variation. Proc Natl Acad Sci U S A. 2018;115(16):3992–4.

52. Charlton S, Booth T, Barnes I. The problem with petrous? A consideration of the potential biases in the utilization of pars petrosa for ancient DNA analysis.World Archaeol. 2019;51(4):3992–4.

53. Miller I, Gianazza E, Eberini I. Encore – Sex dependency of the proteome. J Proteomics. 2020;212:574.

54. Commission E, Research D-Gf. Innovation. Gendered innovations: how gender analysis contributes to research: report of the expert group 'Innovation through gender': Publications Office. 2013.

55. Ziganshin RH, Berezina NY, Alexandrov PL, Ryabinin VV, Buzhilova AP. Optimization of method for human sex determination using peptidome analysis of teeth Enamel from teeth of different biological

generation, Archeological Age, and degrees of taphonomic preservation. Biochemistry Moscow. 2020;85(5):614–22.

56. Jágr M, Eckhardt A, Pataridis S, Broukal Z, Dušková J, Mikšík I. Proteomics of human teeth and saliva. Physiol Res. 2014;63(Suppl 1):614–22.

57. Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. AM J HUM GENET. 1992;50(2):S141–54.

58. Gil-Bona A, Bidlack FB. Tooth Enamel and Its Dynamic Protein Matrix. Int J Mol Sci. 2020;21(12):4458.

59. Fincham AG, Bessem CC, Lau EC, Pavlova Z, Shuler C, Slavkin HC, et al. Human developing enamel proteins exhibit a sex-linked dimorphism. Calcif Tissue Int. 1991;48(4):288–90.

60. Porto IM, Laure HJ, Tykot RH, de Sousa FB, Rosa JC, Gerlach RF. Recovery and identification of mature enamel proteins in ancient teeth. Eur J Oral Sci. 2011;119 (Suppl 1):288–90.

61. Porto IM, Laure HJ, Sousa FBd, Rosa JC, Gerlach RF. New techniques for the recovery of small amounts of mature enamel proteins. J Archaeol Sci. 2011;38(12):83–7.

62. Yi L, Piehowski PD, Shi T, Smith RD, Qian WJ. Advances in microscale separations towards nanoproteomics applications. J Chromatogr A. 2017;1523:3596–604.

63. Castiblanco GA, Rutishauser D, Ilag LL, Martignon S, Castellanos JE, MejíaW. Identification of proteins from human permanent erupted enamel. Eur J Oral Sci. 2015;123(6):40–8.

64. Stewart NA,Molina GF, Mardegan Issa JP, Yates NA, Sosovicka M, Vieira AR, et al. The identification of peptides by nanoLCMS/MS from human surface tooth enamel following a simple acid etch extraction. RSC Adv. 2016;6(66):390–5.

65. Stewart NA, Gerlach RF, Gowland RL, Gron KJ, Montgomery J. Sex determination of human remains from peptides in tooth enamel. Proc Natl Acad Sci U S A. 2017;114(52):61673–9.

66. Lugli F, Di Rocco G, Vazzana A, Genovese F, Pinetti D, Cilli E, et al. Enamel peptides reveal the sex of the Late Antique 'Lovers of Modena'. Sci Rep. 2019;9(1):13649.

67. Lugli F, Figus C, Silvestrini S, Costa V, Bortolini E, Conti S, et al. Sex-related morbidity and mortality in non-adult individuals from the Early Medieval site of Valdaro (Italy): the contribution of dental enamel peptide analysis. JArchaeol Sci Rep. 2020;34.

68. Gowland R, Stewart NA, Crowder KD, Hodson C, Shaw H, Gron KJ, et al. Sex estimation of teeth at different developmental stages using dimorphic enamel peptide analysis. Am J Phys Anthropol. 2021;174(4):859–69.

69. Gasparini A, Lugli F, Silvestrini S, Pietrobelli A, Marchetta I, Benazzi S, et al. Biological sex VS. Archaeological Gender: Enamel peptide analysis of the horsemen of the Early Middle age necropolises of Campochiaro (Molise, Italy). JArchaeol Sci Rep. 2022;859.

70. Avery LC, Prowse TL, Findlay S, de Seréville-Niel CC, Brickley MB. Pubertal timing as an indicator of early life stress in Roman Italy and Roman Gaul. Am J Biol Anthropol. 2023;180(3):548–60.

71. Haas R, Watson J, Buonasera T, Southon J, Chen JC, Noe S, et al. Female hunters of the early Americas. Sci Adv. 2020;6(45):548.

72. Casas-Ferreira AM, del Nogal-Sánchez M, Arroyo ÁE, Vázquez JV, Pérez-Pavón JL. Fast methods based on mass spectrometry for peptide identification. Application to sex determination of human remains in tooth enamel. Microchem J. 2022;181.

73. Rebay-Salisbury K, Janker L, Pany-Kucera D, Schuster D, Spannagl-Steiner M, Waltenberger L, et al. Child murder in the Early Bronze Age: proteomic sex identification of a cold case from Schleinbach, Austria. Archaeol Anthropol Sci. 2020;12(11):265.

74. Rebay-Salisbury K, Bortel P, Janker L, Bas M, Pany-Kucera D, Salisbury RB, et al. Gendered burial practices of early Bronze Age children align with peptide-based sex identification: a case study from Franzhausen I, Austria. J Archaeol Sci. 2022;139.

75. Wasinger VC, Curnoe D, Bustamante S, Mendoza R, Shoocongdej R, Adler L, et al. Analysis of the preserved amino acid bias in peptide profiles of Iron Age teeth from a tropical environment enable sexing of individuals using amelogenin MRM. Proteomics. 2019;19(5).

76. Froment C, Hourset M, Sáenz-Oyhéréguy N, Mouton-Barbosa E, Willmann C, Zanolli C, et al. Analysis of 5000 year-old human teeth using optimized large-scale and targeted proteomics approaches for detection of sex-specific peptides. J Proteomics. 2020;211:1800341.

77. Parker GJ, Yip JM, Eerkens JW, Salemi M, Durbin-Johnson B, Kiesow C, et al. Phinney BS. Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. J Archaeol Sci. 2019;101:169–80.

78. Štamfelj I. Sex estimation based on the analysis of enamel peptides: false assignments due to AMELY deletion. J Archaeol Sci. 2021;130:169.

79. Parker GJ, Buonasera T, Yip JM, Eerkens JW, Salemi M, Durbin-Johnson B, et al. AMELY deletion is not detected in systematically sampled reference populations: a reply to Štamfelj. J Archaeol Sci. 2021;130:105345.

80. Welker F, Ramos-Madrigal J, Gutenbrunner P, Mackie M, Tiwary S, Rakownikow Jersie-Christensen R. The dental proteome of Homo antecessor. Nature 2020;580(7802):235–8.

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 5.6). 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 [114]. 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 [115]. Another study showed that a loss of av. 6.1 μm/min of enamel under phosphoric acid conditions (50%, 3 min) [116]. However, in the case of this study, low concentrations of H_2O_2 (3%) and HCl (5%) were used [117]. These concentrations are lower than the H_2O_2 concentration used in cosmetic tooth whitening (6%) in accordance with European directives [118].

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.

6.2.1 Scanning electron microscope

Scanning electron microscope (SEM) is a non-destructive method that uses electrons instead of light to form an image in one-dimension. Each tooth was scanned before and after chemical etching, using the same scanning parameters. The comparison of these scans allowed us to examine the microscopic changes in the enamel surface that occurred during etching.

6.2.2 Micro-computed tomography

Micro-computed tomography (micro-CT) is another non-destructive method that uses X-rays. This method involves scanning a tooth in two-dimensions. The rotation of the tooth for complementary scans allows the creation of digital reconstruction in three-dimensions. Each tooth was also scanned before and after the chemical etchings with the same scanning parameters. The comparison of these scans allowed us to calculate enamel volume loss during etching.

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 [119]. 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 [120]. In addition, 32 teeth from non-adult individuals were selected because of the impossibility of sex estimation using traditional morphological methods [121, 122]. Finally, scanning electron microscope and micro-computed tomography were used to evaluate the minimally-invasive proteomics method.

Undertaking biological sex assessment of human remains: Applicability of a minimally-invasive methods for the proteomic sex estimation from enamel peptides
Jaroslav Brůžek*, Ivan Mikšík*, Anežka Pilmann Kotěrová, Marine Morvan, Sylva Drtikolová Kaupová, Jana Velemínská, Frédéric Santos, Alžběta Danielisová, Eliška Zazvonilová, Bruno Maureille, Petr Velemínský
Journal of Cultural Heritage, **2023**, *in revisions*.

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)
Sylva Drtikolová Kaupová*, Jaroslav Brůžek, Jiří Hadrava, Ivan Mikšík, Marine Morvan, Lumír Poláček, Lenka Půtová, Petr Velemínský
Archaeological and Anthropological Sciences, **2022**, *under review*.

Article

Undertaking the biological sex assessment of human remains: the applicability of minimally-invasive methods for proteomic sex estimation from enamel peptides

Jaroslav Brůžek[1,2,*], Ivan Mikšík[3,*], Anežka Pilmann Kotěrová[1], Marine Morvan[3,4], Sylva Drtikolová Kaupová[5], Frédéric Santos[2], Alžběta Danielisová[6], Eliška Zazvonilová[1,6], Bruno Maureille[2], Petr Velemínský[5]

[1] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, 128 00 Prague 2, Czech Republic

[2] Univ. de Bordeaux, CNRS, Ministère de la Culture, PACEA, UMR 5199, F-33600 Pessac, France

[3] Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 532 10 Pardubice, Czech Republic

[4] Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20 Prague 4, Czech Republic

[5] Department of Anthropology, National Museum, Václavské náměstí 1700/68, 110 00 Prague 1, Czech Republic

[6] Institute of Archaeology of the Czech Academy of Sciences, Letenská 4, 118 00 Prague 1, Czech Republic

* These authors contributed equally

Corresponding author: koterova@natur.cuni.cz (A.P.K.)

Preprint citation: Brůžek, J., Mikšík, I., Pilmann Kotěrová, A., Morvan, M., Drtikolová Kaupová, S., Santos, F., Danielisová, A., Zazvonilová, E., Maureille, B., Velemínský, P., Undertaking Sex Assessment of Human Remains within Cultural Heritage: Applicability of Minimally-Invasive Methods for Proteomic Sex Estimation from Enamel Peptides, **2023**. Available at SSRN 4439221. http://dx.doi.org/10.2139/ssrn.4439221

Abstract: Being a part of the cultural heritage, skeletal human remains and grave objects are often the only evidence of people who lived many years, or even centuries or millennia, ago, and their preservation for future generations is thus of the utmost importance. The first task in analyzing skeletal remains is to build a biological profile of the individual, including in particular a sex estimation. Recently developed proteomic sex analysis, based on the detection of two sex dependent forms of the amelogenin protein in tooth enamel, could offer a minimally-invasive and reliable approach applicable to both recent and past populations.

The aims of the present study are: 1) to validate the proteomic sex estimation approach with a delicate, minimally-destructive protocol using protein etching in recent and sub-recent identified samples of adult individuals; 2) for the first time, to evaluate the invasiveness of the extraction of amelogenin protein from teeth for proteomic analysis via scanning electron microscope (SEM) and microcomputed tomography (micro-CT); 3) to apply the method to an archaeological sample of unknown adult and juvenile individuals.

An assemblage of 60 teeth (32 males and 28 females) of recent and sub-recent origin was used to validate the approach. A sub-sample of 20 teeth (10 males and 10 females) was used to assess the invasiveness of the amelogenin extraction procedure. For the application of the method, samples of 15 adult and 32 juvenile teeth, both originating from medieval populations, were used.

Proteomic sex estimation achieved 100% accuracy in this sample. An SEM and micro-CT comparison of the dental surfaces before and after chemical treatment showed an approximately 10% loss of enamel and only 2% loss of dentine. The suitability and minimally-invasive character of the protocol for proteomic analysis in biological sex estimation was demonstrated, as was its applicability to archaeological samples.

Keywords: Human skeletal collections, cultural heritage, proteomics, sex estimation, sampling

1. Introduction

Human skeletal remains are an inseparable part of the cultural heritage, and are often the only information about people living in the past. They are preserved in museum and university collections, and are used for elaborating biological profiling methods [1]. Unfortunately, there are differences in legislation between countries [2–5] that affect the handling of skeletons, as well as differences in ethical standards e.g. [6,7]. The common denominators include some laws, such as the Valletta Treaty (Council of Europe, ETS No. 143, 1992), according to which all biological remains, artefacts, objects and any other remains of humankind from past epochs are subsumed under Cultural Heritage and are protected within the scope of a common European legacy. We must preserve for the future the status of archaeological human remains as part of the cultural heritage, and as indispensable empirical sources for the reconstruction of human population history through time and space [8].

The uniqueness of skeletal remains from any period leads conservators to minimize any interference with their integrity. Knowledge of the methods used and the degree of damage they cause are of enormous importance, but such damage is not always fully known [9]. Within the field of human osteology, the debate on ethical issues regarding the sampling, research and display of archaeological human and animal remains has been ongoing for decades, and has become more vigorous with the recent increase in ancient biomolecular studies [10,11].

Although the amount of sample needed for analysis can be just a few milligrams of bone or dental tissue, museums too are being inundated with destructive sampling requests. With respect to anthropological collections, we must consider the need to reduce or completely eliminate the mechanical destruction of skeletal remains deemed useful for scientific evaluation.

The preservation of organic molecules in teeth and bones has proven crucial for understanding the human past [12]. In recent years, proteomics has become an attractive method for studying human, animal, and biological profiles and origins. It is an alternative to DNA analysis, which is limited by the DNA amplification that is present in ancient samples, contamination, its high cost, and the limited preservation of nuclear DNA [13–15]. Currently, three approaches are available for estimating biological sex: osteology, genomics, and proteomics [16], but little is known about the relative reliability of these methods in applied settings [17].

Correct and reliable biological sex estimation of skeletal remains is important in various areas, from bioarchaeology to forensic science. The varying degree of preservation of skeletal remains limits the methods that can be used [18]. Morphological methods are based on the existence of sexual dimorphism in the skeleton, and are subject to population specificity which causes various error rates

[19,20]. The only part of the skeleton that allows a reliable estimate of sex is the pelvis [21]; in archaeological human skeletal assemblages, however, the pelvis is often heavily damaged or completely absent [22]. In addition, morphological methods have another major limiting factor: the inability to reliably or accurately estimate sex from immature elements with any degree of consistency [23,24]. According to Buonasera et al., biological sex estimation was possible for 100% of an archaeological sample by proteomics, for 91% by genomics, and for 51% by osteology; the agreement among the methods was high, but there were conflicts [17].

Proteomics provides a new, seemingly simple and cost-effective way to conduct sex estimation without the risk of contamination. It uses two sexually distinct forms of the amelogenin protein found in tooth enamel, detectable by liquid chromatography-tandem mass spectrometry (LCMS/MS); the AMELY protein (amelogenin Y isoform) is present in enamel dental tissue only in males, while AMELX (isoform X) can be found in both sexes [13]. Enamel is the most mineralized part of the tooth, with mineral content making up about 97% of mature enamel; the rest is formed of proteins and other components, such as water. During enamel formation and maturation the matrix is removed almost completely through enzymatic degradation by proteases, resulting in its hardening and the extensive deposition of calcium-based minerals.

Amelogenin is the predominant protein in the developing extracellular enamel matrix [25]. Although studies estimating biological sex by proteomic analysis of dental enamel use the same principle, they differ in their levels of sampling and pre-treatment. Some studies use more destructive sampling in the form of small enamel chunks extracted with a dentist's drill [26–30]; others offer a less invasive approach in the form of etching the enamel surface with hydrochloric acid [13,31,32].

1.1. Research aim

In this study, we aimed to (1) validate peptide-based enamel sex identification both in teeth from individuals of known sex from a recent population and in dissected individuals from the mid-20th century [33]; (2) assess the effect of protein extraction on the structural integrity of human dental enamel (damage to the enamel surface) via scanning electron microscope (SEM) and micro-computed tomography (micro-CT); and (3) to apply the method to an archaeological assemblage of unknown adult and immature individuals.

2. Material

To obtain and validate a predictor for biological sex based on enamel proteins, we used two different assemblages:

(1) a validation sample of teeth of individuals of known biological sex to verify the reliability of the method; a sub-sample of 20 teeth randomly selected from the validation sample served to verify the effect of protein etching on tooth enamel quality;

(2) an application sample of archaeological teeth from a museum collection, for which sex (in the case of adult individuals) was originally estimated by a morphological method.

2.1. Validation sample

This contains the identified teeth of adult individuals from two samples of the Czech population (Tab. 1). Group (1a) consists of 30 teeth of adult individuals of known age and sex (15 male and 15 female) from the recent Prague population, provided by dentists. Extractions were performed for medical reasons.

Group (1b) originates from the Pachner collection, housed in the Department of Anthropology and Human Genetics of the Faculty of Science, Charles University, Prague. It too comprises 30 identified teeth of individuals of known age and sex from the Prague population (17 male and 13 female), but who were autopsied in the 1940s. This collection dates to the 1930s and consists of individuals of known sex, age (date of birth and date of autopsy), and status (they represented the lower socio-economic classes of Bohemia) [33,34].

2.2. Application sample

This contains 15 adult and 32 non-adult teeth from the burial grounds of an Early Medieval population (9–11th century AD) at the Mikulčice agglomeration in Moravia [35], stored in the Department of Anthropology of the National Museum, Prague (Tab. 1). We selected adult individuals based on the conflicting results of morphological sex estimation by osteological methods between those published in the 1960s [36–38] and the revised sex estimates published in 2020 [39]. We also included non-adult individuals whose biological sex cannot be assessed by morphological methods [32,40]; they were selected to study the differences in diet between boys and girls among the early Slavs, which is the subject of another study [41]. This research was approved by the Institutional Review Board of Charles University.

Table 1. Composition of validation sample, sub-sample subjected to micro-CT and SEM and application sample.

	Validation sample of identified individuals (n=60)							Application sample (n=47)	
	Extraction for medical reasons			Pachner's Identified Collection				Medieval population (9-11th century AD)	
	F (n)	M (n)	F and M (n)	F (n)	M (n)	F and M (n)	Total (n)	Adults	Non-adults
PSE	10	10	20	8	12	20	40	15	32
PSE + micro-CT and SEM	5	5	10	5	5	10	20	not performed	not performed

PSE = proteomic sex estimation; SEM = scanning electron microscope; micro-CT = microtomodensitometry

3. Methods

3.1. Biological sex estimation

Proteomic estimation of sex was done following the method of Stewart et al. [13], which is based on the analysis of two dimorphic peptides of amelogenin: AMELY-(58-64) peptide and AMELX-(44-52)

peptide from tooth enamel. The teeth of males contain both peptides (AMELY and AMELX), while female teeth contain only one (AMELX).

1 step – Sample preparation

The tooth surface was cleaned with water to remove obvious surface contaminants. The tooth crown was then placed in the cap of a 2-mL Eppendorf tube, washed with 3% H2O2 (200 µL) for 30 seconds and then rinsed with ultrapure water. It was then treated with 200 µL of 5% (v/v) HCl. An initial etch was performed by lowering the tooth onto the HCl and maintaining contact for 2 minutes. This first etch was discarded. A second 2-minute etch was collected for the analysis.

2 step – Protein etching

For peptide extraction, a C18 resin-loaded ZipTip (ZTC18S096; EMD Millipore) was used. This had previously been conditioned three times with acetonitrile (100%), and three times with formic acid 0.1% (v/v) (each draw discarded). Peptide solution was then pipetted 10 times to maximize binding to the C18 resin of the ZipTip material. Finally, the ZipTip was washed six times with formic acid 0.1% (v/v) (each wash discarded). Bonded peptides were eluted by 4-µL of acetonitrile/formic acid (60%/0.1%, v/v) and the resulting solution was collected into small vials. This fraction was lyophilized and dissolved in 20 µL of formic acid (2%) in ultrapure water, centrifuged on a benchtop centrifuge for 5 minutes to remove any particulate matter, and transferred (18 µL) to autosampler vials. This sample was injected for analysis by reversed-phase nano-liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS).

3 step – Analysis method

The nanoLC apparatus was a Proxeon Easy-nLC (Proxeon, Odense, Denmark) coupled to a MaXis Q-TOF (quadrupole – time of flight) mass spectrometer with ultra-high resolution (Bruker Daltonics, Bremen, Germany) by nanoelectrosprayer. Five microliters of the peptide mixture were injected into a NS-AC-11 BioSphere C18 column (particle size: 5µm, pore size: 12 nm, length: 152 mm, inner diameter: 75 µm), with an NS-MP-10 BioSphere C18 pre-column (particle size: 5 µm, pore size: 12 nm, length: 20 mm, inner diameter: 100 µm), both obtained from NanoSeparations (Nieuwkoop, Netherlands).

Peptide separation was achieved via a linear gradient between mobile phase A (ultrapure water) and B (acetonitrile), both containing formic acid 0.1% (v/v). Separation was started with a gradient elution from 5% to 30% mobile phase B at 70 minutes. The next step was gradient elution to 50% B in 10 minutes, followed by a gradient to 100% B in 10 minutes. Finally, the column was eluted with 100% B for 30 minutes. Equilibration before the subsequent run was achieved by washing the column with 3 µl of 5% mobile phase B for 5 minutes. The flow rate was 0.25 µL.min-1, and the column was held at ambient temperature (25°C). Online nano-electrospray ionization (easy nano-ESI) in the positive mode was used. The ESI voltage was set at +4.5 kV, spectra rate 3 Hz. Operating conditions: drying gas (N2), 4 L.min-1; drying gas temperature, 180°C; nebulizer pressure, 100 kPa. Experiments were performed by scanning from 420 to 550 m/z.

Proteomic determination of biological sex was done according to the peak area of EIC (extracted ion chromatogram) at a retention time of 35 minutes for m/z 440.22 (AMELYpeptide, marked as Y in the tables) and at a retention time of 43 minutes for m/z 540.28 (AMELX-peptide, marked as X in the tables).

3.2. Evaluation of the effect of protein etching on dental enamel

The effect of etching on enamel was performed in a subset of 20 teeth of individuals of known sex.

Scanning electron microscope (SEM)

Samples were scanned on a Hitachi S-3700N Scanning Electron Microscope (SEM) at the National Museum in Prague. The teeth were documented overall, and then two areas on each were selected and scanned at several resolutions without tooth surface treatment (mechanical or ultrasonic cleaning). After the protein etching of the teeth, the same samples and the same areas on them were scanned again so that the changes that had occurred could be evaluated.

Microcomputed tomography (micro-CT)

Each tooth was scanned twice: before and after the application of the protocol of extraction for proteomic sex estimation with the same scanning parameters. The first scan took place under the same conditions as the SEM analysis of the surface (without mechanical or ultrasonic cleaning). Microtomodensitometric (micro-CT) data for these teeth were acquired using the micro-CT platform in the PACEA (Bordeaux). They were obtained with a "v|tome|x s 240" micro-CT scanner microfocus tube (GE Sensing and Inspection Technologies Phoenix X ray). A total of 1750 radiographic projections (i.e. 1750 angular increments for 360° rotation) were acquired with the following scan parameters: voltage 100 kV, current 160 µA, exposure 500 ms, voxel size 27.5 µm. The micro-CT data were reconstructed using Phoenix datos|x reconstruction 2 software and then exported as a 16-bit TIFF image stack. VG studio max software (Volume Graphics, release 2.2, Heidelberg, Germany) was used for the virtual slice visualization and three-dimensional rendering.

Using Avizo 9.5 (Thermo Fisher Scientific), a semi-automatic threshold-based segmentation was performed on the reconstructed images with subsequent manual corrections. After segmentation, volumetric reconstruction and visualisation of the micro-CT slices were performed using Avizo v. 9.5 software. Enamel thickness mapping before and after protein etching was performed with the Surface Distance module on the Avizo 9.5 software.

4. Results:

4.1. Biological sex estimation via peptides in dental enamel

A biological female determination of the sample is indicated by the detection of only one diagnostic peptide (AMELX); a biological male determination is indicated by the detection of both diagnostic peptides derived from each isoform (AMELX-(44-52), or AMELY-(58-64)) of amelogenin proteins. For the total 60 samples of individuals of known sex, the accuracy of sex estimation was absolute (Tab. 2).

Table 2. Amelogenin-sex estimation in a validation sample of adult teeth of known age and sex from a Czech population (30 males and 30 females).

Sample number	Sub-sample	Sex declared	Tooth sampled	Age (yrs)	Peak area		Proteomic sex
					X	Y	
Proteomic sex estimation							
8	A	F	$_3$M	23	946,208	Not present	F
13	A	F	$_2$M	39	1,896,641	Not present	F

15	A	F	^{3}M	20	1,154,743	Not present	F
16	A	F	$_{3}$M	21	1,527,914	Not present	F
17	A	F	$_{3}$M	40	1,805,977	Not present	F
18	A	F	$_{3}$M	25	874,56	Not present	F
19	A	F	P^{2}	64	3,008,422	Not present	F
20	A	F	M$_{3}$	21	1,325,183	Not present	F
21	A	F	M$_{3}$	55	3,032,252	Not present	F
22	A	F	M^{3}	24	673,982	Not present	F
1	A	M	$_{3}$M	17	748,564	550,882	M
2	A	M	M^{3}	41	1,371,502	2,604,241	M
23	A	M	^{3}M	28	973,924	1,296,839	M
24	A	M	$_{3}$M	29	1,033,295	2,386,597	M
25	A	M	^{3}M	55	2,330,862	3,339,066	M
26	A	M	$_{3}$M	21	734,265	820,057	M
27	A	M	M$_{3}$	36	1,609,180	2,396,535	M
28	A	M	^{3}M	35	1,295,788	2,389,316	M
29	A	M	^{3}M	61	1,565,209	865,203	M
30	A	M	$_{3}$M	30	1,490,656	2,013,183	M
37	P	F	P$_{2}$	34	2,323,905	Not present	F
40	P	F	^{2}P	adult	929,849	Not present	F
43	P	F	M$_{3}$	51	4,721,920	Not present	F
50	P	F	$_{3}$M	38	2,488,777	Not present	F
51	P	F	I^{1}	adult	258,116	Not present	F
52	P	F	M^{1}	70	4,448,210	Not present	F
55	P	F	^{3}M	65	2,227,581	Not present	F
58	P	F	C	adult	2,259,169	Not present	F
33	P	M	C	adult	1,157,578	1,258,985	M
34	P	M	P^{1}	43	1,140,553	2,664,629	M
35	P	M	^{2}P	38	1,967,799	3,615,260	M

36	P	M	^{2}P	adult-	1,254,277	2,890,431	M
38	P	M	P^{2}	62	2,984,328	4,386,024	M
39	P	M	^{2}P	32	2,070,259	4,178,386	M
41	P	M	M^{1}	39	2,536,816	3,926,950	M
44	P	M	^{1}P	adult	903,339	1,473,504	M
45	P	M	P^{1}	32	4,197,536	7,638,730	M
56	P	M	^{1}M	78	1,487,187	1,250,741	M
57	P	M	P_2	61	3,344,463	4,761,168	M
60	P	M	^{2}M	35	2,141,434	2,029,143	M
Proteomic sex estimation with micro-CT and SEM analysis							
9	A*	F	M^{3}	32	1,384,671	Not present	F
10	A*	F	$_{3}$M	25	1,271,997	Not present	F
11	A*	F	M_3	33	1,322,932	Not present	F
12	A*	F	M_3	26	1,831,753	Not present	F
14	A*	F	M_3	26	1,258,762	Not present	F
3	A*	M	$_{3}$M	48	1,366,483	3,291,729	M
4	A*	M	M_3	20	627,076	550,347	M
5	A*	M	$_{3}$M	22	863,334	1,037,607	M
6	A*	M	^{3}M	31	611,035	1,317,945	M
7	A*	M	$_{3}$M	24	551,819	1,299,878	M
31	P*	F	^{3}M	34	2,347,440	Not present	F
32	P*	F	$_{3}$M	28	3,000,280	Not present	F
46	P*	F	M^{3}	adult	778,678	Not present	F
48	P*	F	^{2}M	32	5,039,565	Not present	F
54	P*	F	M^{1}	adult	1,210,069	Not present	F
42	P*	M	M^{1}	73	1,092,764	876,029	M
47	P*	M	M^{1}	33	2,718,667	3,445,208	M
49	P*	M	^{2}M	adult	2,905,233	1,947,298	M
53	P*	M	^{1}M	42	3,697,825	2,906,365	M

59	P*	M	^{2}M	72	2,000,080	1,784,672	M

A = group 1a (teeth from extractions in dental clinics in 2019–2021), P = group 1b (teeth from Pachner's Identified collection from the first half of the 20th century), * = sub-sample for the effect of protein etching on tooth enamel quality. SEM = scanning electron microscope; micro-CT = microtomodensitometry. Tooth nomenclature: I = incisor, C = canine, P = premolar, M = molar, upper index = upper jaw, lower index = lower jaw, side of the index indicates side of the tooth (left or right).

After this validation of the proteomic sex estimation method, the method was applied to a sample of 15 teeth of adult individuals of archaeological provenance. For the needs of another study [41], we also estimated the sex of 32 permanent teeth of juvenile individuals from a medieval population (in which morphological sex estimation is impossible). Table 3 shows the results of the proteomic sex estimation of the application sample from the Early Medieval period. In the case of adults, 7 teeth showed both peptide signals and corresponded to males, while 8 teeth with only AMELX-peptide belonged to females. In the case of immature individuals, the nanoLC-MS/MS analysis showed that 16 teeth with only an AMELX-peptide signal belonged to female individuals; 16 individuals with the presence of both AMELXpeptide and AMELY-peptide signals corresponded to males.

Table 3. Amelogenin-sex estimation in an application sample of 15 adult teeth and a sample of 32 non-adult teeth, both from burial grounds of the Early Medieval (9–11th century AD) population from the Mikulčice agglomeration in Moravia.

Skeleton number	Age-at-death (yrs) (Stloukal 1963, 1964, 1967)	Tooth class	Peak area		Proteomic sex
			X	Y	
H87	40–50	C	3,496,055	3,991,318	M
H170	50–60	$_{2}$P	3,745,990	4,076,427	M
H0171	40–59	C	3,209,222	Not present	F
H0187	50–60	P_2	693,47	1,445,756	M
H0292	30–59	P_2	835,227	Not present	F
H0314(bis)	adult	P_2	4,902,831	Not present	F
H0314	30–40	P^2	2,952,606	3,665,371	M
H0324	30–40	I_1	7,071,775	3,665,371	F
H0352	50–60	P^2	4,283,238	Not present	F
H0363	30–40	P_1	1,986,341	2,610,787	M
H0406	40–50	^{2}M	2,089,384	Not present	F
H0457	50–60	P_1	4,678,292	Not present	F

H0647	40–50	$_1$P	2,551,808	3,039,339	M
H0718	adult	P^2	5,232,203	Not present	F
H1088	adult	^{1}P	2,893,400	3,510,927	M
H73-VI	5–6	M^1	404,29	408,317	M
H143	6–7	^{1}M	985,473	1,057,251	M
H160-VI	6–7	$_1$M	806,03	Not present	F
H207	10–11	^{1}M	713,327	1,009,648	M
H247	5–6	M^1	799,429	Not present	F
H253	4	M$_1$	1,254,321	Not present	F
H266	9–10	M^1	187,41	446,308	M
H296	2	^{1}M	770,229	1,102,988	M
H315	3–4	$_1$M	1,211,179	Not present	F
H343	2–3	M^1	1,260,716	1,923,015	M
H393	4–5	$_1$M	1,186,517	Not present	F
H444	15–17	^{1}M	799,936	1,221,249	M
H447	2–3	$_1$M	776,379	644,6	M
H454	6–7	$_1$M	949,155	Not present	F
H455	2–3	$_1$M	533,605	Not present	F
H462	4–5	M$_1$	1,498,764	Not present	F
H473	4–5	M$_1$	556,056	702,444	M
H489	9	$_1$M	759,236	1,023,086	M
H496	3–4	^{1}M	560,303	Not present	F
H497	6–7	$_1$M	999,491	694,996	M
H526	2–3	M^1	348,887	593,259	M
H538	9	$_1$M	297,597	Not present	F
H550	4–5	M^1	566,669	Not present	F
H582	12–14	$_1$M	733,286	1,023,541	M
H751	5–6	$_1$M	388,359	Not present	F
H792	5–6	$_1$M	1,006,139	1,233,096	M

H878	15–16	$_1$M	915,958	825,651	M
H881	6	$_1$M	949,087	Not present	F
H1041	6	M_1	367,721	Not present	F
H1058	5	M_1	465,283	Not present	F
H1154	5	M_1	671,2	1,374,627	M
H1171	5–6	$_1$M	640,652	Not present	F

Tooth nomenclature: I = incisor, C = canine, P = premolar, M = molar, upper index = upper jaw, lower index = lower jaw, side of the index indicates side of the tooth (left or right)

4.2. Tooth enamel changes due to protein etching

From a macroscopic point of view (visual evaluation), protein extraction does not affect the observed details of the tooth surface. It is only necessary to mention subtle colour changes and the reduction of the gloss of the enamel surface. SEM analysis was used to examine and compare the surface morphology of each tooth from the validation group before protein etching and after etching. Figures 1 and 2 show SEM images of the enamel surface of two examples: tooth 42 from the Pachner collection and tooth 6 from recent extractions in dental clinics. The collected SEM images show that all the samples showed variable and distinct surface changes. Because the teeth were not treated in any way before applying the etching protocol, after protein extraction from the enamel all traces of tartar and other substances present on the enamel surface disappeared.

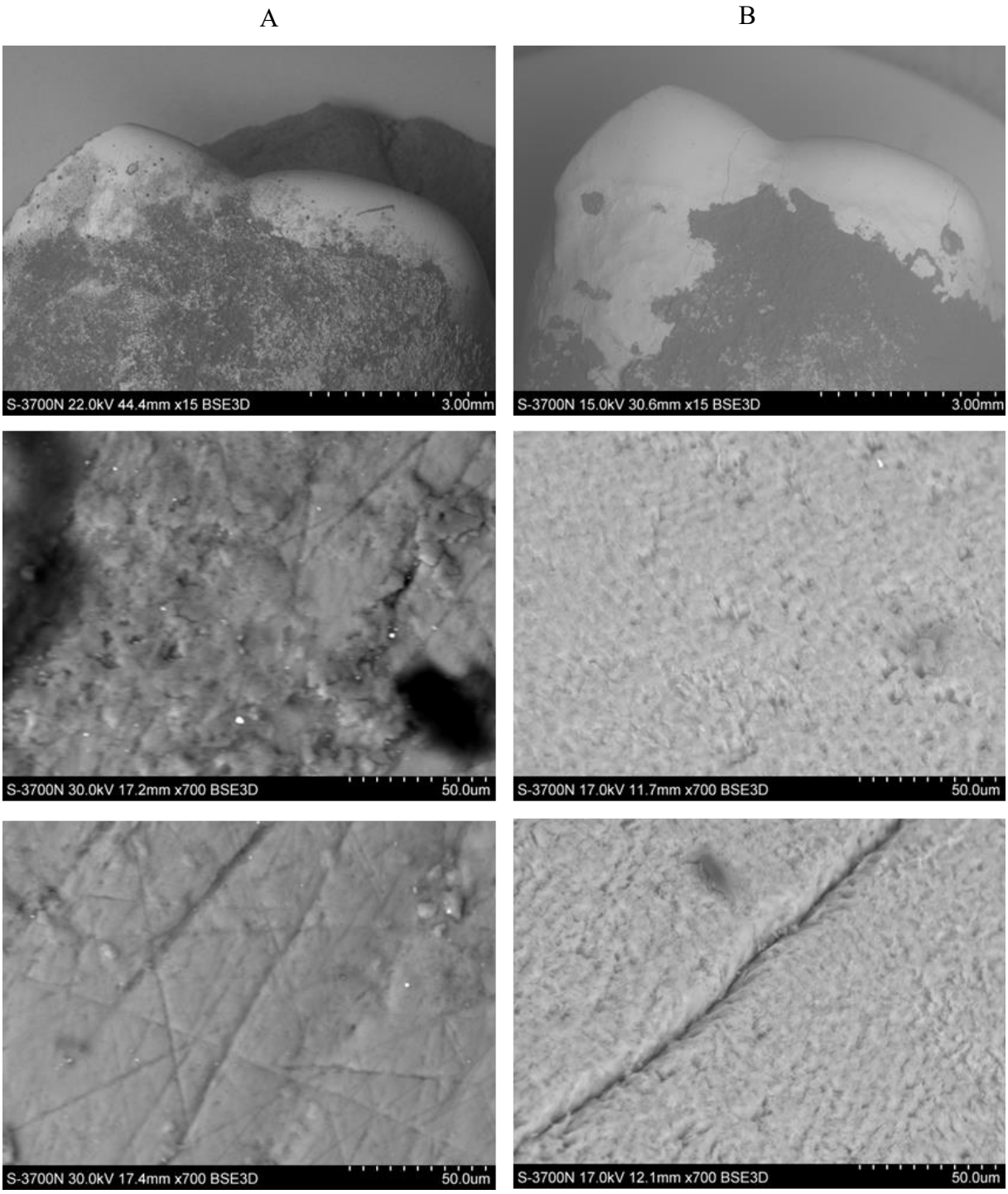


Fig. 1. Comparison of enamel surface alteration obtained by SEM before (A) and after (B) protein etching. Upper row – natural size; Middle row – area 1 at 500x magnification; Bottom row – area 2 at 500x magnification. Sample number 42 (from Pachner's Identified collection, from the first half of the 20th century).

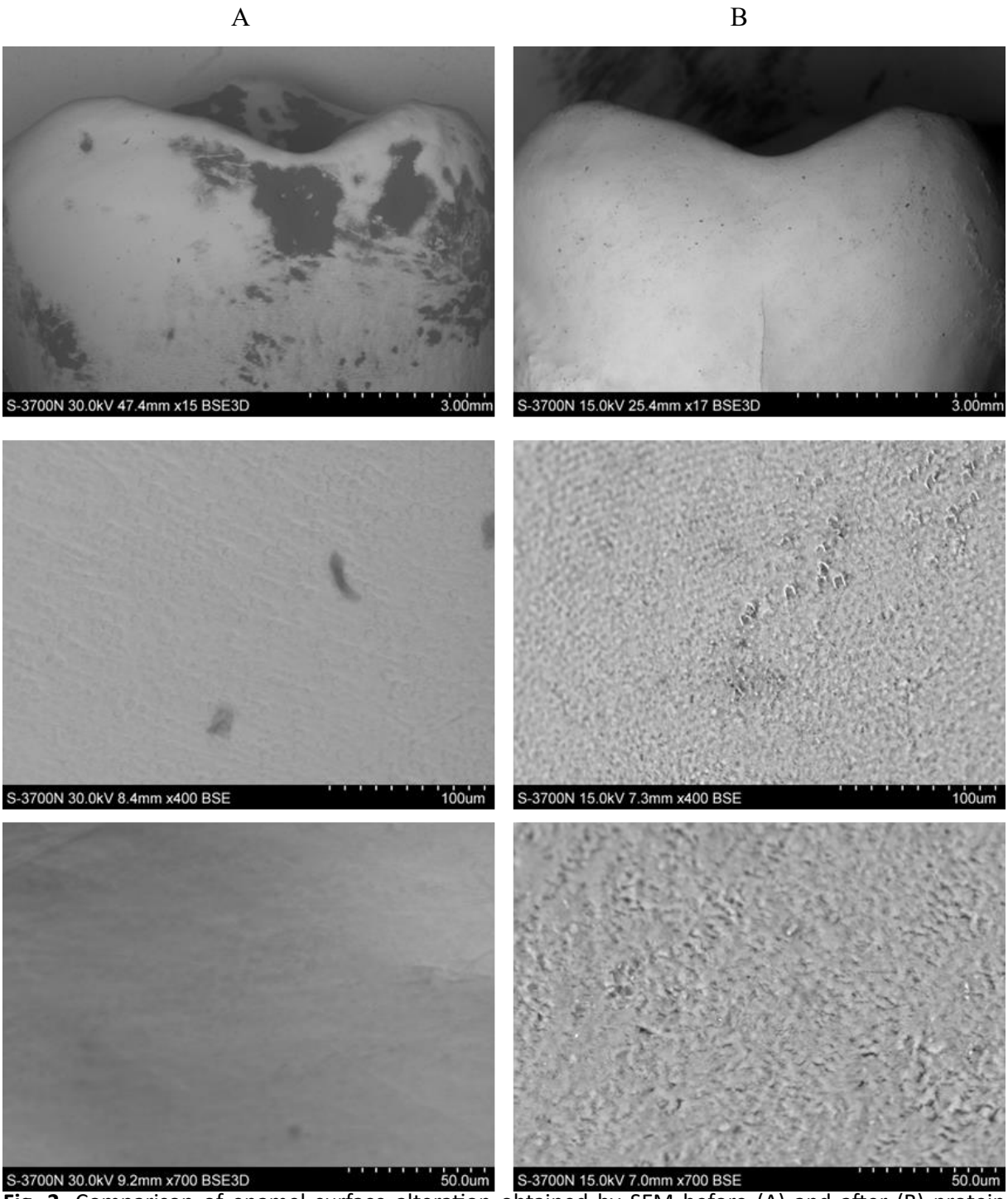


Fig. 2. Comparison of enamel surface alteration obtained by SEM before (A) and after (B) protein etching. Upper row – natural size; Middle row – area 1 at 400x magnification; Bottom row – area 1 at 700x magnification. Tooth sample number 6 (tooth from extraction in dental clinics in 2019–2021).

Additionally, most of the images from the SEM showed slight alterations of the surface with the presence of microporosity, possibly with deepening grooves and defects. The method of sample preparation for proteomic sex estimation appears to weaken the sample structure, which is also evident from the micro-CT analysis. We observed microcracks in some teeth before the analysis, which after protein etching were slightly wider (Figure 3).

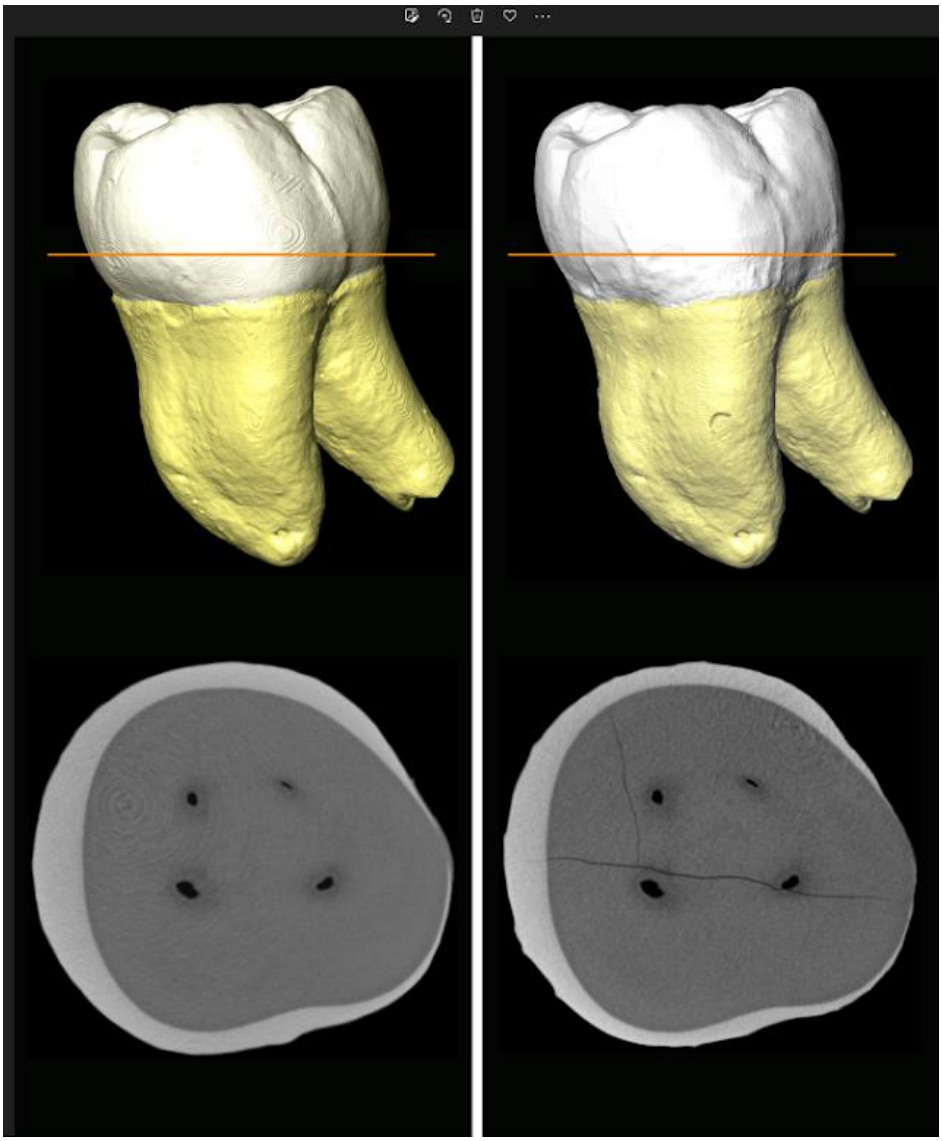


Fig. 3. Structural changes in tooth due to protein etching (Tooth sample number 3). Top left: 3D reconstruction of Tooth 3 before protocol application. Top right: after enamel etching according to the protocol to obtain protein for sex estimation. The horizontal line shows the plane of the section.

Quantitative changes in enamel thickness using colour maps are shown in Fig. 4. In the given example, the enamel surface on the occlusal plane shows visible enlargement of a cavity caused by tooth decay, which is caused by the action of the extraction solution. The enamel thickness is less, and the cavity after caries has increased due to the action of the extraction solution.

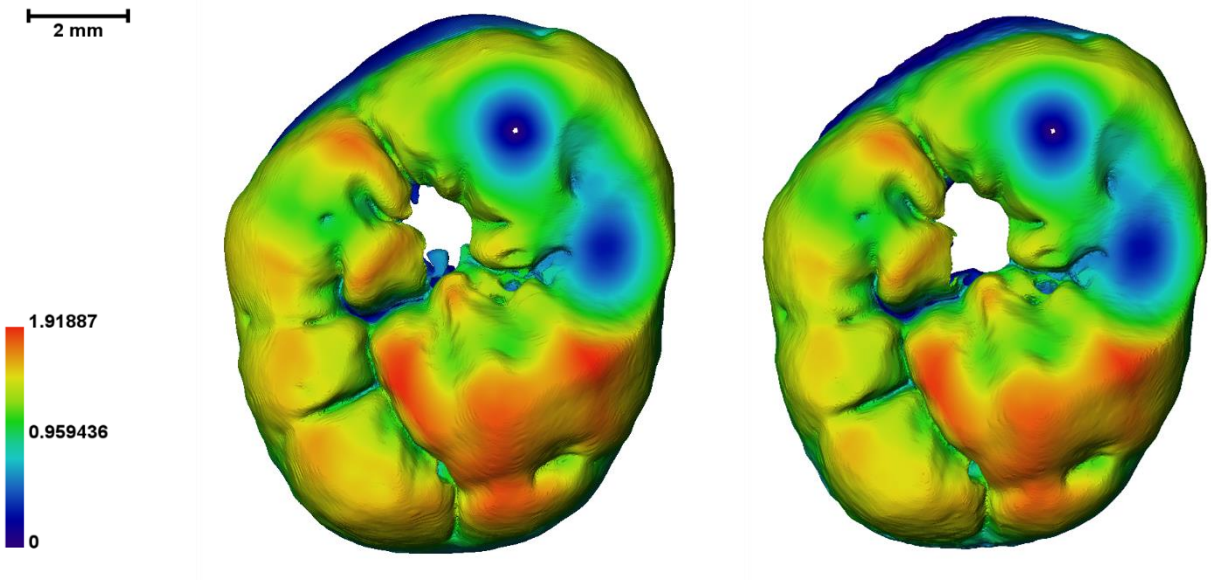


Fig. 4. Colour map of the enamel thickness on the molar occlusal surface. Red indicates areas of thicker enamel surface. The white area in the middle corresponds to an opening caused by dental caries. As a result of the extraction and the action of the acid, the opening became larger (left before etching, right after etching). Tooth sample number 46.

Another example of quantitative changes in enamel thickness using colour maps is shown in Fig. 5. After the application of the protein extraction protocol on the tooth enamel, there is a reduction in the thickness of the enamel, which is manifested by a change in colour from red to blue. This reduction in enamel thickness affects all of the teeth that have been treated. The colours are always less intense on the right colour map. Volume changes of both enamel and dentine, as well as of the entire tooth, are shown in Tab. 4. Due to the application of the extraction protocol, approximately 10% of the enamel is lost, but we also observed a 2% loss of dentine.

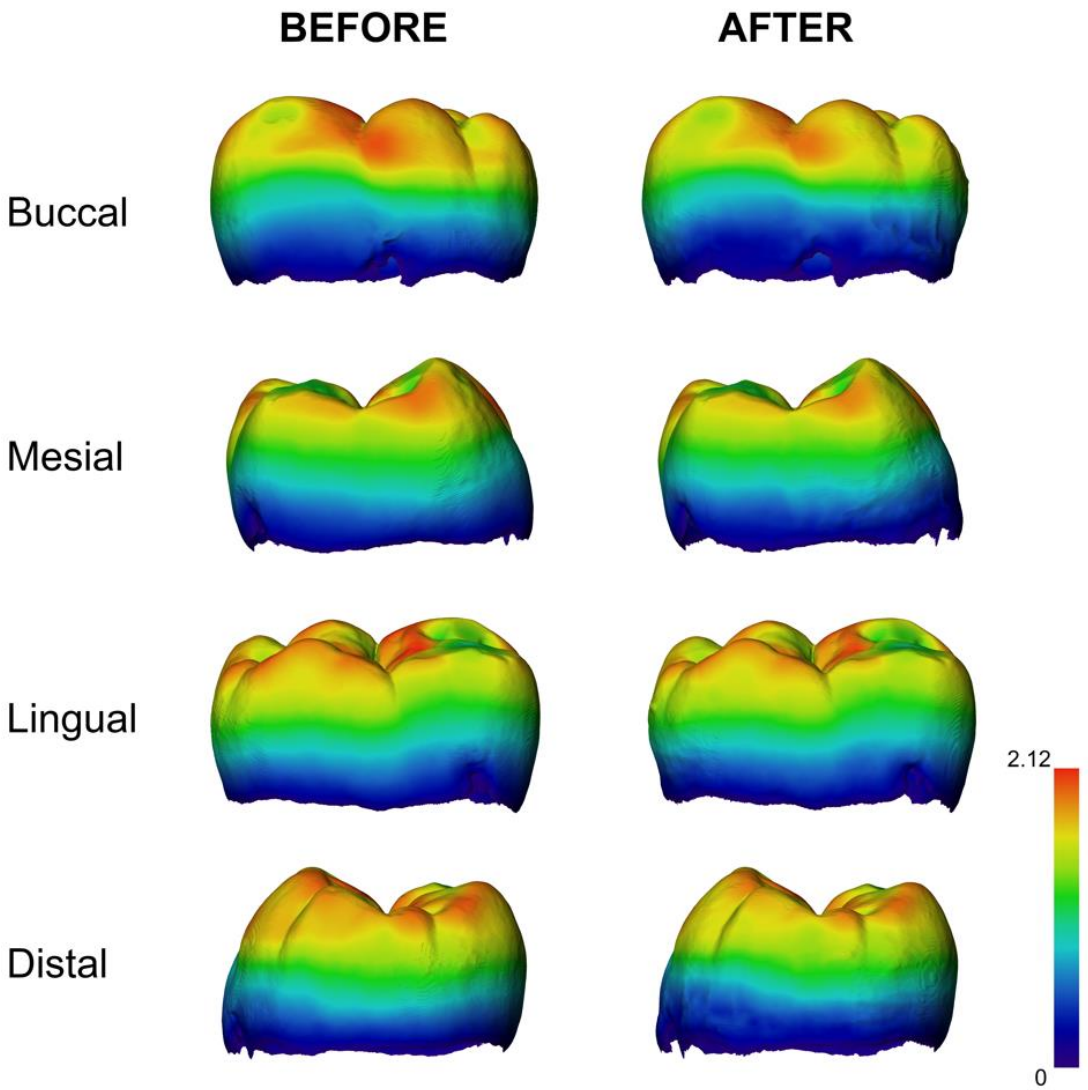


Fig. 5. Colour map of the tooth showing the enamel thickness before and after protein etching. Red indicates areas of thicker enamel surface. Tooth sample number 7.

Table 4. Micro-CT of selected teeth before and after protein etching. Absolute and relative values of tooth volume, enamel and dentine volume.

Sample number*	Tooth class	Enamel volume (mm^3)			Dentin volume (mm^3)			Total volume (mm^3)		
		Before	After	Difference (%)	Before	After	Difference (%)	Before	After	Difference (%)
3	$_3$M	248.03	224.25	9.6	475.14	463.93	2.4	723.2	688.18	4.8
4	M_3	285.38	264.53	7.3	632.08	623.11	1.4	917.5	887.64	3.3
5	$_3$M	311.76	287.41	7.8	706.53	697.15	1.3	1018.3	984.56	3.3
6	^{3}M	191.38	167.6	12.4	499.1	485.5	2.7	690.5	653.1	5.4
7	$_3$M	290.16	270.04	6.9	738.75	728.22	1.4	1028.9	998.26	3.0

9	M^3	151.3	131.57	13.0	403.34	393.32	2.5	554.6	524.89	5.4
10	$_3M$	205.22	178.91	12.8	390.81	383.48	1.9	596.0	562.39	5.6
11	M_3	191.6	166.52	13.1	620.17	603.85	2.6	811.8	770.37	5.1
12	M_3	182.37	157.88	13.4	465.99	453.56	2.7	648.4	611.44	5.7
14	M_3	162.68	139.6	14.2	389.2	372.79	4.2	551.9	512.39	7.2
31	^3M	155.17	139.73	10.0	480.56	473.3	1.5	635.7	613.03	3.6
32	$_3M$	177.24	160	9.7	476.13	471.85	0.9	653.4	631.85	3.3
42	M^1	196.09	180.22	8.1	728.44	720.29	1.1	924.5	900.51	2.6
46	M^3	155.55	141.07	9.3	531.76	516.31	2.9	687.3	657.38	4.4
47	M^1	277.28	260.76	6.0	936.66	928.53	0.9	1213.9	1189.29	2.0
48	^2M	237.48	220.53	7.1	553.83	547.77	1.1	791.3	768.3	2.9
49	^2M	185.94	166.71	10.3	688.07	675.21	1.9	874.0	841.92	3.7
53	^1M	171.39	148.34	13.4	798.43	786.98	1.4	969.8	935.32	3.6
54	M^1	100.7	91.35	9.3.	1005.0	984.73	2.0	1105.7	1076.08	2.7
59	^2M	237.44	218.2	8.1	1197.51	1178.42	1.6	1435.0	1396.62	2.7
	Mean			10.10			1.92			4.02
	Standard deviation			2.59			0.84			1.36
	95% confidence interval for the mean			[8.88, 11.30]			[1.53, 2.31]			[3.38, 4.65]

*Numbers correspond to the sample number in Table 2

Tooth nomenclature: I = incisor, C = canine, P = premolar, M = molar, upper index = upper jaw, lower index = lower jaw, side of the index indicates side of the tooth (left or right)

As can be seen in Fig. 6, the percentage of total volume loss depends on the proportion of enamel in the total volume of the tooth (this proportion is calculated as enamel volume before etching / total volume before etching). This relationship is of moderate intensity ($r = 0.46$) and weakly significant ($p = 0.042$).

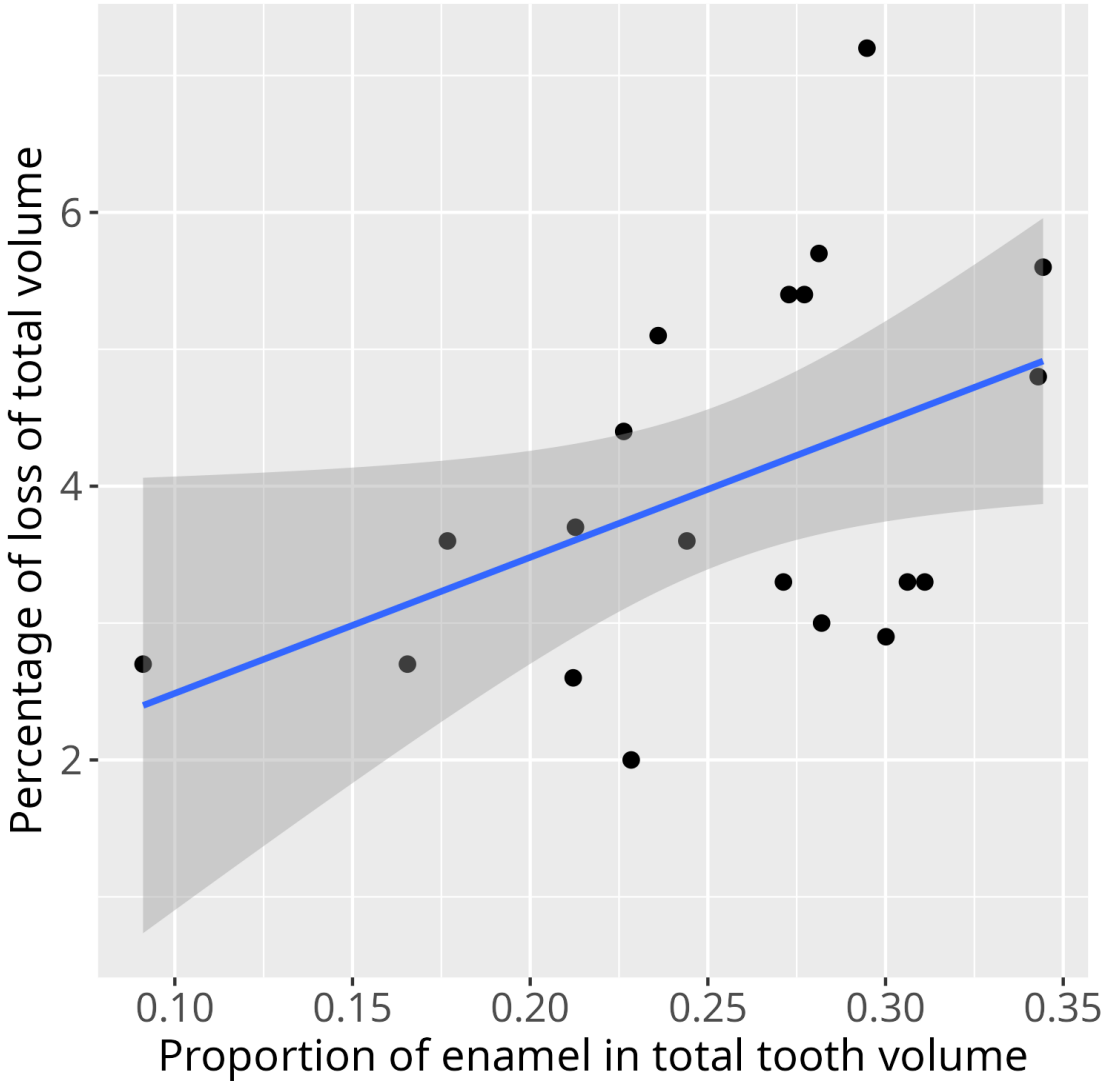


Fig. 6. Relationship between percentage of loss of total volume and the proportion of enamel in total tooth volume.

5. Discussion

The validity of proteomic sex estimation

Proteomics is an approach with a wide range of applications. The main strength of tandem mass spectrometry (MS/MS) is its ability to analyse complex protein mixtures [42]. The use of the resulting MS/MS spectra to determine the sequence of peptides is increasingly common not only in biological sex estimation [27,29,30,32], but also in art history [43–45] and the material analysis of historical textiles [46,47]. The use of proteomic methods in palaeontology is also increasing rapidly , and it is expected that these methods will be helpful for many general applications, connecting molecular biology, palaeontology, archaeology, palaeoecology, and history [42].

The advantage of proteomics in sex estimation is its high reliability and, compared to DNA analysis, the low risk of contamination and relative cost-effectiveness [17]; it is also less destructive. The absolute accuracy of the method used for the extraction and analysis of sexspecific proteins (Tab. 2) proved the suitability of the method for examining human skeletal remains, which, like cultural artefacts, form an

integral part of the cultural heritage. Reliable estimation of sex in sub-adults (which is impossible to assess directly from skeletal remains due to the undevelopment of sexually dimorphic traits), so far uncommon in bioarchaeological practice, significantly increases the explanatory potential of sub-adult skeletons and helps to avoid interpretation bias resulting from unknown sex ratios in specific age classes or socioeconomic groups [41]. Therefore, if sex estimation is needed, it is advisable to use protein extraction methods [13,32] without the need for tooth destruction [26,30]. Other reasons for its use include the low impact of manipulation on the archaeological skeletal material, which manifests itself in minimal invasiveness, as also proven in our results.

In addition, the technique employed allows the first two steps (sample preparation and protein etching) to be carried out in a laboratory with minimal instrumentation and equipment. Only the third and last stage, the protein analysis itself, must take place at a workplace with the appropriate liquid chromatography-tandem mass spectrometry equipment.

Invasiveness of the amelogenin protein etching protocol

The issue of the invasiveness of protein etching is a major topic in terms of cultural heritage preservation efforts. The influence of the various chemicals to which the hard dental tissue, in the case of proteomics the enamel, is exposed should be known. However, we are not aware of any previous study evaluating the extent of enamel loss when applying proteomics for sex estimation; to the best of our knowledge, the effect of protein extraction on the structural integrity of enamel assessed here for the first time. The protocol used in the present study uses low concentrations of H2O2 (3%) for 30 seconds for cleaning and demineralization of the tooth surface to remove calcium phosphate salts (calculus), and HCl (5%) etching for only 2 minutes [13] with immersion of only the dental crown. Such a concentration of H2O2 is even lower than the maximum concentration (up to 6% H2O2) in tooth whitening products, which a 2011 European directive considered to be cosmetic products according to Dias et al. [48]. With regard to HCl concentration and time to its exposition, the only possibility for comparison is offered by studies that deal with the effect of acids on teeth dissolution in a forensic context. The destructive effect of highly concentrated HCl (37%) on human dentition is indisputable. Several studies have established that highly concentrated HCl (depending, among other things, on the type of tooth) cause teeth to completely dissolve after a few hours [49–51]. Concerning the exposure time, Mazza et al. have reported no visible effect after 5 min of immersion in 37% HCl [49]. Gupta and Johnson observed morphological and radiographic changes to enamel after 30mins – 1 hour [52]. Jones et al. 2020 reported almost complete disintegration of enamel after 4 hours and no enamel present after 12 hours in HCl (37%) [51]. It has been confirmed that the acid concentrations as well as the length of exposure is important for the final morphological and chemical impact [49,50], and left no traces behind. In our case, a much lower concentration and a much shorter exposure time were used, which leave no visible traces. We observed enamel loss of 10% and dentine loss of only 2% using micro-computed tomography (micro-CT) (see Tab. 4). However, it should be noted that this also involves the removal of surface dirt and dental calculus, which are included in the volume of dental tissue in the first scanning and are removed only before the protein etching itself.

Importance for sexing in archaeological assemblages

The application of proteomic sex estimation in an archaeological collection to an individual with questionable or contradictory results of biological sex estimation by anthropological methods allows the obtaining of correct information about biological sex, which can be used to study a whole range of issues such as demography, diet and burial rites [15,53]. In the present study, adult individuals, for whom the morphological sex assessments performed by two teams half a century apart showed differences, were selected for the application sample. Table 5 shows the results of proteomic sex

estimation in 15 adults from the Early Medieval period with biological sex estimate discrepancy. The proteomic analysis agrees with the first morphological sex estimate by Stloukal in approximately half of the cases [36–38], while the same ratio of morphological sex agreement occurs in the case of the second morphological sex estimate by Zazvonilová et al. [39]. Sexing by proteomic analysis was successful for all individuals, contrary to the morphological sex estimates where some individuals were not determined or where sex was estimated with uncertainty (in Table 5 indicated by question marks). Unlike morphological sex estimation, which provides a probability of the estimated sex, in proteomic analysis sex can be assigned.

Table 5. Contradictory morphological sex estimation in an application sample of 15 adult teeth performed by Stloukal (1963, 1964, 1967) and by Zazvonilová et al. (2020), in comparison with amelogenin-sex estimation. Concordance of the anthropological estimation with the proteomic sex estimation is indicated by an asterisk.

Skeleton number	Anthropological sex estimation		Proteomic sex estimation
	Stloukal (1963, 1964, 1967)	Zazvonilová et.al. (2020)	Present study
H87	F?	M*	M
H170	M*	F	M
H0171	M	F*	F
H0187	F?	M?*	M
H0292	M	F*	F
H0314(bis)	Not determined	M	F
H0314	Not determined	M*	M
H0324	M	F*	F
H0352	F*	M	F
H0363	M*	F	M
H0406	F*	M	F
H0457	F?*	M	F
H0647	M*	F	M
H0718	M?	F*	F
H1088	F?	M*	M

Regarding the non-adult individuals described in detail elsewhere [41], the current sample of sexed non-adults with known dietary history helped to reveal that there were no dietary differences between Great Moravian boys and girls during the first decades of their lives. Proteomics provides a new, relatively simple, and quite inexpensive method of sex estimation without the risk of contamination.

6. Conclusion

Our results demonstrated the suitability of a protocol that uses protein etching from an intact tooth and does not require its mechanical destruction, or that of parts of the crown. Results show the absolute accuracy of biological sex estimation in a sample of 60 individuals of known sex. Based on

micro-CT analysis (n=20), the etching protein process is minimally invasive and does not cause visible changes to the dental enamel surface. Application to an archaeological sample of non-adult (n=32) and adult (n=15) individuals proved the suitability of proteomic biological sex estimation. Proteomics provides a convenient way to estimate sex in juveniles where anthropological methods cannot be applied. The status of such a procedure allows for wider dissemination of the method, which can thus become a routine technique in the analysis of archaeological skeletal material.

Acknowledgements

The research was carried out with the financial support of the Laboratory of 3D Imaging and Analytical Methods, Department of Anthropology and Human Genetics, Faculty of Science of Charles University in Prague, Ministry of Culture of the Czech Republic (DKRVO 2019-2023/7.I.e, 00023272). This research also benefited from the scientific framework of the University of Bordeaux's Idex "Investments for the Future" program / GPR "Human Past". The authors would like to thank Jiřina Dašková of the Palaeontology Department of the National Museum, Prague for obtaining the scanning electron microscope (SEM) images. We also thank Nicolas Vanderesse and Adrien Thibaut (UMR 5199 PACEA) for help with micro-CT, tooth tissue segmentation and the generation of colour maps. Our thanks also go to Mrs. Armenuhi Kirakosjan of the Institute of Physiology, Czech Academy of Sciences in Prague for her help with the proteomic analysis, and Alastair Millar for linguistic revision of the text.

References

[1] M.G. Belcastro, A. Pietrobelli, T. Nicolosi, M. Milella, V. Mariotti, Scientific and Ethical Aspects of Identified Skeletal Series: The Case of the Documented Human Osteological Collections of the University of Bologna (Northern Italy), Forensic Sci. 2 (2022) 349–361. https://doi.org/10.3390/forensicsci2020025.

[2] C. Knüsel, B. Maureille, Archaeological approaches to human remains: France, in: B. O'Donnabhain, B. Lozada (Eds.), Archaeol. Hum. Remain., Springer, Cham, 2018: pp. 57–80. https://doi.org/doi:10.1007/978-3-319-89984-8_5.

[3] A. Santos, Skulls and skeletons from documented, overseas and archaeological excavations: Portuguese trajectories, in: B. O'Donnabhain, B. Lozada (Eds.), Archaeol. Hum. Remain., Springer, Cham, 2018: pp. 111–125.

[4] M. Licata, A. Bonsignore, R. Boano, F. Monza, E. Fulcheri, R. Ciliberti, Study, conservation and exhibition of human remains: the need of a bioethical perspective, Acta Bio Medica Atenei Parm. (2020) 91.

[5] M.G. Belcastro, V. Mariotti, The place of human remains in the frame of cultural heritage: The restitution of medieval skeletons from a Jewish cemetery, J. Cult. Herit. 49 (2021) 229–238.

[6] K. Squires, D. Piombino-Mascali, Ethical Considerations Associated with the Display and Analysis of Juvenile Mummies from the Capuchin Catacombs of Palermo, Sicily, Public Archaeol. 20 (2021) 66–84. https://doi.org/10.1080/14655187.2021.2024742.

[7] D.G. Jones, Anatomists' uses of human skeletons: Ethical issues associated with the India bone trade and anonymized archival collections, Anat. Sci. Educ. 16 (2023) 610– 617. https://doi.org/https://doi.org/10.1002/ase.2280.

[8] G. Grupe, J. Wahl, Changing Perceptions of Archaeological Human Remains in Germany, in: B. O'Donnabhain, B. Lozada (Eds.), Archaeol. Hum. Remain., Springer, Cham, 2018: pp. 81–92.

[9] M. Sponheimer, C.M. Ryder, H. Fewlass, E.K. Smith, W.J. Pestle, S. Talamo, Saving Old Bones: a non-destructive method for bone collagen prescreening, Sci. Rep. 9 (2019) 13928. https://doi.org/10.1038/s41598-019-50443-2.

[10] R.M. Austin, S.B. Sholts, L. Williams, L. Kistler, C. Hofman, To curate the molecular past, museums need a carefully considered set of best practices, PNAS. 116 (2019) 1471–1474. https://doi.org/doi.org/10.1073/pnas.1822038116.

[11] A.H. Pálsdóttir, A. Bläuer, E. Rannamäe, S. Boessenkool, J.H. Hallsson, Not a limitless resource: ethics and guidelines for destructive sampling of archaeofaunal remains, R. Soc. Open Sci. 6 (2019) 191059. https://doi.org/10.1098/rsos.191059.

[12] M. Buckley, Proteomics in the Analysis of Forensic, Archaeological, and Paleontological Bone, in: Appl. Forensic Proteomics Protein Identif. Profiling, American Chemical Society, 2019: pp. 125-141 SE–8. https://doi.org/doi:10.1021/bk2019-1339.ch008.

[13] N.A. Stewart, R.F. Gerlach, R.L. Gowland, K.J. Gron, J. Montgomery, Sex determination of human remains from peptides in tooth enamel, Proc. Natl. Acad. Sci. 114 (2017) 13649–13654.

[14] F. Bray, S. Flament, G. Abrams, D. Bonjean, K.D. Modica, C. Rolando, C. Tokarski, P. Auguste, Extinct species identification from Upper Pleistocene bone fragments not identifiable from their osteomorphological studies by proteomics analysis, BioRxiv. (2020). https://doi.org/10.1101/2020.10.06.328021.

[15] K. Rebay-Salisbury, P. Bortel, L. Janker, M. Bas, D. Pany-Kucera, R.B. Salisbury, C. Gerner, F. Kanz, Gendered burial practices of early Bronze Age children align with peptide-based sex identification: A case study from Franzhausen I, Austria, J. Archaeol. Sci. 139 (2022) 105549. https://doi.org/https://doi.org/10.1016/j.jas.2022.105549.

[16] I. Mikšík, M. Morvan, J. Brůžek, Peptide analysis of tooth enamel – A sex estimation tool for archaeological, anthropological, or forensic research, J. Sep. Sci. 46 (2023) 2300183. https://doi.org/https://doi.org/10.1002/jssc.202300183.

[17] T. Buonasera, J. Eerkens, A. de Flamingh, L. Engbring, J. Yip, H. Li, R. Haas, D. DiGiuseppe, D. Grant, M. Salemi, C. Nijmeh, M. Arellano, A. Leventhal, B. Phinney, B.F. Byrd, R.S. Malhi, G. Parker, A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation, Sci. Reports. 10 (2020) 1–15. https://doi.org/10.1038/s41598-020-68550-w.

[18] A.R. Klales, Sex Estimation of the Human Skeleton, Academic Press, 2020.

[19] A. Kotěrová, J. Velemínská, J. Dupej, H. Brzobohatá, A. Pilný, J. Brůžek, Disregarding population specificity: its influence on the sex assessment methods from the tibia, Int. J. Legal Med. 131 (2016) 251–261.

[20] E.K. Oikonomopoulou, E. Valakos, E. Nikita, Population-specificity of sexual dimorphism in cranial and pelvic traits: evaluation of existing and proposal of new functions for sex assessment in a Greek assemblage, Int. J. Legal Med. 131 (2017) 1731–1738.

[21] J. Brůžek, P. Murail, Methodology and reliability of sex determination from the skeleton, in: A. Schmitt, E. Cunha, J. Pinheiro (Eds.), Forensic Anthropol. Med. Complement. Sci. From Recover. to Cause Death, Humana Press Inc., Totowa, 2006: pp. 225–242.

[22] T. Waldron, The relative Survival of the Human Skeleton:Implications for Palaeopathology, in: A. Boddington, A. Garland, R. Janaway (Eds.), Death, Decay Recon-Struction Approaches to Archaeol. Forensic Sci., Manchester University Press, 1987: pp. 55–64.

[23] L.K. Corron, F. Santos, P. Adalian, K. Chaumoitre, P. Guyomarc'h, F. Marchal, J. Brůžek, How low can we go? A skeletal maturity threshold for probabilistic visual sex estimation from immature human os coxae, Forensic Sci. Int. 325 (2021) 110854. https://doi.org/https://doi.org/10.1016/j.forsciint.2021.110854.

[24] M.E. Lewis, The Bioarchaeology of Children: Perspectives from Biological and Forensic Anthropology, Cambridge University Press, Cambridge, 2006. https://doi.org/DOI: 10.1017/CBO9780511542473.

[25] M. Jágr, A. Eckhardt, S. Pataridis, Z. Broukal, J. Duskova, I. Miksik, Proteomics of human teeth and saliva, Physiol. Res. 63 (2014) S141.

[26] F. Lugli, G. Di Rocco, A. Vazzana, F. Genovese, D. Pinetti, E. Cilli, M.C. Carile, S. Silvestrini, G. Gabanini, S. Arrighi, L. Buti, E. Bortolini, A. Cipriani, C. Figus, G. Marciani, G. Oxilia, M. Romandini, R. Sorrentino, M. Sola, S. Benazzi, Enamel peptides reveal the sex of the Late Antique 'Lovers of Modena,' Sci. Rep. 9 (2019) 13130. https://doi.org/10.1038/s41598-019-49562-7.

[27] F. Lugli, C. Figus, S. Silvestrini, V. Costa, E. Bortolini, S. Conti, B. Peripoli, A. Nava, A. Sperduti, L. Lamanna, L. Bondioli, S. Benazzi, Sex-related morbidity and mortality in non-adult individuals from the Early Medieval site of Valdaro (Italy): the contribution of dental enamel peptide analysis, J. Archaeol. Sci. Reports. 34 (2020) 102625. https://doi.org/https://doi.org/10.1016/j.jasrep.2020.102625.

[28] 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, D.M. Rocke, B.S. Phinney, Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel, J. Archaeol. Sci. 101 (2019) 169–180. https://doi.org/https://doi.org/10.1016/j.jas.2018.08.011.

[29] C. Froment, M. Hourset, N. Sáenz-Oyhéréguy, E. Mouton-Barbosa, C. Willmann, C. Zanolli, R. Esclassan, R. Donat, C. Thèves, O. Burlet-Schiltz, C. Mollereau, Analysis of 5000 year-old human teeth using optimized large-scale and targeted proteomics approaches for detection of sex-specific peptides, J. Proteomics. 211 (2020) 103548. https://doi.org/https://doi.org/10.1016/j.jprot.2019.103548.

[30] A. Gasparini, F. Lugli, S. Silvestrini, A. Pietrobelli, I. Marchetta, S. Benazzi, M.G. Belcastro, Biological sex VS. Archaeological Gender: Enamel peptide analysis of the horsemen of the Early Middle age necropolises of Campochiaro (Molise, Italy), J. Archaeol. Sci. Reports. 41 (2022) 103337. https://doi.org/https://doi.org/10.1016/j.jasrep.2021.103337.

[31] M. Fonović, T. Leskovar, I. Štamfelj, Determination of Sex Based on Sexually Dimorphic Amelogenin Peptides in Human Tooth Enamel, J. Crim. Investig. Criminol. 72 (2021) 2.

[32] R. Gowland, N.A. Stewart, K.D. Crowder, C. Hodson, H. Shaw, K.J. Gron, J. Montgomery, Sex estimation of teeth at different developmental stages using dimorphic enamel peptide analysis, Am. J. Phys. Anthropol. 174 (2021) 859–869. https://doi.org/https://doi.org/10.1002/ajpa.24231.

[33] P. Pachner, Pohlavní rozdíly na lidské pánvi (Sexual Differences in Human Pelvis), Česká akademie věd a umění, Praha, 1937.

[34] L. Borovanský, Pohlavní rozdíly na lebce člověka (Sexual differences in human skulls), Česká akademie věd a umění, Praha, 1936.

[35] L. Poláček, Great Moravian Elites From Mikulčice, Czech Academy of Sciences, Institute of Archaeology, Brno, 2021.

[36] M. Stloukal, První pohřebiště na hradišti „Valy “u Mikulčic, Památky Archeol. 54. (1963) 114–140.

[37] M. Stloukal, Čtvrté pohřebiště na hradišti „Valy “u Mikulčic, Památky Archeol. 55. (1964) 479–505.

[38] M. Stloukal, Druhé pohřebiště na hradišti „Valy “u Mikulčic, Památky Archeol. 58. (1967) 272–319.

[39] E. Zazvonilová, P. Velemínský, J. Brůžek, Paleodemografická interpretace kosterních souborů minulých populací: nové hodnocení raně středověkých pohřebišť u 3. a 6. kostela v Mikulčicích (Palaeodemographic interpretation of skeletal assemblages of past populations: a new evaluation of early mediev, Archeol. Rozhl. 72 (2020) 67–101.

[40] D. Ferembach, I. Schwidetzky, M. Stloukal, Recommendations for age and sex diagnoses of skeletons, J. Hum. Evol. 9 (1980) 517–549.

[41] S.D. Drtikolová 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 (9th-10th centuries AD, Czechia), Archaeol. Anthropol. Sci. Preprint (2022).

[42] C. Warinner, K. Korzow Richter, M.J. Collins, Paleoproteomics, Chem. Rev. 122 (2022) 13401–13446. https://doi.org/10.1021/acs.chemrev.1c00703.

[43] A. Lluveras-Tenorio, R. Vinciguerra, E. Galano, C. Blaensdorf, E. Emmerling, M. Perla Colombini, L. Birolo, I. Bonaduce, GC/MS and proteomics to unravel the painting history of the lost Giant Buddhas of Bāmiyān (Afghanistan), PLoS One. 12 (2017) e0172990. https://doi.org/10.1371/journal.pone.0172990.

[44] W. Fremout, S. Kuckova, M. Crhova, J. Sanyova, S. Saverwyns, R. Hynek, M. Kodicek, P. Vandenabeele, L. Moens, Classification of protein binders in artist's paints by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry: an evaluation of principal component analysis (PCA) and soft independent modelling of class analogy (SIMCA), Rapid Commun. Mass Spectrom. 25 (2011) 1631–1640. https://doi.org/https://doi.org/10.1002/rcm.5027.

[45] I.K. Levy, R. Neme Tauil, A. Rosso, M.P. Valacco, S. Moreno, F. Guzmán, G. Siracusano, M.S. Maier, Finding of muscle proteins in art samples from mid-18th century murals by LC–MSMS, J. Cult. Herit. 48 (2021) 227–235. https://doi.org/https://doi.org/10.1016/j.culher.2020.11.005.

[46] A.K. Popowich, T.P. Cleland, C. Solazzo, Characterization of membrane metal threads by proteomics and analysis of a 14th c. thread from an Italian textile, J. Cult. Herit. 33 (2018) 10–17. https://doi.org/https://doi.org/10.1016/j.culher.2018.03.007.

[47] C. Solazzo, S. Clerens, J.E. Plowman, J. Wilson, E.E. Peacock, J.M. Dyer, Application of redox proteomics to the study of oxidative degradation products in archaeological wool, J. Cult. Herit. 16 (2015) 896–903. https://doi.org/https://doi.org/10.1016/j.culher.2015.02.006.

[48] S. Dias, A. Mata, J. Silveira, R. Pereira, A. Putignano, G. Orsini, R. Monterubbianesi, D. Marques, Hydrogen Peroxide Release Kinetics of Four Tooth Whitening Products— In Vitro Study, Materials (Basel). 14 (2021). https://doi.org/10.3390/ma14247597.

[49] A. Mazza, G. Merlati, C. Savio, G. Fassina, P. Menghini, P. Danesino, Observations on dental structures when placed in contact with acids: Experimental studies to aid identification processes, J. Forensic Sci. 50 (2005) JFS2004292-5.

[50] M. Raj, K. Boaz, N. Srikant, Are teeth evidence in acid environment, Forensic Dent. Sci. 5 (2013) 7.

[51] C. Jones, T. Bracewell, A. Torabi, C.C. Beck, T.B. Harvey, The effect of hydrochloric acid (HCl) on permanent molars: A scanning electron microscope (SEM) and energy dispersive X-ray spectroscopy (EDS) study, Med. Sci. Law. 60 (2020) 172–181. https://doi.org/10.1177/0025802420905981.

[52] K.K. Gupta, A. Johnson, Morphologic and radiographic effects of acids on the teeth: An in-vitro forensic study, Indian J. Forensic Med. Toxicol. 14 (2020) 28–33.

[53] C. McFadden, The past, present and future of skeletal analysis in palaeodemography, Philos. Trans. R. Soc. B Biol. Sci. 376 (2020) 20190709. https://doi.org/10.1098/rstb.2019.0709.

Article

Early life histories of Great Moravian children – carbon and nitrogen isotopic analysis of dentine serial sections from the Early Medieval population of Mikulčice (9th-10th centuries AD, Czechia)

Sylva Drtikolová Kaupová[1], Jaroslav Brůžek[2], Jiří Hadrava[2], Ivan Mikšík[3], Marine Morvan[3,4], Lumír Poláček[5], Lenka Půtová[1], Petr Velemínský[1]

[1] Department of Anthropology, National Museum, Václavské náměstí 1700/68, 110 00 Prague 1, Czech Republic

[2] Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, 128 00 Prague 2, Czech Republic

[3] Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 532 10 Pardubice, Czech Republic

[4] Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20 Prague 4, Czech Republic

[5] Institute of Archaeology of the Czech Academy of Sciences, Letenská 4, 118 00 Prague 1, Czech Republic

Corresponding author: sylva.kaupova@nm.cz (S.D.K.)

Preprint citation: Drtikolová Kaupová, S; Brůžek, J.; Hadrava, J.; Mikšík, I.; Morvan, M.; Poláček, L.; Půtová, L.; Velemínský, P., Early life histories of Great Moravian children – carbon and nitrogen isotopic analysis of dentine serial sections from the Early Medieval population of Mikulčice (9th-10th centuries AD, Czechia), *Archaeological and Anthropological Sciences*, **2022**, https://doi.org/10.21203/rs.3.rs-1913554/v1

Abstract: In order to compare the early life experiences of different population subgroups from the Early Medieval centre of Mikulčice, carbon and nitrogen isotopic values were measured in dentine serial sections from the first permanent molar of 78 individuals. Age-at-death, sex (estimated in subadults with the help of proteomics) and socio-economic status were considered as explicative variables. Average values of both nitrogen and carbon maximal isotopic offset within the isotopic profile were higher than the recommended range for weaning under healthy circumstances: 3.1 ± 0.8‰ for Δ15Nmax and 1.6 ± 0.8‰ for Δ13Cmax. Individuals who died during the first decade of life showed earlier ages at the final smoothing of the nitrogen isotopic curve (suggesting complete weaning) than older individuals. Most individuals (n = 43) showed positive covariance between δ15N and δ13C values during the period of breastfeeding. The average δ15N values from the post-weaning period were similar to those of bone, while post-weaning δ13C values were significantly higher.

Though an increased Δ15Nmax suggests a common presence of physiological stress, the intra-population comparison of early life experiences does not suggest that individuals who died during their first decade experienced greater levels of environmental stress during infancy.

The predominance of positive covariance between carbon and nitrogen isotopic values during the breastfeeding period, together with an increased Δ13Cmax and increased post-weaning δ13C, suggest

that millet was either a part of a special diet preferred during lactation or was introduced as a first dietary supplement.

Keywords: Breastfeeding, physiological stress, diet, stable isotopes, Middle Ages

1. Introduction

Since the pioneering work by Fogel et al. (1989) described the relationship between the hair isotopic values of mothers and their breastfed babies, the reconstruction of breastfeeding and weaning behaviour has naturally attracted the attention of bioarchaeologists. The duration of breastfeeding, as well as the timing of the introduction of complementary food, clearly affect the health and physical well-being of the child both in the short term (Lamberti et al. 2011, Shamir 2016, Wilson et al. 2006) and the long term (Berti et al. 2017, Demmelmair et al. 2006, Kendall et al. 2021, Lamberti et al. 2011, McDade 2005, Palou and Picó 2009). Moreover, as a bio-socio-cultural phenomenon, infant and young child feeding practices are influenced by a number of cultural, religious, economic and also environmental factors (Fildes 2017, Quinlan 2007, Thorvaldsen 2008, Tomori et al. 2017, Yovsi and Keller 2003).

Last but not least, the duration of breastfeeding has a considerable impact on women's health and fertility (Bentley et al. 2001, Jay 2009). For all these reasons, information on this aspect of childcare helps substantially to understand the population dynamics and living conditions of past populations.

For more than a decade, researchers generally followed a cross-sectional approach, analysing carbon and nitrogen isotopes in the bone collagen of infants and young children of various ages, and comparing their values to female population averages (e.g. Fogel et al. 1989, Katzenberg et al. 1996, Mays 2010, Pearson et al. 2010, Prowse et al. 2008, Schurr 1997), often attempting to link observed isotopic patterns with the mortality and morbidity profiles of subadults.

Subsequently, however, some hidden pitfalls of this approach were addressed, such as the omission of the intra-population dietary variation in lactating mothers or potential mortality biases in weaning interpretations in terms of the osteological paradox (Fuller et al. 2003, Wood et al. 1992). A solution to these problematic issues was seen in the enforcement of an intra-individual approach, recovering dietary information from different periods of an individual's life. Samples were taken from different mineralized tissues (Herrscher 2013, Howcroft et al. 2012, Kaupová et al. 2014), from different parts of bone (Waters-Rist et al. 2011), and finally from serial sections of dental tissues, with the number of serial samples increasing along with development in mass spectrometry reducing the sample size (Beaumont et al. 2013, Eerkens et al. 2011, Fuller et al. 2003, Howcroft et al. 2012).

Gradually, scientists' attention moved to dentine tissue, which, due to the absence of turnover, retains the isotopic signal from the period of tooth development throughout life (Balasse et al. 2001, Richards et al. 2002), and thus allows the inclusion of adult individuals into studies while avoiding the risk of mortality biases. Sampling horizontal sections of dentine results in an isotopic profile covering the entire period of the tooth development (Beaumont et al. 2013). However, rather than resolving the issue of breastfeeding, the more detailed sampling methodology revealed that the dentine isotopic record results from a complex interaction of a number of dietary and non-dietary factors, among which physiological stress plays a key role (Beaumont et al. 2013, 2018, Craig-Atkins et al. 2018, King et al. 2018).

Potential sources of isotopic variation in dentine isotopic values

Breastfeeding

The isotopic effect of breastfeeding is a reflection of a phenomenon called the “trophic level effect” (Ambrose and Norr 1993, DeNiro and Epstein 1978, 1981, Schoeninger and DeNiro 1984), whereby heavier isotopes are discriminated against at each level of the food chain, resulting in isotopic enrichment of the consumers' tissues above those of their prey. In bone collagen, enrichment by 3-5 ‰ was observed with each trophic level for δ15N values (direct estimates for human give slightly higher estimates of 5.5-6 ‰, O’Connell et al. 2012) and by 1 ‰ for δ13C values. However, focusing on breastfeeding, direct observation of mother-infant pairs reported somewhat lower shifts in δ15N values associated with breastfeeding – between 2-3 ‰ (Fuller et al. 2006, Herrscher et al. 2017).

Prior to birth, the foetus is an integral part of the maternal organism (Fogel et al. 1989). Although a small offset has been observed in the tissues formed in-utero in mother-offspring pairs (de Luca et al. 2012, but see Herrscher et al. 2017), much larger deviations occur once breastfeeding starts. Due to the trophic level effect, breastfed infants exhibit an elevation of both the δ13C and δ15N values above maternal values. During weaning, both δ13C and δ15N isotopic values of the newly formed tissues drops continuously, along with the decreasing dietary proportion of breast milk (Fogel et al. 1989, Fuller et al. 2006).

Nutritional or physiological stress

As demonstrated by a number of controlled-feeding experiments on different animal species (for review see e.g. Reitsema 2013), as well as by direct observations of humans with eating disorders, pregnancy complications, and a number of serious diseases (Fuller et al. 2005, Mekota et al. 2006, Tea et al. 2021), changes in nitrogen balance under physiological stress are responsible for notable isotopic shifts. To compensate for protein insufficiency under conditions of negative nitrogen balance, body tissues are catabolized and recycled. Thus, the fractionating processes of transamination and deamination are repeated, resulting in enrichment in the 15N of the newly formed tissues (Reitsema 2013). The impact of decreased protein bioavailability on carbon isotopic values is, however, less unequivocal. In some studies, the above-described shifts in δ15N values have been found to be accompanied by decreased δ13C values, this results from the combination of the altered ratio between routed dietary and endogenously synthesized amino acids and of the mobilization of fat stores, which are isotopically depleted (Beaumont et al. 2018, Mekota et al. 2006, Neuberger et al. 2013, Schwarcz 2002). Based on these findings, a typical isotopic "stress pattern" has been described, which is commonly used especially in studies of dentine isotopic profiles, stressing the opposing covariance between δ15N and δ13C values during the stress episode (Beaumont et al. 2018, Beaumont and Montgomery 2016, Craig-Atkins et al. 2018).

An increasing number of studies (Canterbury et al. 2020, D’Ortenzio et al. 2015, Drtikolová Kaupová et al. 2021, Fuller et al. 2005, Katzenberg and Lovell 1999), however, have found no change in δ13C values during the stress episodes affecting δ15N values. There are even studies (Drtikolová Kaupová et al. in press, Neuberger et al. 2013) reporting positive covariance between δ13C and δ15N values during such stress episodes.

Other factors

The next point potentially deconstructing the paradigm of the "typical isotopic stress pattern" is that there are a number of other factors of both dietary and physiological origin, concerning mother and/or child, which may evoke isotopic shifts emulating or overriding isotopic reflection of both breastfeeding and stress. The use of distinct weaning foods is quite common throughout the world (e.g. Sellen 2001), as are specific dietary rules imposed on pregnant or lactating women (Baumslag 1987). Alternatively, annual or seasonal variation in the isotopic composition of the mothers' diets, and therefore of breastmilk, can never be excluded. This argument is especially important in all the contexts with documented consumption of C4 plants and/or both marine and freshwater products (King et al. 2018). Further, although there is limited information on the isotopic composition of breastmilk, it seems that there are isotopic shifts in breastmilk values over the course of breastfeeding independent of maternal diet or health (Herrscher et al. 2017), probably linked with changing macronutrient composition over the course of lactation (Czosnykowska-Łukacka et al. 2018). The practice of wet-nursing, well-known in continental Europe since Antiquity (Fildes 2017), can also affect the isotopic reflection of breastfeeding, when the isotopic values of the wet-nurse are different from those of the mother (Herrscher 2004).

For all the reasons mentioned above, it is evident that the reduction of isotopic variation in infants and young children to a simple typology of isotopic curves reflecting "weaning" or "physiological stress" (e.g. Craig-Atkins et al. 2018) is not viable. In this study, we thus attempt to avoid this descriptive approach and rather analyse the interplay between δ15N and δ13C isotope values in early life without such categorization, to compare the isotopic values and their shifts between biologically and socio-culturally defined groups and to highlight the multiple potential interpretations of incremental profile shapes.

For that reason, we have explored the early life experiences of a sufficiently numerous (n = 78) population sample from the Great Moravian settlement agglomeration of Mikulčice, defining a number of simple questions: first, do all the individuals exhibit a isotopic peak in the early life period attributable to breastfeeding? Second, due to the well-described impact of early life experience on health and physical well-being in a long-term perspective (Demmelmair et al. 2006, McDade 2005, Palou and Picó 2009), we searched for the potential differences in isotopic profiles between individuals who died during M1 formation, those who died after M1 formation but before adulthood (i.e. during the second decade of the life) and finally individuals who survived to adulthood. To check for different parental investment in infants of different sexes, which are common in populations with high gender inequality (Eerkens and Bartelink 2013, Jayachandran and Kuziemko 2011), we explored the potential differences in isotopic profile between males and females. And finally, we explored the whether the level of parental effort and/or environmental risk differed between members of Great Moravian elites vs. common folk.

2. Materials and methods

Great Moravia (9th to beginning of the 10th centuries AD) was the first Slavic proto-state structure in Central Europe (Figure 1). Along with rapid political consolidation and the introduction of the first proto-urban centers, Christianisation probably extended into Moravia at the beginning of the 9th century (Herold 2012, Kalhous 2020, Macháček 2013). The skeletal material analysed in this study

comes from the settlement agglomeration at Mikulčice (Czechia, N 48°48'15.9", E 17°05'08.5", Kuna et al. 2018), which is believed to be one of the prominent power centres of the Great Moravian Empire (Poláček 2018). It attained a degree of urbanization unprecedented in the region, with a high concentration of ecclesiastical buildings. The settlement complex consisted of a fortified acropolis or "castle" and a bailey, surrounded by unfortified suburbs. Ongoing archaeological research begun in the 1950s has uncovered more 2500 graves both in the suburbs and at the acropolis, including presumably dynastic graves in the interior of the main churches as well as a number of richly equipped graves, suggesting the presence of the true elites of Great Moravian society (Poláček 2008). The adult and peripherally the subadult diets of the Mikulčice population have previously been explored (Halffman and Velemínský 2015, Jílková et al. 2019, Kaupová et al. 2018), providing a good framework for comparison with current results and documenting the notable input of millet in the Great Moravian diet.

The dataset included 46 adults and 32 subadults, of which 25 died during M1 formation, while 7 died after the completion of the M1 root. There were 41 males and 37 females. In terms of socio-economic status, 38 individuals were classified as elite, the rest (n = 40) as non-elite. The age-at death distribution of the adult dataset was affected by dental wear, so individuals younger than 40 years were strongly prevalent. In adults, sex estimation was based primarily on the morphology (Brůžek 2002, Phenice 1969) and metrics of the innominate bones (Brůžek et al. 2017, Murail et al. 2005). Where these were absent or poorly present, an evaluation of the morphological traits of the skull was used (Walker 2008). Age-at death estimation was based on the evaluation of senescent changes of the auricular surface (Schmitt 2005), pubic symphysis (Schmitt 2008) and the acetabulum (Calce 2012). In young adults, indicators of skeletal maturation of iliac crests and clavicles were used (Scheuer et al. 2008). For subadults, sex was estimated with the help of sex chromosome-linked isoforms of the peptide amelogenin from human tooth enamel. The modification of the method by Stewart et al. (2017) was used, employing a minimally destructive acid etching procedure and subsequent nano liquid chromatography tandem mass spectrometry. Further details of the method and the sex estimates for the subadults can be found in the Online Resource 1. The reliable sex estimation in the subadults, still not common in bio-archaeological practice, greatly increases the testimonial power of the subadult skeletons and helps to avoid the interpretative biases resulting from unknown sex ratios in particular age classes or socio-economic groups. Age-at-death for subadults was assessed using dental development (Smith 1991).

The character of grave goods was used for the categorization of socio-economic status. Elites were considered to be individuals from well-equipped graves containing gold, luxury jewellery and textiles, belt strap-ends or cutlery, gilded buttons, metal weapons, spurs, or iron coffin accessories. Individuals from graves with objects of daily use, such as knives, ceramics, glass buttons or beads, simple jewellery or from graves lacking any grave goods, were considered to be non-elite.

The criterion for inclusion into the dataset was the preservation of at least one first permanent molar (M1) and a low degree of dental wear, admitting max. stage E according to Lovejoy (1985). Samples for stable isotope analysis were taken preferentially from the lower M1, but in cases of absence, damage (e.g. by dental caries) or notable dental wear, the upper M1 was sampled. After removing adhered dirt, the M1 was halved along the mesio-distal (lower M1) or vestibulo-palatinal (upper M1) axis, with half of the tooth preserved for further analyses. Enamel was removed from one half of the M1 using dental burs and saws. Collagen was extracted using Method 2 described by Beaumont et al. (2013) in the modification by van der Haas et al. (2018). The demineralized dentine was sectioned into 10 horizontal sections (in the case of fully formed teeth) reflecting dental developmental stages (Smith 1991). As far as possible the calculations of the approximate ages for each section considered the distinct rate of

dentine secretion at various stages of tooth development, being calculated separately for each particular tooth segment (Czermak et al. 2020): the crown (Cri-Crc, divided into 5 slices), the superior half of the root (Crc-R1/2, divided into 2 slices), the inferior half of the root (R1/2-A1/2, divided into 2 slices) and the closing apex (A1/2-Ac, one slice). According to the Smith (1991) developmental scheme, the M1 crown forms approximately from birth, with the age at crown completion being 2.2±0.5 years. The fully formed tooth crowns were cut in 5 horizontal increments, each representing 1/5th of 2.2 years. Analogously, calculations were made for the M1 root.

All the sample preparations were carried out at the Department of Anthropology, National Museum, Prague, CZ. EA-IMRS (Elemental Analysis – Isotope Ratio Mass Spectrometry) was performed at Iso-Analytical, Crewe, UK. Stable carbon and nitrogen isotopic compositions were calibrated relative to the VPDB and AIR scales using IAEA-CH-6 and IAEA-N-1 inter-laboratory comparison standards. Measurement uncertainty was monitored using in-house standards: IA-R068 (soy protein, δ13CV-PDB = –25.22 ‰, δ15NAIR = 0.99 ‰), IA-R038 (L-alanine, δ13CV-PDB = –24.99 ‰, δ15NAIR = –0.65 ‰), IA-R069 (tuna protein, δ13CV-PDB = –18.88 ‰, δ15NAIR = 11.60 ‰) and a mixture of IAEA-C7 (oxalic acid, δ13CV-PDB = –14.48 ‰) and IA-R046 (ammonium sulphate, δ15NAIR = 22.04 ‰). Precision was determined to be ± 0.14‰ for both δ13C and δ15N values based on repeated measurements of calibration standards, check standards, and sample replicates. Accuracy or systematic error was determined to be ± 0.07 for δ13C and ± 0.11 for δ15N values based on the difference between the observed and known δ values of the check standards and the long-term standard deviations of these check standards. The total analytical uncertainty as defined by Szpak et al. (2017) was estimated to be ± 0.16 ‰ for δ13C values and ± 0.18 for δ15N values.

To describe the carbon and nitrogen isotopic profiles with respect to dietary and physiological changes, several parameters were chosen: for δ15N values, we noted the presence of the initial peak in isotopic values as primary evidence for breastfeeding. We calculated the maximal isotopic offset (Δ15Nmax) to describe the height of this peak and noted the ages at i) the first notable decrease from the peak value and ii) the smoothing of the isotopic profile to describe its extension (Figure 2A). In both cases 0.4 ‰ (a double of the analytical error at two standard deviations) was considered to be a significant isotopic change. For δ13C values, we also calculated the maximal isotopic offset (Δ13Cmax). We then evaluated the shape of the carbon isotopic profile in relation to that of nitrogen during the initial peak described above. Four basic types of carbon isotopic profiles were defined: i: positive covariance between δ15N and δ13C values, ii: negative covariance when δ13C values are low during the time of peaking δ15N and subsequently increase along with decreasing δ15N values, iii: Flat carbon profile, and finally iv: positive covariance between δ15N and δ13C values in most of the slices, but with an initial decrease in δ13C values preceding the shift in δ15N values (Figure 2B).

To describe the isotopic values after the period of breastfeeding, we calculated the average post-weaning isotopic value (i.e. the average value from all the slices developed after smoothing the nitrogen isotopic curve as described above). This was to describe childhood diet in the period following weaning, and to compare it with previously published dietary info from later life periods (Jílková et al. 2019, Kaupová et al. 2018). For each individual, we also searched for two potential patterns described in previous isotopic studies of dentine incremental profiles: i) the presence of the so-called 'post-weaning dip' in nitrogen isotopic values, which is usually linked to the special character of the post-weaning diet, with low input from animal products (Tsutaya and Yoneda 2015), and ii) cases of notable shifts in both carbon and nitrogen isotopic values showing a pattern of a negative covariance, usually attributed to biological stress (Beaumont and Montgomery 2016).

Statistical analyses were examined in R software version 3.3.3 (R Core Team 2017). The effect of age, sex, and socio-economic status on the values of nitrogen and carbon post-weaning averages, as well as

age at first decrease and age at final smoothing, were modelled using linear models and tested with ANOVA. The values of age at first decrease and age at final smoothing, which were classified into 3 and 6 levels respectively (i.e. an ordinal scale), were treated as numerical variables in the analyses, because there were relatively high numbers of levels with only a few observations per level.

As maximal offset of nitrogen, maximal offset of carbon, age of weaning, and age at first decrease correlated with each other (with a maximal correlation coefficient of 0.31 between the the maximal offset of nitrogen and maximal offset of carbon), the joint effect of age, sex, survival, and socio-economic status was modeled with RDA analysis using the library ''vegan'' (Oksanen et al. 2015). The effect of the predictors was tested with an ANOVA-like permutation test with the function “anova.cca”. The incidence of distinct carbon isotopic profiles in particular age-groups was compared by Fisher exact test.

3. Results

Complete isotopic results are given in full in the Online Resource 2. Several increments (N = 6) did not yield enough collagen for analysis, this concerns exclusively apical sections including only a small portion of dentine from dentine horns. All the analysed samples met the criteria for good collagen preservation. Individuals analysed in this study exhibit highly varied δ15N and δ13C profiles, which are shown for each individual in the Online Resource 2.

Nitrogen isotopic profiles

Almost all the individuals show some sort of decrease in δ15N values during the earliest life period, which could be viewed as a reflection of breastfeeding and subsequent weaning. However, notable variation exists in both the timing and magnitude of these shifts. Based on the study by Fuller et al. (2006), an isotopic offset of 2-3 ‰ could be seen as primary evidence of weaning from full breastfeeding under healthy circumstances. Shifts beyond these margins (in both directions) suggest the combined impact of the trophic level effect and biological stress, and/or the other factors listed above. In this study, the average Δ15Nmax was slightly higher than the recommended range (3.1 ± 0.8 ‰). In 37 (of the 78) individuals, Δ15Nmax was above 3 ‰, with a maximum of 5.1 ‰. In nine individuals, Δ15Nmax was below 2 ‰ with six of them not showing a typical isotopic curve shape attributable to weaning: five individuals, (Graves 112-VI, 462, 673, 1058, 1171) show very low δ15N in the earliest dentine slice following by a delayed peak in δ15N (Figure 3). No. 207 shows a specific nitrogen isotopic profile with low δ15N in the earliest dentine slice followed by the two subsequent peaks in δ15N of smaller extent (Figure 3).

When comparing the Δ15N between the defined population subgroups, there were no statistically significant differences between age classes, sexes or socio-economic groups (Table 1, Figure 4). In the case of age, the result was close to the 0.05 % level of significance, with non-survivors showing surprisingly lower nitrogen isotopic offsets that older individuals.

Concerning the start and end points of the above described isotopic shift attributable to weaning, in most of the individuals (n = 37) δ15N values decreased significantly in the second slice, representing isotopic values at the age of 5-11 months. In 30 individuals the first decrease was observed in the third slice (representing isotopic values from approx. 11-16 months), while in five individuals, δ15N values did not decrease before 16-21 months (4th slice). In detailed comparison, females showed the first

isotopic decrease at a later age than males (Table 1). Age-at-death and socio-economic status had no statistically significant relationship to this parameter (Figure 4).

Table 1. Relationship between defined parameters of both carbon and nitrogen isotopic curve and the explicative variables (p – values[a])

	δ15N				δ13C	
	Δ15Nmax	Age at first decrease	Age at final smoothig	Post-weaning average	Type of the profile	Post-weaning average
Age	0.064	0.897	0.007	0.222	0.008	0.096
Sex	0.835	0.008	0.411	0.038	x	0.097
Socio-economic status	0.617	0.550	0.030	<0.001	x	0.058

[a] results significant at 0.05 level are in bold

The age at the final smoothing of the isotopic curve showed much higher variation, with the youngest individuals (n = 7) being 11-16 months old (slice 3). In the majority of individuals the final smoothing was relatively uniformly dispersed between the age classes of 16-21 months (slice 4, n = 20), 21 months to 2.2 years (slice 5, n = 25) and 2.2-3.7 years (slice 6, n = 18). Finally, in two individuals the final smoothing occurred as late as at 3.7-5.3 years. Surprisingly, individuals, who died during the first decade of life showed an earlier age at final smoothing than older individuals (Table 1). Also, in elites, the final smoothing occurred on average earlier than in non-elites. The sex of the individual had no important impact at this parameter (Figure 4).

According to the RDA analysis, which included all four factors (maximal nitrogen and carbon offsets, age at first decrease, age at final smoothing), the RDA1 axis shows that adults and males display a higher maximal offset, but non-survivors a higher age at first decrease. RDA1 however explains only 7 % of the variability and this result is not statistically significant (ANOVA-like permutation test, F = 1.45, p = 0.848).

Carbon isotopic profile and the covariance between carbon and nitrogen isotopic values

The maximal carbon isotopic offset (Δ13Cmax) ranged between 0.3 and 3.6 ‰, with an average value of 1.6 ± 0.8 ‰, which is substantially higher than the trophic level effect of exclusive breastfeeding (Fuller et al. 2006).

Concerning the shape of the carbon isotopic profile and the relationship to nitrogen isotopic values, most of individuals (n = 43) follow the pattern of positive covariance between δ15N and δ13C values at least in the initial portion of the isotopic profile, i.e. in the period of supposed breastfeeding and weaning (Profile i). Instances where carbon and nitrogen isotope values negatively covary during this period, with decreasing δ15N values along with increasing δ13C values, are relatively scarce (n = 7, Profile ii). A flat carbon profile (Profile iii) occurs in nine cases. Lastly, profile iv, where carbon and nitrogen isotopic values covary positively for most of the time, but the decrease in δ13C values precedes the decrease in δ15N values, occurs in eleven cases. There are eight cases (graves nos. 112/VI,

207, 182, 462, 625, 673, 727, 1171) in which the carbon isotopic profile does not corresponds with any of the types described above. In 112/VI and 673 carbon isotopic values covary positively with δ15N, but do not follow the typical weaning scenario (see Figure 3). In 462 and 1171 atypical carbon profiles occur along with atypical nitrogen profiles, but neither positive nor negative covariance between δ15N and δ13C values is present. In 182, 625 and 727, unexpected shifts in δ13C values occur along the nitrogen isotopic peak typical for breastfeeding. Finally, in individual 207 carbon isotopic values covary negatively with nitrogen isotopic values during both episodes of increased δ15N values (Figure 3).

The type of the carbon isotopic curve differed significantly between individuals, who died before M1 completion and others, with a flat carbon isotopic profile being more common in non-survivors (seven of the nine cases). It was not possible to assess the potential relation to sex or socio-economic status or to include this variable in the multifactorial analysis due to the dominance of profile I, and the subsequently low number of cases in some categories.

Post-weaning isotopic values

The average nitrogen isotopic values from all the slices developed after the smoothing of the nitrogen isotopic curve ranged between 9.1 and 12.5‰ with a mean of 11.2 ± 0.8 ‰. The average carbon isotopic values ranged between –19.1 and –16.3 ‰ with a mean of –17.6 ± 0.6 ‰. Seven individuals died too early to identify the point of the final smoothing of the nitrogen isotopic curve and thus were not included in this analysis. Among adults (n = 44) we were able to compare these "post-weaning averages" with previously published adult isotopic values from bone collagen (Jílková et al. 2019, Kaupová et al. 2018). Childhood nitrogen isotopic values were similar to those from adulthood (mean = 11.0 ± 0.9 ‰, $p = 0.297$), while carbon isotopic values in dentine samples were significantly higher than those from bone (mean = –18.0 ± 0.5 ‰, $p < 0.001$, Figure 5). In the case of nitrogen, post-weaning averages differed significantly between socio-economic classes (Table 1, Figure 5), with elites showing higher δ15N values (mean = 11.5 ± 0.6 ‰) than non-elites (mean = 10.9 ± 0.8‰). Sex too appeared to have a significant impact on δ15N with males (mean = 11.4 ± 0.7‰) showing higher δ15N values than females (mean = 11.0 ± 0.9 ‰). Multifactorial analysis, however, suggests that this probably results from the distinct distribution of both sexes into socio-economic classes (ANOVA, $p = 0.001$ for socio-economic status and 0.139 for sex). Age-at-death had no significant impact on post-weaning δ15N. None of the studied factors show a significant impact on post-weaning carbon isotopic values (Table 1).

Episodes of negative covariance between δ15N and δ13C values during the post-weaning period were observed in twelve cases, including two of the evaluated 18 non-survivors. As stated above, the seven individuals who died too early to identify the exact end point of the isotopic shift associated with weaning were not included in these counts.

Fisher exact test showed no significant difference in the incidence of negative covariance between those who died during M1 formation and others ($p = 0.490$). The number of cases with opposing covariance was too low to effectuate any deeper statistical comparison concerning differences between sexes or socio-economic classes.

In a number of individuals, however, we observed a mild increase in δ15N in later childhood following the period of depletion in 15N during the post-weaning period. The presence of this event, known as the "post-weaning dip" (Tsutaya and Yoneda 2015) was evaluated only in individuals who died after M1 completion, thus presenting a complete isotopic profile. A post-weaning dip was observed in 25 of the 46 individuals. There were no differences in the incidence of post-weaning dip between sexes (Fisher exact test, $p = 0.758$) or between socio-economic groups ($p = 0.773$).

4. Discussions

Nitrogen isotopic values during the period of supposed breastfeeding

In this study, we have considered three potential indicators of physiological stress. Firstly, a peak in nitrogen isotopic values during the period of infancy higher than 3 ‰ (Fuller et al. 2006) may suggest physiological stress during full breastfeeding or around the introduction of the first dietary supplements (which form a relatively small proportion of the diet and are thus still isotopically invisible). This pattern was present in 37 individuals (47 %). This percentage is relatively high, taking into account the known information on the background levels of stress acting on the Mikulčice subadult population. In general, the expansion of Great Moravia corresponded to a favourable era of mild and stable weather, which, together with other favourable natural conditions in South Moravia, enabled high productivity on cultivated land and thus enabled strong population growth (Hladík 2020). On the other hand, Mikulčice's inhabitants surely had to face the negative aspects of urbanization, such as poor sanitation, parasitic infestations, an elevated risk of infection, or dependency on food supplies from the hinterland (Walter and DeWitte 2017). However, the degree of urbanization never exceeded a relatively extensive, proto-urban formation. Thus, though life in a newly-established centre clearly brought problems and challenges previously unknown in a rural and tribal community, the level of stress most probably did not reach that observed in the purely urban formations of the High Middle Ages. This is supported by the results of osteological analysis, according to which the prevalence of non-specific stress indicators in Mikulčice infants and young children was comparable to those from the rural hinterland (Kaupová et al. 2014).

Traditionally, the first introduction of dietary supplementation is viewed as a highly risky period in respect to infection, as the child's immune system encounters a range of new pathogens (Lamberti et al. 2011). It has however been proven that the immunological buffering of breastfeeding still acts during complementary feeding (Kendall et al. 2021).

Moreover, isotopic data from later childhood suggest that the observed pattern could be augmented by the character of post-weaning food. As a ''post-weaning dip'' in nitrogen isotopic values was observed in more than half of the individuals, the lower proportion of animal products consumed during early childhood may boost the Δ15Nmax of these individuals above the 3 ‰ margin.

Secondly, the delayed starting point of the isotopic decrease in δ15N values could suggest biological stress during the early phase of weaning. This parameter, however, has a less precisely defined range than the previous one. According to current medical recommendations, an insufficiency of breastmilk to meet an infant's dietary needs in terms of both micronutrients (e.g. iron) and calories may threaten after just six months of age (Fewtrell et al. 2007, Pérez-Escamilla et al. 2019). Although the portions of supplementary foods could be negligible at the beginning, it is hardly believable that a child could prosper under a regime of exclusive breastfeeding after 1 year of age. In our sample, such a delayed start of the decline in nitrogen isotopic values, with the first decrease observed at 1.3-1.8 years, was present in five individuals (6 %).

Finally, physiological stress in the later phases of weaning would be expressed by a slower decline and the delayed final smoothing of the nitrogen isotopic profile. However, due to the extreme variation reported in the duration of breastfeeding, which could be over 6 years (Dettwyler 2004, Fildes 2017,

Piovanetti 2001), it is not possible identify a point on the nitrogen isotopic curve (before the final smoothing) beyond which isotopic enrichment cannot be due to partial breastfeeding.

Looking at the data from a dietary point of view, the first decrease in δ15N values was observed in most of the individuals at approx. 5-10 months. This suggest a practice of weaning roughly in accordance with current medical recommendations (WHO 2009), implying exclusive breastfeeding for 6 months. Occasionally, infants could have received a notable amount of dietary supplements even earlier, as suggested by the presence of individuals with maximal nitrogen isotopic offset lower than 2 ‰ (n = 3), in these cases, however, other explanations such as change in maternal diet or the use of a wet nurse, cannot be excluded. Due to the competing impact of physiological stress, the detection of the latest age for the first dietary supplements is impossible, and won't be attempted in this paper.

The earliest observed age-at-smoothing of the nitrogen isotopic curve suggests that the first children could well have been fully weaned before 11 months, which is much earlier than the current medical recommendation of partial breastfeeding for 2 years (WHO 2009). This was observed in seven individuals, which is a minor but not negligible proportion of the population. Even if our estimates on age-at-complete weaning are considered to be rather maximalistic (as the presence of biological stress could mimic longer breastfeeding), all except two individuals appeared to have been weaned before 2.2 years. This appears to be relatively early in the perspective of historical populations (Dettwyler 2004, Fildes 2017, Thorvaldsen 2008), but it must be kept in mind that minor supplements of breast milk below 10% of the dietary input would be unobservable isotopically (King et al. 2018). It should also be mentioned here that, due to the chosen sampling strategy, the sixth slice corresponds to relatively long period between 2.2 and 3.75 years. This could further help to blur the potential minor consumption of breastmilk for some time after 2.2 years.

Finally, even with the incremental sampling methods, time averaging inevitably occurs as multiple incremental boundaries are crossed (Czermak et al. 2020). Also, a new dietary source may be introduced well in advance of a detectable isotopic shift. Further, there is some intra-individual variation in teeth development (e.g. Smith 1991). This means that the ages noted in this study as the periods of notable dietary/or physiological change are only rough estimates and should be used rather for comparative purposes.

At the individual level, six cases showed a nitrogen isotopic pattern not compatible with the breastfeeding and weaning scenario. All of these showed low nitrogen isotopic values in the first slice, which can be interpreted as a failure of breastfeeding. However, these originally low δ15N values are in all cases followed by one or two subsequent peaks in δ15N values extending between 0.6 and 3 ‰, which are difficult to interpret. Episodes of physiological stress had to be extremely common in the case of bottle-fed babies consuming non-sterilized food. Neither experimental studies focused on the effect of physiological stress (D'Ortenzio et al. 2015, Drtikolová Kaupová et al. 2021, Fuller et al. 2005, Katzenberg and Lovell 1999, Mekota et al. 2006) nor observations of subadult famine victims (Beaumont and Montgomery 2016), report isotopic shifts associated with malnutrition and/or disease higher than 2 ‰, but four of these six cases show a nitrogen isotopic peak ranging between 2-3 ‰ (see Figure 3, for example). Although an isotopic shift of this extent could well correspond to exclusive breastfeeding (Fuller et al. 2006), this explanation is highly improbable, as the first slice provides the dietary information from approx. 5 months of life. Though there could be variation in the growth pattern of the teeth (Smith 1991), it is extremely unlikely, that an infant could retain the ability to effectively suckle after several weeks (or more probably months) of artificial feeding. This suggests the role of some unknown factor (or a combination of several factors) affecting the isotopic values of these

children in early childhood. In one case only (No. 207, age-at-death = 8-10 years, Figure 3) the extent of the nitrogen isotopic shifts (together with the presence of negative covariance with carbon isotopic values) offers a typical image of repeated episodes of physiological stress (Beaumont and Montgomery 2016, Craig-Atkins et al. 2018), which however, was the child able to overcome. This, as well as the age of two of the afore-mentioned individuals (112-VI and 673), who survived till 40-50, and 20-30 years respectively, contradicts the general view that in the past there was no choice other than breastfeeding to ensure the survival of the child (Dettwyler 2004).

Intra-population comparison of early life δ15N

An intra-population comparison of early life experiences does not suggest that individuals who died during the first decade of life experienced a greater level of environmental stress during early childhood. None of the three parameters listed above was higher in non-survivors. In fact, the age at complete smoothing of the nitrogen isotopic curve was lower in non-survivors, as was the maximal nitrogen isotopic offset (though here, the difference was on the borderline of statistical significance). As lower Δ15Nmax could result from an earlier start to the weaning process (before 5 months), the observed pattern may theoretically demonstrate lower parental investment (Quinlan 2007), impacting the life expectancy of a child.

The delayed onset of the nitrogen isotopic decrease observed in females in comparison to males is probably caused by the higher prevalence of individuals with a first decrease observed as late as in the fourth slice, with four of the five cases being female. As noted above, such a pattern is probably linked with physiological stress, but considering the low number of cases, it is unclear whether this can be seen as a reflection of a higher level of stress during the weaning process, imposed systematically on female children e.g. by the lower quality of complementary food.

The earlier occurrence of final smoothing observed in elites in comparison to non-elites has at least two equally valid explanations: the first relates to the higher incidence of stress during the weaning process in non-elites, the second to the earlier weaning of elite children. The latter may well be linked to the ongoing Christianization of the Great Moravian population. Although there is no direct written testimony from the Great Moravian context, certain rules concerning family life and childcare probably existed (Thorvaldsen 2008): as a unique example, the 9th century document "The Responses of Pope Nicholas I to the Questions of the Bulgars" recommends sexual abstinence during the entire period of breastfeeding (Bartoňková et al. 1971). As the pressure to abide by Christian rules was probably higher in elite groups, this could have led at least some elite women in Mikulčice to shorten the breastfeeding period. To estimate the strength of this argument, however, is beyond the scope of this paper.

In any case, as RDA analysis including all four explicative factors gave non-significant results, explaining only a small portion of the data variation, the results of uni-factorial analysis suggesting a relationship between some parameters and age-at-death, sex or social status must be considered with caution.

Carbon isotopic values during the period of supposed breastfeeding

The presence of negative covariance, usually regarded as an indication of physiological stress, in the initial part of the isotopic profile was relatively rare in this dataset, being observed in only seven cases. This however does not necessarily mean that the level of physiological stress during infancy and early childhood was low. As stated above, the results of experimental studies suggest that the isotopic changes in a carbon body pool in relation to physiological stress may vary (Canterbury et al. 2020, D'Ortenzio et al. 2015, Drtikolová Kaupová et al. 2021, Fuller et al. 2005, Neuberger et al. 2013). This inconsistency can theoretically be caused by intra-individual differences in the amount of body fat and the phase of malnutrition. At first, a decrease in δ13C values may reflect the use of carbon from energy

sources to synthesize non-essential amino acids, but later body protein catabolism may predominate and cause an increase in δ13C values (Drtikolová Kaupová et al. in press, King et al. 2018, Salesse et al. 2019). Moreover, even if we accept negative covariance between carbon and nitrogen isotopic values as being the primary response of a subadult organism to biological stress, this concept could be valid only a in purely terrestrial C3 plant-based ecosystem, where isotopically distinct food groups such as fish (either freshwater or marine) or C4 plants were not accessible in substantial quantities. This was clearly not the case for the Great Moravian population (Halffman and Velemínský 2015, Jílková et al. 2019, Kaupová et al. 2018).

In fact, taking into account both the potential roles of stress and millet consumption, it is quite surprising that most of the individuals (n = 43) show a positive covariance between carbon and nitrogen isotopic values, as can be observed during the process of breastfeeding and weaning under healthy circumstances and with a stable and isotopically analogous diet for both mother and child. This could mean that millet played a special role in the diet of infants or potentially also their breastfeeding mothers. As will be explained in detail bellow, carbon isotopic values from the post-weaning period suggest, that the childhood diet included more millet than that of adults. The persistence of the carbon isotopic peak in these circumstances is thus quite surprising. If an infant was being weaned onto resources with a higher millet proportion than those in the maternal diet, we would expect a rise in δ13C values, "cancelling out" the trophic level shift due to weaning. The result would be either a flat carbon isotopic profile (as observed in nine cases in this study) or the occurrence of negative covariance mimicking the stress pattern described above (King et al. 2018). As this is not true in a majority of the sample, it means that either millet was part of a special diet preferred for some reason by lactating women, or was introduced as the very first dietary supplement. Consumed in the form of porridge (Adamson 2004), millet has the ideal consistency for first dietary supplementation, it is also easily digestible, and in comparison to other cereals has a higher nutritional value with a high content of vitamins, minerals and antioxidants (Aurelia et al. 2020, Kulp 2000, Weber and Fuller 2008). While this information was not known to Great Moravian people, the consumption of millet could have reduced the risk of both malnutrition and disease in their children during the period of complementary feeding and – from the immunological point of view – the extremely risky time of the complete cessation of breastfeeding (Kendall et al. 2021). At the same time, due to a low protein content in comparison to breastmilk, millet-based dietary supplements have the potential to impact carbon isotopic values (as carbohydrates could be partially used to build the carbon skeletons of non-essential amino-acids), while breast milk remains the main source of nitrogen (Fuller et al. 2006). Another point suggesting a special role for millet impacting the isotopic values from infancy (either directly or indirectly via maternal diet) is the extent of maximal carbon isotopic offset occurring during the supposed breastfeeding period, which in the vast majority of individuals (n = 55) surpassed the maximum of 1 ‰ which could be attributable to the trophic level effect of breastfeeding (Fuller et al. 2006). In 16 individuals Δ13Cmax was even higher than 2 ‰.

In cases showing profile iv, where carbon and nitrogen covary positively but the decrease in δ13C precedes the decrease in δ15N (eleven cases), as well as in eight cases not corresponding with any of the outlined scenarios, a number of equally plausible explanations may be offered and thus, we will not attempt any interpretation here.

Incidence of distinct carbon isotopic profiles in particular age groups

The substantially higher prevalence of flat carbon isotopic profiles observed in those who died before M1 completion has no simple explanation. The absence of breastfeeding as a factor impacting further survival would be a valid cause (Craig Atkins et al. 2018), but all the non-survivors with a flat carbon isotopic profile show a nitrogen isotopic peak within or very close to the 2-3 ‰ range observed in fully

breastfed babies (Fuller et al. 2006). As discussed above, such a high nitrogen isotopic offset is unlikely to have been caused solely by physiological stress. A dietary explanation might be the weaning of the child onto resources with a higher millet proportion than in the maternal diet. However this, like other explanations arising from dietary factors, does not explain, why this scenario was more prevalent in non-survivors.

Carbon and nitrogen isotopic values in the post-weaning period

The intra-individual comparison of the post-weaning isotopic averages vs bone isotopic values confirms the earlier results of Jílková et al. (2019), reporting a higher importance of millet in childhood diet. While Jílková's team focused on the later period of childhood (approx. 8-10 years as a portion of the M2 root was sampled), this study documents the importance of millet in earlier phases of life. The span of the post-weaning period actually varies between 1.3-9.4 and 3.75-9.4 years, based at the age on the final smoothing of the isotopic curve.

On the other hand, the proportion of animal products consumed by individuals included in this study was stable throughout life, opposing the results from Jílková et al. (2019), who observed a lower proportion of animal products in the subadult diet of individuals buried around Mikulčice 6th church. It needs to be stated, however, that both studies focused on different population subgroups buried within the Mikulčice settlement agglomeration. While the 6th church is located in the suburb of Mikulčice, this study includes mainly individuals buried at the fortified Acropolis i.e. within a supposed residential area of the highest elites of Great Moravian society (Poláček 2008). Thus, the combined results of the two studies suggest that while the higher input of millet in a child's diet was a widespread practice among Mikulčice inhabitants, the access of children to animal products may have differed in particular parts of Mikulčice.

Intra-population variation in post-weaning isotopic values

The absence of any systemic differences in post-weaning averages between non-survivors vs individuals who died after M1 completion suggests that the quality of the post-weaning diet did not impact the chance of survival, and that there was no higher incidence of long-term physiological stress in non-survivors. However, it must be noted that severe and long-lasting physiological stress would have had to act to significantly impact the post-weaning values, averaging the isotopic signal from several years. The distinct nitrogen isotopic values of elite vs non-elite graves attest that the socio-economic differences in diet documented isotopically in later life periods (Jílková et al. 2019, Kaupová et al. 2018) were already present during childhood.

Exploring the latter part of the isotopic profile in greater detail, the episodes of negative covariance, generally viewed as evidence of physiological stress, had no higher prevalence in non-survivors. Cases of the so called "post-weaning dipp", i.e. the notable increase in δ15N in later childhood, which is commonly observed in agricultural populations and is generally viewed as testimony to the lower input of animal products in the period immediately following weaning (Tsutaya and Yoneda 2015), were much more common than episodes of opposing covariance (in more than half of cases). While this could also be viewed as a reflection of anabolic activity linked to mid-growth spurt (Kendall et al. 2021), but the knowledge on the role of growth velocity is still limited, with current experimental studies not providing convincing evidence of the assertion of this factor on nitrogen isotopic values (Reitsema and Muir 2015, Waters-Rist and Katzenberg 2010). Further, the presence of the "post-weaning" dip appeared to be linked to the subsistence strategy (Tsutaya and Yoneda 2015). In a Romano-British population sample, for which we have written records on the perception of childhood, the timing of the dip coincides with the change in social status of children (Cocozza et al. 2021). Both of these findings further enhance the first explanation.

And finally, as emerges from the above discussion on the testimonial power of carbon isotopic values in respect to stress, it appears evident that neither can the presence of physiological stress be excluded in these individuals.

Limitations of the study

The interpretative difficulties due to equifinality (King et al. 2018), unavoidable by current methods, have been repeatedly stressed over the course of the discussion, so need no recapitulation here.

As a much more important point, we would like to stress the limited knowledge on the action of physiological stress in the subadult organism. Most of the experimental studies on physiological stress carried out on humans concern adult individuals only (D'Ortenzio et al. 2015, Drtikolová Kaupová et al. 2021, Fuller et al. 2005, Katzenberg and Lovell 1999, Mekota et al. 2006, Tea et al. 2021). The only exception is the work by Beaumont and Montgomery (2016), using incremental sampling of dentine tissue from developing teeth, this study is however based on an archaeological population, albeit from the well-documented context of a 19th century famine event. Moreover, it has to be kept in mind that the isotopic reflection of famine, where starvation is the primary cause of death (though it surely acted in synergy with a number of diseases, e.g. Solomons 2007), may well differ from the less extreme consequences, where the pathophysiological origin of tissue catabolism may prevail.

Moreover, the limited number of experimental studies based on subadult individuals from different animal species suggest that a meaningful level of nutritional stress need not necessarily induce an isotopic response in either δ15N or δ13C (Ambrose 2002, Kempster et al. 2007, Williams et al. 2007). This could be due to the potentially competing influence of negative and positive nitrogen balance (Waters-Rist and Katzenberg 2010). Also, a reduction in growth, rather than tissue catabolism, may be the primary response of subadults to dietary stress (Ambrose 2002). Thus, more experimental studies on subadult organisms are greatly needed to confirm that the pattern of negative covariance between nitrogen and carbon isotopic values is the real response of subadult body to physiological stress.

5. Conclusion

Alhough the current data – mainly maximal nitrogen isotopic offsets commonly increased over 3‰ – suggest that physiological stress might have been relatively common during early childhood, there are no indications that the occurrence of such physiological stress in this life phase affected life expectancy in the long term, or that the stress level in early life differed substantially between sexes or socio-economic classes.

The isotopic data also suggest that in the majority of the studied sample, a notable proportion of supplementary food was introduced between 5–10 months, i.e. roughly in the interval recommended by current medical advice. The first children in our dataset were fully weaned before 10 months of age. The latest age at the final smoothing of the nitrogen isotopic curve suggesting weaning was as late as 3.75–5.3 years, but this was observed in 2 individuals only. All the other individuals showed final smoothing of the nitrogen isotopic curve at 2.2–3.75 years. With respect to age-at-weaning, this has to be viewed rather as the upper range of the possible estimate, as the impact of physiological stress cannot be excluded in these seemingly late weaned children. Also, it has to be kept in mind that minor supplementation by breastmilk would not be observable by isotopic analysis. Six individuals do not show the nitrogen isotopic shift attributable to breastfeeding at all.

Carbon isotopic values from infancy and early childhood suggest that millet probably had an important role at the very beginning of the weaning process, and/or in the diet of lactating women. The isotopic data from later childhood, together with the data from a previous study (Jílková et al. 2019), suggest that a higher input of millet was characteristic of childhood idet during the first decade of life.

Nitrogen isotopic values from the post-weaning period suggest that though a small drop in the dietary proportion of animal products may occur in the period immediately following weaning, the average proportion of animal products consumed during childhood is comparable to that consumed in adulthood. The socio-economically motivated differences in dietary behaviour previously observed in adults (Kaupová et al. 2018) were already present during childhood.

In future, a comparison of the set of potential indicators of physiological stress listed above with other population samples with distinct documented levels of environmental stress will be carried out. Such comparisons may bring more knowledge about the level of environmental stress that the subadult population of Mikulčice had to face.

Funding and competing interests

This study was supported financially by Czech Science Foundation (Project Id: 19-13265S) and Ministry of Culture of the Czech Republic (DKRVO 2019-2023/7.I.d, 00023272). The authors have no conflicts of interest to declare that are relevant to the content of this article.

Acknowledgements

We would like to thank the following institutions for their financial support: the Czech Science Foundation (Grant number: 19-13265S), and the Ministry of Culture of the Czech Republic (Grant number: DKRVO 2019-2023/7.I.d, 00023272). Alastair Millar helped to revise the English.

References

1. Adamson MW (2004) Food in medieval times. Greenwood Publishing Group, Westport
2. Ambrose SH (2002) Controlled diet and climate experiments on nitrogen isotope ratios of rats. In: Ambrose SH, Katzenberg MA (eds) Biogeochemical approaches to paleodietary analysis. Springer, New York, pp 243-259. https://doi.org/10.1007/0-306-47194-9_12
3. Ambrose SH, Norr L (1993) Experimental evidence for the relationship of the carbon isotope ratios of whole diet and dietary protein to those of bone collagen and carbonate. In: Lambert JB, Grupe G (eds) Prehistoric human bone: Archaeology at the molecular level. Springer, New York, pp 1-37. https://doi.org/10.1007/978-3-662-02894-0_1
4. Aurelia LN, Mihaela H, Anca G, Lavinia I (2020) Use of millet grain in weaning pigs diet: Effects on performance and health status. Arch Zootech 23:143-154. https://doi.org/10.2478/azibna-2020-0019
5. Balasse M, Bocherens, H, Mariotti A, Ambrose SH (2001) Detection of dietary changes by intra-tooth carbon and nitrogen isotopic analysis: an experimental study of dentine collagen of cattle (Bos taurus). J Archaeol Sci 28:235-245. https://doi.org/10.1006/jasc.1999.0535

6. Bartoňková D, Haderka K, Havlík L, Ludvíkovský J, Vašica J, Večerka R (1971) Magnae moraviae fontes historici IV. Universita J. Evangelisty Purkyně, Brno
7. Baumslag N (1987) Breastfeeding: Cultural practices and variations. In: Jelliffe D, Jelliffe E (eds) Advances in international maternal and child health 7. Clarendon Press, Oxford, pp 36-50
8. Beaumont J, Atkins E-C, Buckberry J, Haydock H, Horne P, Howcroft R, Mackenzie K, Montgomery J (2018) Comparing apples and oranges: Why infant bone collagen may not reflect dietary intake in the same way as dentine collagen. Am J Phys Anthropol 167:524-540. https://doi.org/10.1002/ajpa.23682
9. Beaumont J, Geber J, Powers N, Wilson A, Lee-Thorp J, Montgomery J (2013) Victims and survivors: Stable isotopes used to identify migrants from the Great Irish Famine to 19th century London. Am J Phys Anthropol 150:87-98. https://doi.org/10.1002/ajpa.22179
10. Beaumont J, Gledhill A, Lee-Thorp J, Montgomery J (2013) Childhood diet: A closer examination of the evidence from dental tissues using stable isotope analysis of incremental human dentine. Archaeometry 55:277-295. https://doi.org/10.1111/j.1475-4754.2012.00682.x
11. Beaumont J, Montgomery J (2016) The Great Irish Famine: Identifying starvation in the tissues of victims using stable isotope analysis of bone and incremental dentine collagen. PLoS One 11:e0160065. https://doi.org/10.1371/journal.pone.0160065
12. Bentley GR, Paine RR, Boldsen J (2001) Fertility changes with the prehistoric transition to agriculture: Perspectives from reproductive ecology and paleodemography. In: Ellison PT (ed) Reproductive ecology and human evolution. Aldine de Gruyter, New York, pp 203-231
13. Berti C, Agostoni C, Davanzo R, Hyppönen E, Isolauri E, Meltzer HM, Steegers-Theunissen RPM, Cetin I (2017) Early-life nutritional exposures and lifelong health: Immediate and long-lasting impacts of probiotics, vitamin D, and breastfeeding. Nutr Rev 75:83-97. https://doi.org/10.1093/nutrit/nuw056
14. Brůžek J (2002) A method for visual determination of sex, using the human hip bone. Am J Phys Anthropol 117:157-168. https://doi.org/10.1002/ajpa.10012
15. Brůžek J, Santos F, Dutailly B, Murail P, Cunha E (2017) Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. Am J Phys Anthropol 164:440-449. https://doi.org/10.1002/ajpa.23282
16. Calce SE (2012) A new method to estimate adult age-at-death using the acetabulum. Am J Phys Anthropol 148:11-23. https://doi.org/10.1002/ajpa.22026
17. Canterbury JA, Beck CW, Dozier C, Hoffmeister K, Magaro J, Perrotti AG, Wright LE (2020) Bariatric surgery as a proxy for nutritional stress in stable isotope investigations of archaeological populations. J Archaeol Sci Rep 30:102196. https://doi.org/10.1016/j.jasrep.2020.102196
18. Cocozza C, Fernandes R, Ughi A, Groß M, Alexander MM (2021) Investigating infant feeding strategies at Roman Bainesse through Bayesian modelling of incremental dentine isotopic data. Int J Osteoarchaeol 31:429-439. https://doi.org/10.1002/oa.2962
19. Craig-Atkins E, Towers J, Beaumont J (2018) The role of infant life histories in the construction of identities in death: An incremental isotope study of dietary and physiological status among children afforded differential burial. Am J Phys Anthropol 167:644-655. https://doi.org/10.1002/ajpa.23691
20. Czermak A, Fernández-Crespo T, Ditchfield PW, Lee-Thorp JA (2020) A guide for an anatomically sensitive dentine microsampling and age-alignment approach for human teeth isotopic sequences. Am J PhysAnthropol 173:776-783. https://doi.org/10.1002/ajpa.24126
21. Czosnykowska-Łukacka M, Królak-Olejnik B, Orczyk-Pawiłowicz M (2018) Breast milk macronutrient components in prolonged lactation. Nutrients 10:1893. https://doi.org/10.3390/nu10121893

22. de Luca A, Boisseau N, Tea I, Louvet I, Robins RJ, Forhan A, Charles M-A, Hankard R (2012) δ15N and δ13C in hair from newborn infants and their mothers: A cohort study. Pediatr Res 71:598-604. https://doi.org/10.1038/pr.2012.3
23. Demmelmair H, von Rosen J, Koletzko B (2006) Long-term consequences of early nutrition. Early Hum Dev 82:567-574. https://doi.org/10.1016/j.earlhumdev.2006.07.004
24. DeNiro MJ, Epstein S (1978) Influence of diet on the distribution of carbon isotopes in animals. Geochim Cosmochim Acta 42:495-506. https://doi.org/10.1016/0016-7037(78)90199-0
25. DeNiro MJ, Epstein S (1981) Influence of diet on the distribution of nitrogen isotopes in animals. Geochim Cosmochim Acta 45: 341-351. https://doi.org/10.1016/0016-7037(81)90244-1
26. Dettwyler KA (2004) When to wean: Biological versus cultural perspectives. Clin Obstet Gynecol 47:712-723. https://doi.org/10.1097/01.grf.0000137217.97573.01
27. D'Ortenzio L, Brickley M, Schwarcz H, Prowse T (2015) You are not what you eat during physiological stress: Isotopic evaluation of human hair. Am J Phys Anthropol 157:374-388. https://doi.org/10.1002/ajpa.22722
28. Drtikolová Kaupová S, Velemínský P, Grossová I, Půtová L, Cvrček J (In press) Stable isotope values of carbon and nitrogen in hair compared to bone collagen from individuals with known medical histories (Bohemia, 19th–21st centuries). International Journal of Osteoarchaeology. https://doi.org/10.1002/oa.3125
29. Drtikolová Kaupová S, Cvrček J, Grossová I, Hadrava J, Půtová L, Velemínský P (2021) The impact of pathological conditions on carbon and nitrogen isotopic values in the bone collagen of individuals with known biographic data and medical history. Int J Osteoarchaeol 31:1105-1124. https://doi.org/10.1002/oa.3022
30. Eerkens JW, Bartelink EJ (2013) Sex-biased weaning and early childhood diet among middle holocene hunter–gatherers in Central California. Am J Phys Anthropol 152:471-483. https://doi.org/10.1002/ajpa.22384
31. Eerkens JW, Berget AG, Bartelink EJ (2011) Estimating weaning and early childhood diet from serial micro-samples of dentin collagen. J Archaeol Sci 38:3101-3111. https://doi.org/10.1016/j.jas.2011.07.010
32. Fewtrell MS, Morgan JB, Duggan C, Gunnlaugsson G, Hibberd PL, Lucas A, Kleinman RE (2007) Optimal duration of exclusive breastfeeding: What is the evidence to support current recommendations? Am J Clin Nutr 85:635S-638S. https://doi.org/10.1093/ajcn/85.2.635S
33. Fildes V (2017) The culture and biology of breastfeeding: An historical review of Western Europe. In: Stuart-Macadam P, Dettwyler KA (eds) Breastfeeding: Biocultural Perspectives, Routledge, New York, pp 75-99
34. Fogel ML, Tuross N, Owsley DW (1989) Nitrogen isotope tracers of human lactation in modern and archaeological populations. In: Annual Report of Geophysical Laboratory Carnegie Institution of Washington. pp 111-117
35. Fuller BT, Fuller JL, Harris DA, Hedges REM (2006) Detection of breastfeeding and weaning in modern human infants with carbon and nitrogen stable isotope ratios. Am J Phys Anthropol 129:279-293. https://doi.org/10.1002/ajpa.20249
36. Fuller BT, Fuller JL, Sage NE, Harris DA, O'Connell TC, Hedges REM (2005) Nitrogen balance and δ15N: Why you're not what you eat during nutritional stress. Rapid Commun Mass Spectrom 19:2497-2506. https://doi.org/10.1002/rcm.2090
37. Fuller BT, Richards MP, Mays SA (2003) Stable carbon and nitrogen isotope variations in tooth dentine serial sections from Wharram Percy. J Archaeol Sci 30:1673-1684. https://doi.org/10.1016/S0305-4403(03)00073-6

38. Halffman CM, Velemínský P (2015) Stable isotope evidence for diet in early medieval Great Moravia (Czech Republic). J Archaeol Sci Rep 2:1-8. https://doi.org/10.1016/j.jasrep.2014.12.006
39. Herold H (2012) Fortified settlements of the 9th and 10th centuries ad in Central Europe: Structure, function and symbolism. Mediev Archaeol 56:60-84. https://doi.org/10.1179/0076609712Z.0000000003
40. Herrscher E (2004) Comportements socioculturels liés à l'allaitement et au sevrage: Le cas d'une population grenobloise sous l'Ancien Régime. Annales de la Fondation Fyssen 20:46-66
41. Herrscher E (2013) Détection isotopique des modalités d'allaitement et de sevrage à partir des ossements archéologiques. Cah Nutr Diet 48:75-85. https://doi.org/10.1016/j.cnd.2012.12.004
42. Herrscher E, Goude G, Metz L (2017) Longitudinal study of stable isotope compositions of maternal milk and implications for the palaeo-diet of infants. Bull Mem Soc Anthropol Paris 29:131139. https://doi.org/10.1007/s13219-017-0190-4
43. Howcroft R, Eriksson G, Lidén K (2012) Conformity in diversity? Isotopic investigations of infant feeding practices in two iron age populations from Southern Öland, Sweden. Am J Phys Anthropol 149:217-230. https://doi.org/10.1002/ajpa.22113
44. World Health Organization (2009) Infant and Young Child Feeding: Model Chapter for Textbooks for Medical Students and Allied Health Professionals. http://www.ncbi.nlm.nih.gov/books/NBK148965/. Accessed 01 June 2022
45. Jay M (2009) Breastfeeding and weaning behaviour in archaeological populations: Evidence from the isotopic analysis of skeletal materials. Child Past 2:163-178. https://doi.org/10.1179/cip.2009.2.1.163
46. Jayachandran S, Kuziemko I (2011) Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. Q J Econ 126:1485-1538. https://doi.org/10.1093/qje/qjr029
47. Jílková M, Kaupová S, Černíková A, Poláček L, Brůžek J, Velemínský P (2019) Early medieval diet in childhood and adulthood and its reflection in the dental health of a Central European population (Mikulčice, 9th–10th centuries, Czech Republic). Arch Oral Biol 107:104526. https://doi.org/10.1016/j.archoralbio.2019.104526
48. Kalhous D (2020) Church Organisation as a Bearer of New Culture and Innovations and Potential Support of Central Power. In: Poláček L (ed) Great Moravian elites from Mikulčice. Archeologický ústav AV ČR Brno, Brno, pp 61-67
49. Katzenberg MA, Herring DA, Saunders SR (1996) Weaning and infant mortality: Evaluating the skeletal evidence. Am J Phys Anthropol 101(S23):177-199. https://doi.org/10.1002/(SICI)1096-8644(1996)23+<177::AID-AJPA7>3.0.CO,2-2
50. Katzenberg MA, Lovell NC (1999) Stable isotope variation in pathological bone. Int J Osteoarchaeol 9:316-324. https://doi.org/10.1002/(SICI)1099-1212(199909/10)9:5<316::AID-OA500>3.0.CO,2-D
51. Kaupová S, Herrscher E, Velemínský P, Cabut S, Poláček L, Brůžek J (2014) Urban and rural infant-feeding practices and health in early medieval Central Europe (9th–10th Century, Czech Republic). Am J Phys Anthropol 155:635-651. https://doi.org/10.1002/ajpa.22620
52. Kaupová S, Velemínský P, Herrscher E, Sládek V, Macháček J, Poláček L, Brůžek J (2018) Diet in transitory society: Isotopic analysis of medieval population of Central Europe (ninth–eleventh century AD, Czech Republic). Archaeol Anthropol Sci 10:923-942. https://doi.org/10.1007/s12520-016-0427-8

53. Kempster B, Zanette L, Longstaffe FJ, MacDougall-Shackleton SA, Wingfield JC, Clinchy M (2007) Do stable isotopes reflect nutritional stress? Results from a laboratory experiment on song sparrows. Oecologia 151:365-371. https://doi.org/10.1007/s00442-006-0597-7
54. Kendall E, Millard A, Beaumont J (2021) The "weanling's dilemma" revisited: Evolving bodies of evidence and the problem of infant paleodietary interpretation. Am J Phys Anthropol 175(S72):57-78. https://doi.org/10.1002/ajpa.24207
55. King CL, Halcrow SE, Millard AR, Gröcke DR, Standen VG, Portilla M, Arriaza BT (2018) Let's talk about stress, baby! Infant-feeding practices and stress in the ancient Atacama desert, Northern Chile. Am J Phys Anthropol 166:139-155. https://doi.org/10.1002/ajpa.23411
56. Kulp K (2000) Handbook of cereal science and technology, revised and expanded. Taylor & Francis, New York
57. Kuna M et al. (2018) Mikulčice. Archeologický atlas ČR. https://www.archeologickyatlas.cz/cs/lokace/mikulcice_ho_hradiste. Accessed 27 July 2022
58. Lamberti LM, Fischer Walker CL, Noiman A, Victora C, Black RE (2011) Breastfeeding and the risk for diarrhea morbidity and mortality. BMC Public Health 11:S15. https://doi.org/10.1186/1471-2458-11-S3-S15
59. Lovejoy CO (1985) Dental wear in the Libben population: Its functional pattern and role in the determination of adult skeletal age at death. Am J Phys Anthropol 68:47-56. https://doi.org/10.1002/ajpa.1330680105
60. Macháček J (2013) Great Moravian central places and their practical function, social significance and symbolic meaning. In: Ettel P, Werther L (eds) Zentrale Orte und Zentrale Räume des Frühmittelalters in Süddeutschland. RGZM – Tagungen 18. Römisch-Germanisches Zentralmuseum, Mainz, pp 235-248
61. Mays S (2010) The effects of infant feeding practices on infant and maternal health in a medieval community. Child Past 3:63-78. https://doi.org/10.1179/cip.2010.3.1.63
62. McDade TW (2005) Life history, maintenance, and the early origins of immune function. Am J Hum Biol 17:81-94. https://doi.org/10.1002/ajhb.20095
63. Mekota A-M, Grupe G, Ufer S, Cuntz U (2006) Serial analysis of stable nitrogen and carbon isotopes in hair: Monitoring starvation and recovery phases of patients suffering from anorexia nervosa. Rapid Commun Mass Spectrom 20:1604-1610. https://doi.org/10.1002/rcm.2477
64. Murail P, Brůžek J, Houët F, Cunha E (2005) DSP: A tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurements. Bull Mem Soc Anthropol Paris 17:167-176. https://doi.org/10.4000/bmsap.1157
65. Neuberger FM, Jopp E, Graw M, Püschel K, Grupe G (2013) Signs of malnutrition and starvation—Reconstruction of nutritional life histories by serial isotopic analyses of hair. Forensic Sci Int 226:22-32. https://doi.org/10.1016/j.forsciint.2012.10.037
66. O'Connell TC, Kneale CJ, Tasevska N, Kuhnle GGC (2012) The diet-body offset in human nitrogen isotopic values: A controlled dietary study. Am J Phys Anthropol 149:426-434. https://doi.org/10.1002/ajpa.22140
67. Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: Community Ecology Package. R package version 2.2-1. https://CRAN.R-project.org/package=vegan. Accessed 10 June 2022
68. Palou A, Picó C (2009) Leptin intake during lactation prevents obesity and affects food intake and food preferences in later life. Appetite 52:249-252. https://doi.org/10.1016/j.appet.2008.09.013
69. Pearson JA, Hedges REM, Molleson TI, Özbek M (2010) Exploring the relationship between weaning and infant mortality: An isotope case study from Aşıklı Höyük and Çayönü Tepesi. Am J Phys Anthropol 143:448-457. https://doi.org/10.1002/ajpa.21335

70. Pérez-Escamilla R, Buccini GS, Segura-Pérez S, Piwoz E (2019) Perspective: Should exclusive breastfeeding still be recommended for 6 months? Adv Nutr 10:931-943. https://doi.org/10.1093/advances/nmz039
71. Phenice TW (1969) A newly developed visual method of sexing the os pubis. Am J Phys Anthropol 30:297-301. https://doi.org/10.1002/ajpa.1330300214
72. Piovanetti Y (2001) Breastfeeding beyond 12 months: An historical perspective. Pediatr Clin North Am 48:199-206. https://doi.org/10.1016/S0031-3955(05)70294-7
73. Poláček L (2008) Great Moravia, the Power Centre at Mikulčice and the Issue of the Socio-economic Structure. In: Velemínský P, Poláček L (eds) Studien zum Burgwall von Mikulčice VIII. Archeologický ústav AV ČR Brno, Brno, pp 11-44
74. Poláček L (2018) The Mikulčice-Valy Stronghold and Great Moravia. Archeologický ústav AV ČR Brno, Brno
75. Prowse TL, Saunders SR, Schwarcz HP, Garnsey P, Macchiarelli R, Bondioli L (2008) Isotopic and dental evidence for infant and young child feeding practices in an imperial Roman skeletal sample. Am J Phys Anthropol:137:294-308. https://doi.org/10.1002/ajpa.20870
76. Quinlan RJ (2007) Human parental effort and environmental risk. Proc Royal Soc B 274:121-125. https://doi.org/10.1098/rspb.2006.3690
77. R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 10 June 2022
78. Reitsema LJ (2013) Beyond diet reconstruction: Stable isotope applications to human physiology, health, and nutrition. Am J Hum Biol 25:445-456. https://doi.org/10.1002/ajhb.22398
79. Reitsema LJ, Muir AB (2015) Growth velocity and weaning δ15N "Dips" during ontogeny in Macaca mulatta. Am J Phys Anthropol 157:347-357. https://doi.org/10.1002/ajpa.22713
80. Richards MP, Mays S, Fuller BT (2002) Stable carbon and nitrogen isotope values of bone and teeth reflect weaning age at the Medieval Wharram Percy site, Yorkshire, UK. Am J Phys Anthropol 119:205-210. https://doi.org/10.1002/ajpa.10124
81. Salesse K, Kaupová S, Brůžek J, Kuželka V, Velemínský P (2019) An isotopic case study of individuals with syphilis from the pathological-anatomical reference collection of the national museum in Prague (Czech Republic, 19th century A.D.). Int J Paleopathol 25:46-55. https://doi.org/10.1016/j.ijpp.2019.04.001
82. Sellen DW (2001) Comparison of infant feeding patterns reported for nonindustrial populations with current recommendations. J Nutr 131:2707-2715. https://doi.org/10.1093/jn/131.10.2707
83. Shamir R (2016) The benefits of breast feeding. Protein in Neonatal and Infant Nutrition: Recent Updates 86:67-76. https://doi.org/10.1159/000442724
84. Scheuer L, Black S, Schaefer MC (2008) Juvenile Osteology: A Laboratory and Field Manual (1st edition). Academic Press, Cambridge
85. Schmitt A (2005) Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque. Bull Mem Soc Anthropol Paris 17:89-101. https://doi.org/10.4000/bmsap.943
86. Schmitt A (2008) Une nouvelle méthode pour discriminer les individus décédés avant ou après 40 ans à partir de la symphyse pubienne. J Med Leg Droit Med 51:15-24.
87. Schoeninger MJ, DeNiro MJ (1984) Nitrogen and carbon isotopic composition of bone collagen from marine and terrestrial animals. Geochim Cosmochim Acta: 48:625-639. https://doi.org/10.1016/0016-7037(84)90091-7

88. Schurr MR (1997) Stable nitrogen isotopes as evidence for the age of weaning at the Angel Site: A comparison of isotopic and demographic measures of weaning age. J Archaeol Sci 24:919-927. https://doi.org/10.1006/jasc.1996.0171
89. Schwarcz HP (2002) Some Biochemical Aspects of Carbon Isotopic Paleodiet Studies. In: Ambrose SH, Katzenberg MA (eds) Biogeochemical approaches to paleodietary analysis. Springer, New York, pp 189-209. https://doi.org/10.1007/0-306-47194-9_10
90. Smith BH (1991) Standards of human tooth formation and dental age assessment. In: Kelley MA, Larsen CS (eds) Advances in dental anthropology. Wiley, New York, pp 143-168
91. Solomons NW (2007) Malnutrition and infection: An update. Br J Nutr 98:S5-S10. https://doi.org/10.1017/S0007114507832879
92. Stewart NA, Gerlach RF, Gowland RL, Gron KJ, Montgomery J (2017) Sex determination of human remains from peptides in tooth enamel. Proc Nat Acad Sci USA 114:13649-13654. https://doi.org/10.1073/pnas.1714926115
93. Szpak P, Metcalfe JZ, Macdonald RA (2017) Best practices for calibrating and reporting stable isotope measurements in archaeology. J Archaeol Sci Rep 13:609-616. https://doi.org/10.1016/j.jasrep.2017.05.007
94. Tea I, De Luca A, Schiphorst A-M, Grand M, Barillé-Nion S, Mirallié E, Drui D, Krempf M, Hankard R, Tcherkez G (2021) Stable isotope abundance and fractionation in human diseases. Metabolites 11:370. https://doi.org/10.3390/metabo11060370
95. Thorvaldsen G (2008) Was there a European breastfeeding pattern? Hist Fam 13:283-295. https://doi.org/10.1016/j.hisfam.2008.08.001
96. Tomori C, Palmquist AEL, Quinn EA (2017) Breastfeeding: New anthropological approaches. Routledge, New York
97. Tsutaya T, Yoneda M (2015) Reconstruction of breastfeeding and weaning practices using stable isotope and trace element analyses: A review. Am J Phys Anthropol 156:2-21. https://doi.org/10.1002/ajpa.22657
98. van der Haas VM, Garvie-Lok S, Bazaliiskii VI, Weber AW (2018) Evaluating sodium hydroxide usage for stable isotope analysis of prehistoric human tooth dentine. J Archaeol Sci Rep 20:80-86. https://doi.org/10.1016/j.jasrep.2018.04.013
99. Walker PL (2008) Sexing skulls using discriminant function analysis of visually assessed traits. Am J Phys Anthropol 136:39-50. https://doi.org/10.1002/ajpa.20776
100. Waters-Rist AL, Bazaliiskii VI, Weber AW, Katzenberg MA (2011) Infant and child diet in Neolithic hunter-fisher-gatherers from cis-baikal, Siberia: Intra-long bone stable nitrogen and carbon isotope ratios. Am J Phys Anthropol 146:225-241. https://doi.org/10.1002/ajpa.21568
101. Waters-Rist AL, Katzenberg MA (2010) The effect of growth on stable nitrogen isotope ratios in subadult bone collagen. Internat J Osteoarchaeol 20:172-191. https://doi.org/10.1002/oa.1017
102. Weber SA, Fuller D (2008) Millets and their role in early agriculture. Pragdhara 18:69-90.
103. Williams CT, Buck CL, Sears J, Kitaysky AS (2007) Effects of nutritional restriction on nitrogen and carbon stable isotopes in growing seabirds. Oecologia 153:11-18. https://doi.org/10.1007/s00442-007-0717-z
104. Wilson W, Milner J, Bulkan J, Ehlers P (2006) Weaning practices of the Makushi of Guyana and their relationship to infant and child mortality: A preliminary assessment of international recommendations. Am J Hum Biol 18:312-324. https://doi.org/10.1002/ajhb.20500
105. Wood JW, Milner GR, Harpending HC, Weiss KM, Cohen MN, Eisenberg LE, Hutchinson DL, Jankauskas R, Cesnys G, Katzenberg MA, Lukacs JR, McGrath JW, Roth EA, Ubelaker DH,

Wilkinson RG (1992) The osteological paradox: Problems of inferring prehistoric health from skeletal samples [and comments and reply]. Curr Anthropol 33:343-370. https://doi.org/10.1086/204084

106. Yovsi RD, Keller H (2003) Breastfeeding: An adaptive process. Ethos 31:147-171. https://doi.org/10.1525/eth.2003.31.2.147

Figure 1. Extent of Great Moravia under the rule of Mojmír I. (836–846), Rostislav (846–870) and Svatopluk I. (871–894), the location of the Mikulčice settlement agglomeration and other Early Medieval central places. Modified according to Poláček (2018)

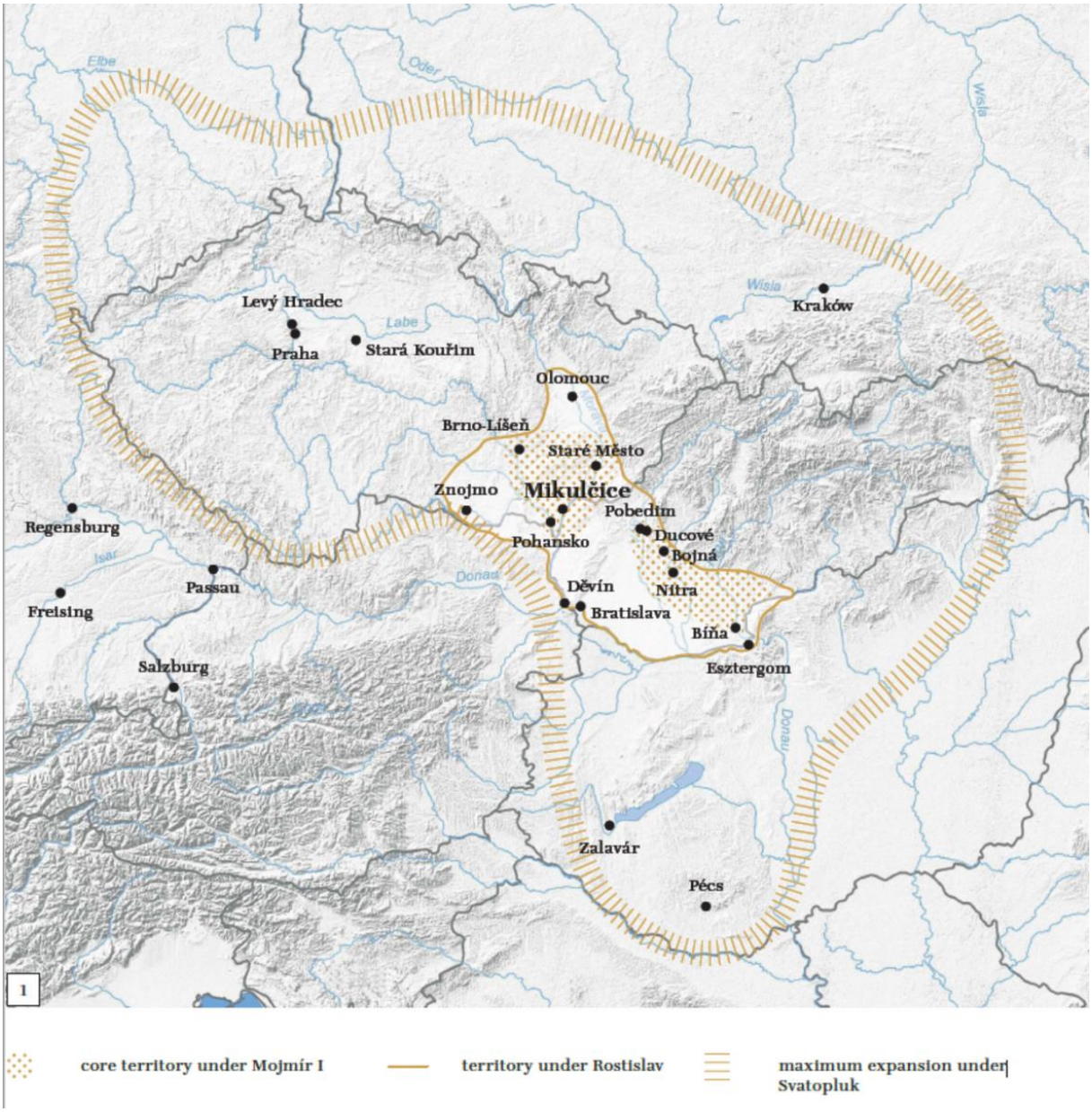
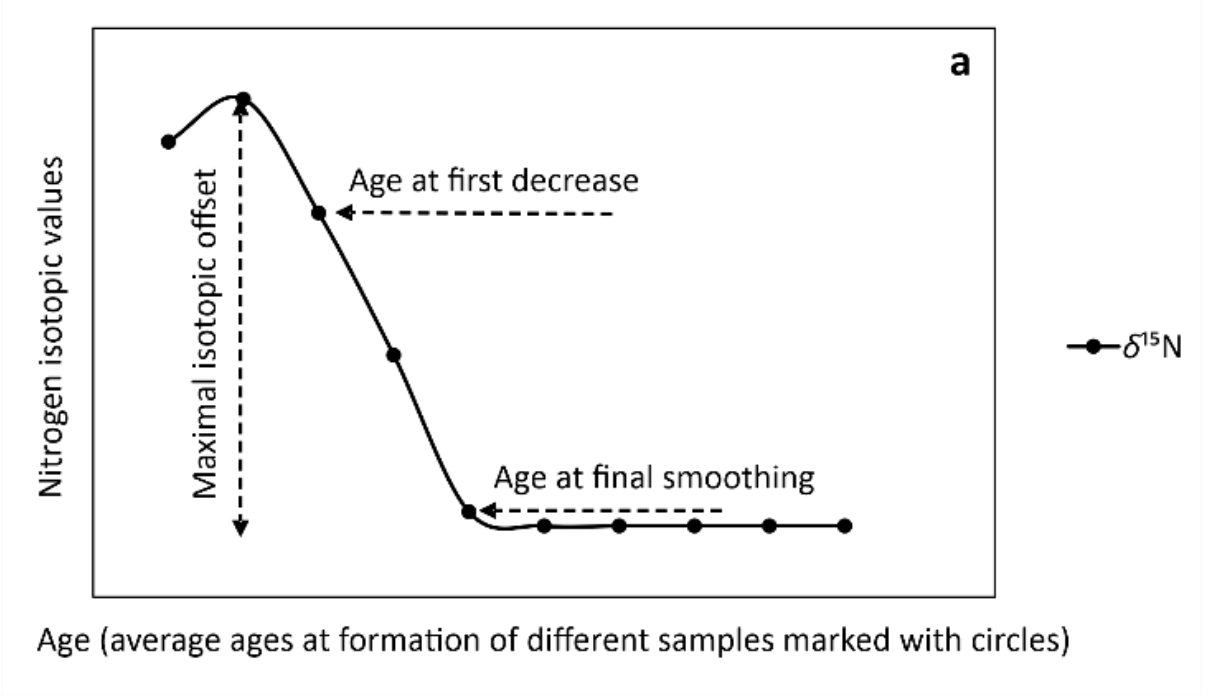


Figure 2. a: Theoretical weaning model with expected incremental nitrogen isotope data for an individual being weaned onto a diet similar to that of the mother. Note that 0.4 ‰was considered a cut-off value of significant isotopic change when defining the points marked on the isotopic curve, b: four categories of incremental carbon isotope data in relation to nitrogen isotopic data present in the studied dataset: i: positive covariance between δ 15N and δ 13C values, ii: negative covariance when δ 13C values are low during the time of peaking δ 15N values and subsequently increase along with decreasing δ 15N values. iii: flat carbon profile and finally iv: positive covariance between δ 15N and δ 13C values in most of the slices, but initial decrease in δ 13C values preceding the shift in δ 15N values



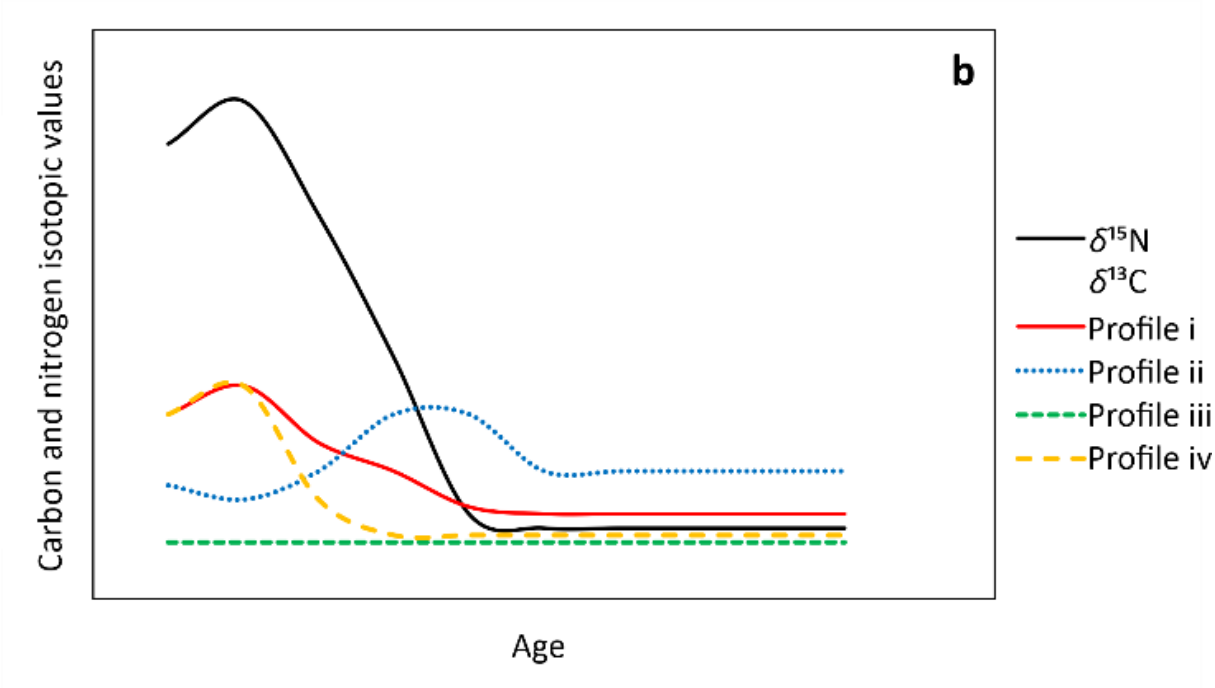


Figure 3. Examples of three individuals from the study cohort: Grave 286 shows a positive covariance between carbon and nitrogen isotopic values, typical for breastfeeding and subsequent weaning. Note the common patterns in the studied dataset: the presence of a small post-weaning dip in δ 15N values and the Δ 13Cmax higher than the maximal trophic level effect of exclusive breastfeeding (Fuller et al. 2006), Grave 112_VI shows an unexpected isotopic pattern with low nitrogen isotopic values in the first slice, suggesting a failure of breastfeeding with a subsequent rise in both carbon and nitrogen isotopic values of unknown origin, Grave 207 shows an isotopic pattern suggesting failure of breastfeeding with subsequent episodes of negative covariance between δ 15N and δ 13C values showing a typical picture of physiological stress as described by Beaumont and Montgomery (2016)

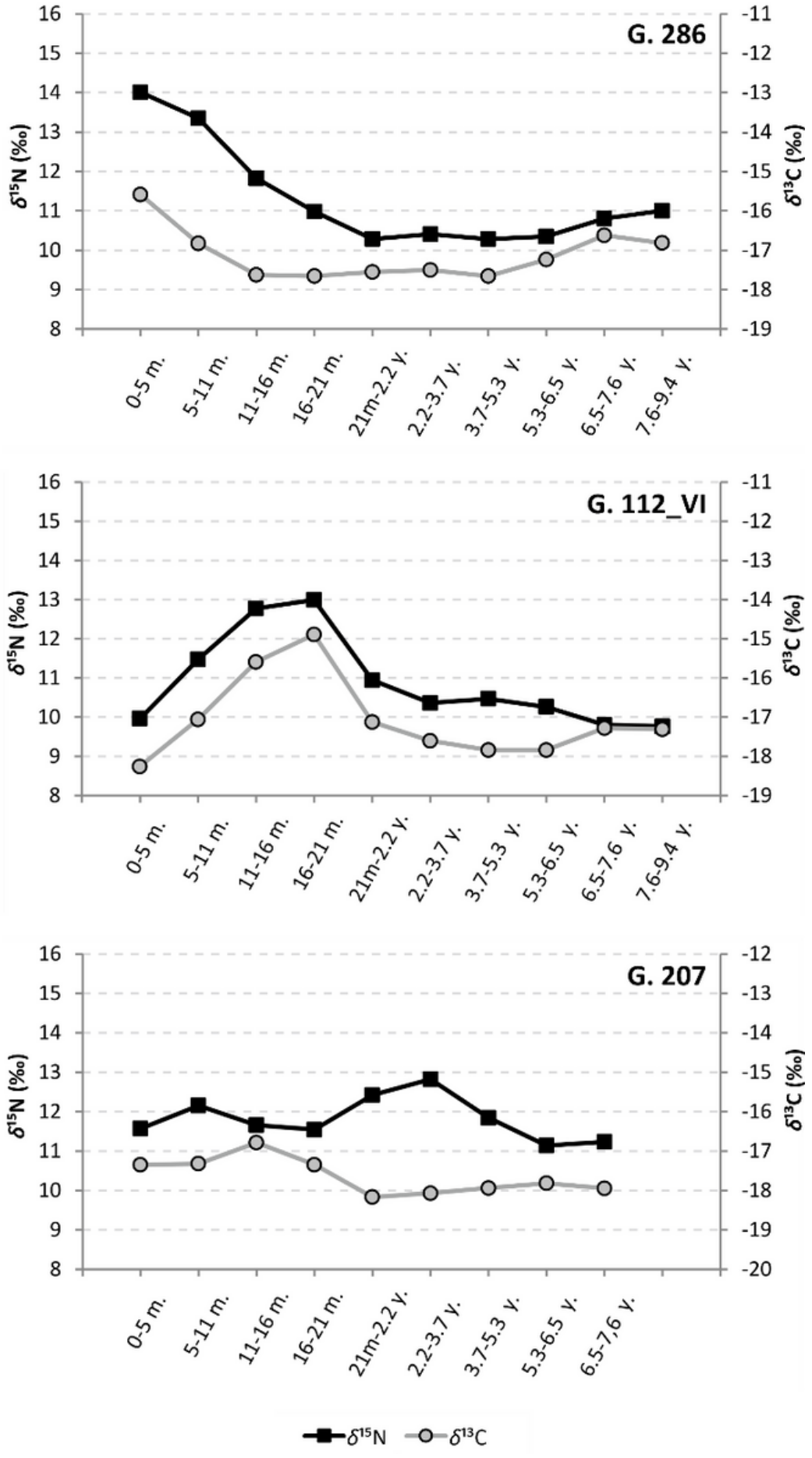

G. 286
G. 112_VI
G. 207
δ15N (‰)
δ13C (‰)
0-5 m.
5-11 m.
11-16 m.
16-21 m.
21m-2.2 y.
2.2-3.7 y.
3.7-5.3 y.
5.3-6.5 y.
6.5-7.6 y.
7.6-9.4 y.
6.5-7,6 y.
δ15N
δ13C

Figure 4. Typical nitrogen isotopic profiles for particular population groups (median values). To illustrate the shape of the isotopic curve while eliminating the confounding factor of underlying dietary variation among children and their lactating mothers, isotopic offsets from the highest δ 15N value within the isotopic profile were used to create these graphs. a: divided according to age-at-death, b: divided according to sex, c: divided according to socio-economic status

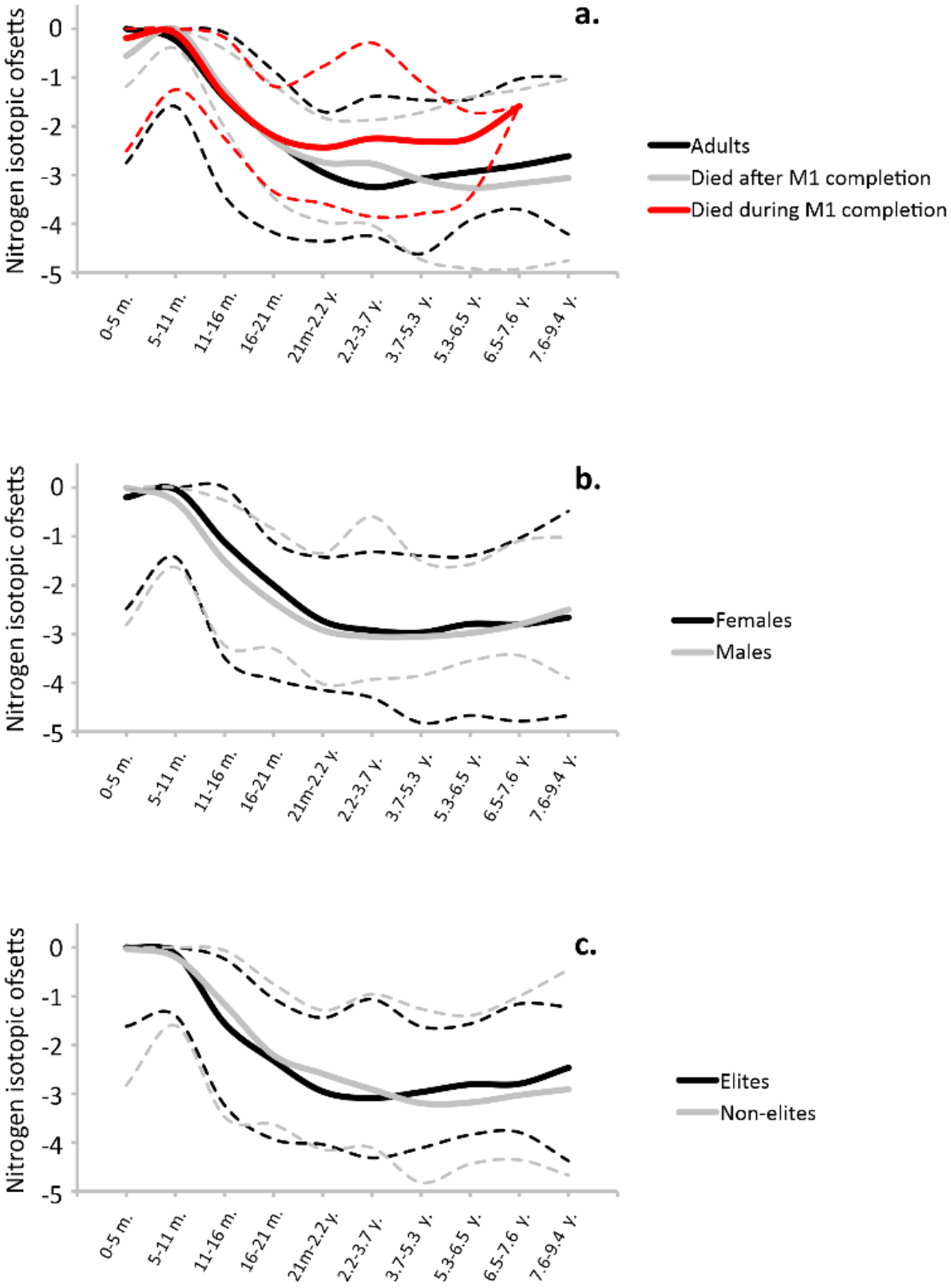
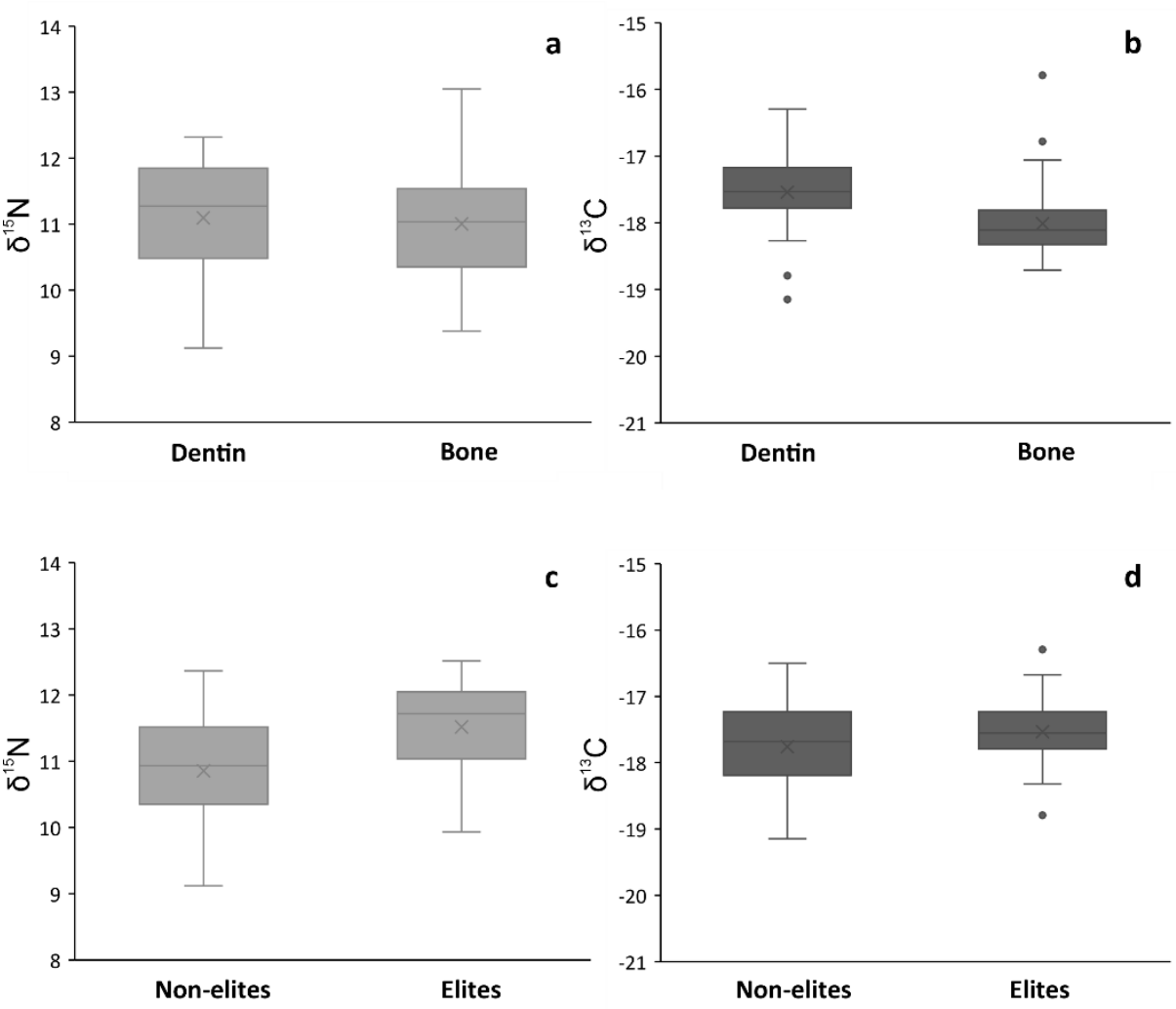

a.
Nitrogen isotopic ofsetts
0
-1
-2
-3
-4
-5
0-5 m.
5-11 m.
11-16 m.
16-21 m.
21m-2.2 y.
2.2-3.7 y.
3.7-5.3 y.
5.3-6.5 y.
6.5-7.6 y.
7.6-9.4 y.
Adults
Died after M1 completion
Died during M1 completion
b.
Nitrogen isotopic ofsetts
0
-1
-2
-3
-4
-5
0-5 m.
5-11 m.
11-16 m.
16-21 m.
21m-2.2 y.
2.2-3.7 y.
3.7-5.3 y.
5.3-6.5 y.
6.5-7.6 y.
7.6-9.4 y.
Females
Males
c.
Nitrogen isotopic ofsetts
0
-1
-2
-3
-4
-5
0-5 m.
5-11 m.
11-16 m.
16-21 m.
21m-2.2 y.
2.2-3.7 y.
3.7-5.3 y.
5.3-6.5 y.
6.5-7.6 y.
7.6-9.4 y.
Elites
Non-elites

Figure 5. Nitrogen and carbon post-weaning averages (defined as an average value from all the dentine slices formed after the termination of the initial decline in δ15N): in comparison with bone isotopic values of the same individuals (A, B), compared between socio-economic classes (C, D)



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.

Two methods based on sexual dimorphism, *i.e.* osteoarchaeology and genomics, which are commonly used for sex estimation, were reviewed. First, in osteoarcheology, sex estimation is based on three methods: visual, metric, and geometric-morphometric. Although these methods have the advantage of being non-destructive, their estimates are not absolute. Moreover, sex estimation is possible only in adulthood, and complete skeletons with a population reference, which considerably reduces the number of skeleton candidates for sex estimation, can lead to misclassifications. Second, in genomics, owing to the strong contamination and degradation of aDNA, sex estimation by discrimination of both forms of the amelogenin gene is not absolute. Additionally, as archaeological remains as rare and precious samples, few of them are available for biological analysis. aDNA analysis is not the method of choice to discriminate both sex-dependent genes because of its destructive and large sample-consuming character.

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.

Appendix A

Biological techniques

A.1 Polymerase chain reaction

The polymerase chain reaction (PCR) is used to amplify a small amount of a particular DNA sequence of interest (gene) based on cycles following three steps. First, the selected gene is denatured. The denaturation of the DNA double helix is achieved by breaking the hydrogen bonds between the complementary bases. Then, the annealing step consists of selecting two different and gene-specific primers that have an oligonucleotide sequence complementary to the 3′ and 5′ ends of the single-stranded gene. Finally, DNA polymerase extends the second complementary single strand using free deoxyribonucleotide triphosphates (dNTPs). Repeating this cycle quickly produces millions to billions of copies of this specific gene. After PCR is performed, the resulting DNA fragments can be analyzed by agarose gel electrophoresis. The precise temperature at each cycle step is a crucial parameter (Figure A.1).

A.2 Gel electrophoresis

Gel electrophoresis can be used for the separation of DNA fragments and proteins. Different types of gel can be used *i.e.* agarose, polyacrylamide (native PAGE), and sodium dodecylsulfate-polyacrylamide (SDS-PAGE). Agarose gel electrophoresis is commonly used for the separation of DNA fragments according to their sequence length under applied voltage (Figure A.2). This separation technique allows us to confirm the presence of PCR products with the expected size. Native-PAGE and SDS-PAGE, as far as they are concerned, are mostly used for protein separations. In native-PAGE, a pH gradient in the gel is formed using the applied high electric potential (10 kV). Proteins are separated according to their isoelectric point, in other words, the pH for which the proteins are in their uncharged form (Figure A.2). On the other hand, in SDS-PAGE, a moderate voltage (0.2 kV) is sufficient to separate proteins according to their molecular masses (Figure A.2). These two methods can also be combined for the separation of proteins in two dimensions.

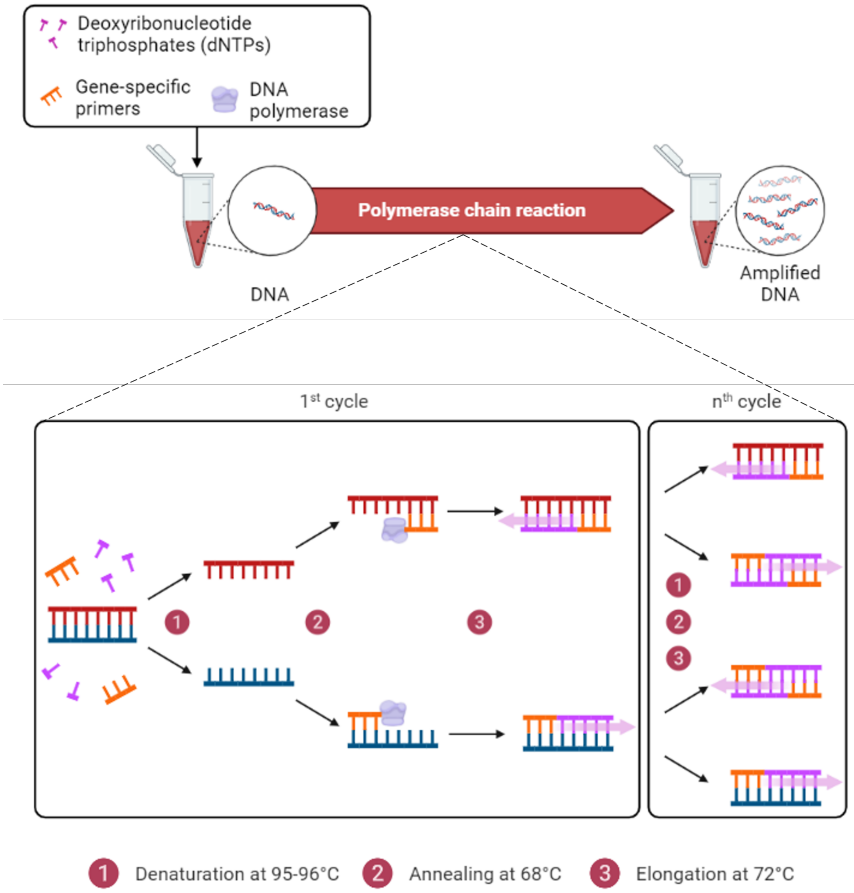


FIGURE A.1: Representation of PCR principle
Created with BioRender.com

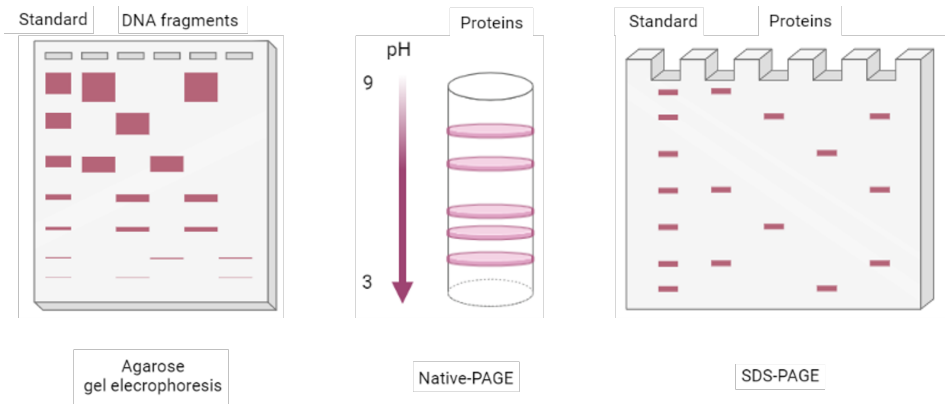


FIGURE A.2: Representation of different gel electrophoresis separations
Created with BioRender.com

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] W.-C. Lee, K. H. Lee, *Analytical biochemistry* **2004**, *324*, 1–10.

[15] M. Azarkan, J. Huet, D. Baeyens-Volant, Y. Looze, G. Vandenbussche, *Journal of Chromatography B* **2007**, *849*, 81–90.

[16] S. Fekete, A. Beck, J.-L. Veuthey, D. Guillarme, *Journal of pharmaceutical and biomedical analysis* **2015**, *113*, 43–55.

[17] J. Tang, M. Gao, C. Deng, X. Zhang, *Journal of Chromatography B* **2008**, *866*, 123–132.

[18] M. T. Davis, J. Beierle, E. T. Bures, M. D. McGinley, J. Mort, J. H. Robinson, C. S. Spahr, W. Yu, R. Luethy, S. D. Patterson, *Journal of Chromatography B: Biomedical Sciences and Applications* **2001**, *752*, 281–291.

[19] S. G. Valeja, L. Xiu, Z. R. Gregorich, H. Guner, S. Jin, Y. Ge, *Analytical chemistry* **2015**, *87*, 5363–5371.

[20] A. Goyon, A. Beck, O. Colas, K. Sandra, D. Guillarme, S. Fekete, *Journal of Chromatography A* **2017**, *1498*, 80–89.

[21] S. Fekete, A. Beck, J.-L. Veuthey, D. Guillarme, *Journal of pharmaceutical and biomedical analysis* **2014**, *101*, 161–173.

[22] W. Cai, T. Tucholski, B. Chen, A. J. Alpert, S. McIlwain, T. Kohmoto, S. Jin, Y. Ge, *Analytical chemistry* **2017**, *89*, 5467–5475.

[23] G. Brusotti, E. Calleri, R. Colombo, G. Massolini, F. Rinaldi, C. Temporini, *Chromatographia* **2018**, *81*, 3–23.

[24] H. Y. Ta, F. Collin, L. Perquis, V. Poinsot, V. Ong-Meang, F. Couderc, *Analytica chimica acta* **2021**, *1174*, 338233.

[25] S. Štěpánová, V. Kašička, *Journal of Separation Science* **2023**, 2300043.

[26] S. Štěpánová, V. Kašička, *Analytica Chimica Acta* **2022**, 339447.

[27] I. Mikšík, *Journal of separation science* **2019**, *42*, 385–397.

[28] V. Kašička, *Journal of Separation Science* **2022**, *45*, 4245–4279.

[29] C. Wang, X. Fang, C. S. Lee in *Capillary Electrophoresis of Biomolecules*, (Eds.: N. Volpi, F. Maccari), Humana Totowa, NJ, **2013**, pp. 1–12.

[30] S. Roca, L. Dhellemmes, L. Leclercq, H. Cottet, *ChemPlusChem* **2022**, *87*, e202200028.

[31] C. Huhn, R Ramautar, M Wuhrer, G. Somsen, *Analytical and bioanalytical chemistry* **2010**, *396*, 297–314.

[32] L. Leclercq, M. Morvan, J. Koch, C. Neusüß, H. Cottet, *Analytica Chimica Acta* **2019**, *1057*, 152–161.

[33] S. Roca, L. Leclercq, P. Gonzalez, L. Dhellemmes, L. Boiteau, G. Rydzek, H. Cottet, *Journal of Chromatography A* **2023**, *1692*, 463837.

[34] L. Dhellemmes, L. Leclercq, A. Höchsmann, C. Neusüß, J.-P. Biron, S. Roca, H. Cottet, *Journal of Chromatography A* **2023**, *1695*, 463912.

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

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

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

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

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

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

[41] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[56] 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.

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

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

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

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

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

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

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

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

[65] 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.

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

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

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

[69] 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.

[70] 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*.

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

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

[73] 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.

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

[75] A. R. Klales, *Sex estimation of the human skeleton: history, methods, and emerging techniques*, Elsevier Press: Waltham, MA, USA, **2020**, pp. 1–364.

[76] T. W. Phenice, *American journal of physical anthropology* **1969**, *30*, 297–301.

[77] T. D. White, *Human Osteology*, Academic Pr, St Louis, Missouri, USA, **1991**.

[78] B. A. Williams, T. L. Rogers, *Journal of forensic sciences* **2006**, *51*, 729–735.

[79] S. R. Kelley, S. D. Tallman, *Forensic Sciences* **2022**, *2*, 321–348.

[80] A. Del Bove, A. Veneziano, *Applied Sciences* **2022**, *12*, 9285.

[81] S. C. Koelzer, I. V. Kuemmel, J. T. Koelzer, F. Ramsthaler, F. Holz, A. Gehl, M. A. Verhoff, *Forensic science international* **2019**, *303*, 109929.

[82] T. Waldron in *Death, Decay and Reconstruction : Approaches to Archaeology and Forensic Science.* (Ed.: A. R. Klales), Elsevier Press: Waltham, MA, USA, **1987**, pp. 55–64.

[83] N. Sangchay, V. Dzetkuličová, M. Zuppello, J. Chetsawang, *Siriraj Medical Journal* **2022**, *74*, 330–339.

[84] J. Brůžek, F. Santos, B. Dutailly, P. Murail, E. Cunha, *American journal of physical anthropology* **2017**, *164*, 440–449.

[85] G. R. Dabbs, P. H. Moore-Jansen, *Journal of forensic sciences* **2010**, *55*, 149–152.

[86] T. Mello-Gentil, V. Souza-Mello, *Forensic Sciences Research* **2022**, *7*, 11–23.

[87] A. Kotěrová, J. Velemínská, J. Dupej, H. Brzobohatá, A. Pilnỳ, J. Brůžek, *International journal of legal medicine* **2017**, *131*, 251–261.

[88] E.-K. Oikonomopoulou, E. Valakos, E. Nikita, *International Journal of Legal Medicine* **2017**, *131*, 1731–1738.

[89] A. C. Stone, G. R. Milner, S. Pääbo, M. Stoneking, *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **1996**, *99*, 231–238.

[90] C. Afonso, D. Nociarova, C. Santos, C. Martinez-Labarga, I. Mestres, M. Duran, A. Malgosa, *American journal of human biology* **2019**, *31*, e23204.

[91] A. Poma, P. Cesare, A. Bonfigli, A. R. Volpe, S. Colafarina, G. Vecchiotti, A. Forgione, O. Zarivi, *Plos one* **2022**, *17*, e0269913.

[92] S. Sasaki, H. Shimokawa, *The International journal of developmental biology* **1995**, *39*, 127–133.

[93] A. K. Bansal, D. C. Shetty, R. Bindal, A. Pathak, *Journal of oral and maxillofacial pathology* **2012**, *16*, 395–399.

[94] A. Mannucci, K. M. Sullivan, P. L. Ivanov, P. Gill, *International journal of legal medicine* **1994**, *106*, 190–193.

[95] F. A. Kaestle, K. A. Horsburgh, *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **2002**, *119*, 92–130.

[96] A. Mittnik, C.-C. Wang, J. Svoboda, J. Krause, *PloS one* **2016**, *11*, e0163019.

[97] F Francès, O Portolés, J. González, O Coltell, F Verdú, A Castelló, D Corella, *Clinica Chimica Acta* **2007**, *386*, 53–56.

[98] M. Arnay de la Rosa, E. González-Reimers, R. Fregel, J. Velasco-Vázquez, T. Delgado-Darias, A. M. González, J. M. Larruga, *Journal of archaeological science* **2007**, *34*, 1515–1522.

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

[100] 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.

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

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

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

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

[105] 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.

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

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

[108] 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.

[109] 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.

[110] 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.

[111] 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.

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

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

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

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

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

[117] 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.

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

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

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

[121] 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.

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