

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/57622>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

FT-ICR-MS Approach to Monitor
Asparagine Deamidation and its Isomers
Products in Collagen from Ancient
Bones

Pilar Perez Hurtado MSc.

A thesis submitted for the degree of Doctor of Philosophy

University of Warwick
Department of Chemistry

July 2013

Table of Contents

List of Figures.....	v
List of Tables	xiii
Acknowledgements.....	xv
Declaration.....	xvii
Abstract.....	xviii
Abbreviations	xix
Chapter One: Introduction	1
1.1. Archaeological Importance of Deamidation	2
1.2. Introduction to Peptide and Protein Deamidation.....	3
1.2.1. Mechanism of Deamidation in Peptides and Proteins.....	5
1.2.2. Mass Spectrometric Methods for Studying Deamidation	6
1.3. Mass Spectrometry.....	8
1.3.1. Ionisation Sources	9
1.3.2. Mass Analysers	16
1.3.3. Fourier Transform-Ion Cyclotron Mass Spectrometry	18
1.3.4. Fragmentation Methods Used for Structural Analysis of Proteins and Peptides in Mass Spectrometry	26
1.4. Differentiation of Isomeric Amino Acid Residues in Proteins and Peptides Using Mass Spectrometry	34
1.4.1. High-energy CAD and hot ECD applied in the differentiation of Ile from Leu.....	38

1.4.2. Use of high-energy CAD, ECD, and ETD to differentiate aspartic from isoaspartic acid in proteins and peptides	42
1.4.3. Differentiation of α -glutamic acid and γ -glutamic acid residues in peptides	50
1.5. Aims and Objectives	54
1.6. Research Papers	55
1.7. References	56
Chapter 2: Development of a methodology to monitor deamidation of collagen I	67
2.1. Introduction	68
2.2. Experimental Methodology	71
2.2.1. Materials.....	71
2.2.2. Digestion of collagen	71
2.2.3. FT-ICR-MS measurements	72
2.3. Results and Discussion.....	73
2.3.1. Peptide Characterization	73
2.3.2. Identification of Deamidation Sites in Collagen by Using FT-ICR-MS	79
2.3.3. Deamidation Rates	86
2.3.4. Differentiation of IsoAsp from Asp.....	91
2.4. Conclusions	93
2.5. References	94
2.6. Mass Accuracy and Peptide Assignments.....	99
Chapter 3: FT-ICR-MS Approach to monitor asparagine deamidation in collagen from archaeological material	113

3.1. Introduction	114
3.2. Experimental Methodology	116
3.2.1. Materials.....	116
3.2.2. Collagen extraction from bovine bones obtained from a Coventry (UK) butcher	115
3.2.3. Digestion of collagen	118
3.2.4. Collagen extraction from archeological bones.....	118
3.2.5. FT-ICR-MS measurements	118
3.3. Results and Discussion	119
3.3.1. Characterization of marker peptides from bovine collagen	119
3.3.2. Identification of deamidation sites in collagen extracted from archeological bones using FT-ICR-MS	122
3.3.3. Determination of the reaction rates for deamidation reaction.....	126
3.3.4. Estimation of the Activation Energy.....	128
3.3.5. Effect of digestion treatment on deamidation	130
3.3.6. Effect of acidic treatment on deamidation	131
3.3.7. Effect of pH and buffers.....	132
3.4. Conclusions	133
3.5. References	134
3.6. Mass Accuracy and peptide Assignments	137
Chapter 4: Monitoring deamidation and differentiation of the reaction products in ancient collagen using ECD	143
4.1. Introduction	144
4.1.1. Non-MS Methods to determine IsoAsp residues	144
4.1.2. MS methods to determine IsoAsp from Asp residues.....	147

4.2. Experimental Methodology	150
4.2.1. Materials.....	150
4.2.2. Collagen extraction from archeological bones.....	150
4.2.3. FT-ICR-MS measurements	153
4.3. Results and Discussion.....	154
4.3.1. Identification of collagen peptides originated from ancient cow bones	154
4.3.2. Differentiation of Asp from IsoAsp in collagen using ECD.....	158
4.4. Conclusions	164
4.5. References	165
4.6. Mass Accuracy and Peptide Assignments.....	168
Chapter 5: Conclusions and future directions	183
5.1. Conclusions	184
5.2. Future Directions	186
APPENDIX: Use of High Resolution Mass Spectrometry in the Analysis of Polymeric Excipients in Drug Delivery Formulations.....	187

List of Figures

FIGURE 1.1: A) Isotopic distribution of the combination of non-deamidated (N) and deamidated (D) species. B) Deconvoluted isotopic distribution of the N. C) Deconvoluted isotopic distribution of D, showing the shift of 0.984 Da	7
FIGURE 1.2: Simulation of the isotopic patterns corresponding to non-deamidated (black) and deamidated species (red).....	8
FIGURE 1.3: Schematic representation of the 12 T Bruker SolariX system. Adapted with permission from Bruker Daltonik GmbH	9
FIGURE 1.4: Electron ionisation source. Adapted from O'Connor, 2010.....	10
FIGURE 1.5: Process of ionisation in FAB.....	11
FIGURE 1.6: Chemical structure for alpha cyano hydroxycinnamic acid and 2,5-dihydroxy benzoic acid (DHB)	12
FIGURE 1.7: Representation of the MALDI ionisation process.....	13
FIGURE 1.8: Electrospray ionisation process. Figure adapted from Kebarle, P. A Brief Overview of the Present Status of the Mechanisms Involved in Electrospray Mass Spectrometry. <i>Journal of Mass Spectrometry</i> . 2000 , <i>35</i> , 804-817	15
FIGURE 1.9: Representation of a quadrupole mass analyser.....	17
FIGURE 1.10: Combined Linear/Reflectron MALDI-TOF mass spectrometer Reprinted with permission from O'Connor	18
FIGURE 1.11: Representation of the process of excitation and detection of the ions in the ion cyclotron resonance (ICR) cell. Reproduced with permission of Bruker Daltonik GmbH.....	19
FIGURE 1.12: Diagram of the “Infinite cell”. Red arrow shows the direction of the magnetic field. Figure was reproduced with the permission of Bruker Daltonik GmbH.....	21

FIGURE 1.13: Direction of the Lorentz force due to the presence of a magnetic field perpendicular to the plane of the paper	22
FIGURE 1.14: Ion motion in the ICR cell. The axis of the magnetic field is parallel to the plane of the paper. Reprinted from Kolhinen, V.S., et al. Commissioning of the double Penning trap system MLLTRAP. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2009, 600, 391-397, with permission from Elsevier	24
FIGURE 1.15: Fourier transformation of the time domain signal into frequency domain signal. Adapted from P. B. O'Connor, 2010.....	26
FIGURE 1.16: ECD spectrum of two peptides BUSM1 and BUSM2, the shaded regions represent the $c^{\bullet}+58$ corresponding to the presence of isoAsp. Reprinted with permission from Cournoyer JJ, Pittman JL, Ivleva VB, Fallows E, Waskell L, Costello CE, O'Connor PB. 2005. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. Protein Science 14:452-463. Copyright © 2005 John Wiley and Sons	48
FIGURE 1.17: ECD spectrum of two peptides BUSM1 and BUSM2, the shaded regions represent the $z-57$ corresponding to the presence of isoAsp. Reprinted with permission from Cournoyer JJ, Pittman JL, Ivleva VB, Fallows E, Waskell L, Costello CE, O'Connor PB. 2005. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. Protein Science 14:452-463. Copyright © 2005 John Wiley and Sons	49
FIGURE 2.1: Representation of the structure of collagen type I. Adapted from http://atlasgeneticsoncology.org/Genes/COL1A2ID411ch7q22.html . Accessed on: May 17 th 2013	68

FIGURE 2.2: Sequence structure of collagen type 1. Figure adjusted from protein data bank: <http://www.rcsb.org/pdb/101/motm.do?momID=4>. Accessed on: May 17th 2013..... 74

FIGURE 2.3: Mass spectrum of collagen tryptic digest. Peptides marked with **N** (asparagines) were selected for deamidation studies. **P** represents Hyp. Insert a) shows the isotopic pattern for the non-deamidated, m/z [696.84696]⁴⁺, and the deamidated peptide, m/z [697.09326]⁴⁺, where the triangles represent the non-deamidated peptide and the circles represent the deamidated peptide. Insert b) shows the theoretical isotopic distribution for the peptides..... 78

FIGURE 2.4: ECD of tryptic digest peptide [coll_I (α 2) -1022-1051]⁴⁺ ion at m/z 697.09326. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Insert shows the charge reduced species at m/z 929.4584. (*) represents noise peaks and (?) represents unknown peaks, (\blacktriangle) represents a different peptide present in the mixture..... 80

FIGURE 2.5: ECD of tryptic digest peptide [coll_I (α 1) 1152-1167]²⁺ ion at m/z 781.89408. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Insert show the region where IsoAsp diagnostic ion appears (z_{13}^{\bullet} -57, m/z 1204.63171) indicating position 13, and the lost CO₂ from z_{16} (z_{16} -44, m/z 1502.77117) represents the diagnostic ion for the presence of aspartic acid in position 16. Ions labelled with blue are originated from another peptide present in the digestion mixture at m/z 780.90998..... 81

FIGURE 2.6: ECD of tryptic digest peptide [coll_I (α 1) 397-414]²⁺ ion at 793.88150 m/z . Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. The insert show the region where the Asp diagnostic ion (z_{16} -57, m/z 1385.68082) appears, this one correspond to position 3 (N-terminal), 16 (C-

terminal) in the peptide. Ions labelled with purple and brown are originated from other two peptides present in the digestion mixture with m/z 790.88673 and m/z 795.90975..... 82

FIGURE 2.7: CAD of tryptic digest peptide [coll_I (α 2) -1022-1051]⁴⁺ ion at m/z 697.09326. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site 83

FIGURE 2.8: CAD of tryptic digest peptide [coll_I(α 1) -1152-1167]²⁺ ion at m/z 781.89408. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Ions labelled with blue are originated from another peptide present in the digestion mixture at m/z 780.90998..... 84

FIGURE 2.9: CAD of tryptic digest peptide [coll_I (α 1) 397-414, coll_I(α 1)]²⁺ ion at m/z 793.88150. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site 85

FIGURE 2.10: Snapshot of the deamidation reaction of peptides at 80 min after denatured time at 62 °C. A. [coll_I α 2) -1022-1051]⁴⁺ ion at m/z 697.09326, B. [coll_I (α 1) -1152-1167]²⁺ ion at m/z 781.89408, and C. [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150. Triangles represent the non-deamidated peptides and the circles represent the deamidated peptides. Inserts to the right are the theoretical isotopic distributions. Inserts to the left are expansion of the A+1 isotope of the non-deamidated and the A isotope of the deamidated species 88

FIGURE 2.11: Exponential decay curves of the non-deamidated peptide at 62 °C. A. [coll_I (α 2) -1022-1051]⁴⁺ ion at m/z 697.09326, B.[coll_I (α 1) -1152-1167]²⁺ ion at m/z 781.89408, and C. [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150. The exponential coefficients represent the rate constants of deamidation for each peptide.

The intercept at the y axis represents the extent of deamidation before sample preparation in each case 89

FIGURE 3.1: Pieces of bovine bone obtained from a Coventry UK butcher..... 117

FIGURE 3.2: Size of the bones after hammering 117

FIGURE 3.3: Resulting bone powder obtained from the grinding process..... 117

FIGURE 3.4: a) Mass spectrum of tryptic digest of collagen extracted from contemporary cow bones. **N/D** represents deamidation site. **P** represents Hyp. The inserts show three possible deamidated peptides corresponding to m/z [697.09326]⁴⁺, m/z [781.89408]²⁺, and m/z [793.88133]²⁺. b) Tryptic digest collagen extracted from 1,050 year-old bone. Inserts (1) and (2) show the experimental and theoretical isotopic distributions respectively..... 120

FIGURE 3.5: ECD-IRMPD of tryptic digest peptide [coll_I (α 1) 1152-1167]²⁺ ion at m/z 781.89408. The deamidation is located in position 4 from the N-terminus (**D**). The P highlighted in green represent the modification of hydroxylation. Ions labelled with superscript (\diamond) originate from another peptide present in the digestion mixture at m/z 780.90998 123

FIGURE 3.6: ECD-IRMPD of tryptic digest peptide [coll_I (α 1) 397-414]²⁺ ion at m/z 793.88150. The deamidation site is located in position 3 from the N-terminus (**D**), in the peptide. Prolines highlighted in green represent the modification of hydroxylation. Ions labelled with superscript (\ddagger) are originated from other peptide present in the digestion mixture with m/z 790.88678 124

FIGURE 3.7: CAD of tryptic digest peptide [coll_I(α 1) -1152-1167]²⁺ ion at m/z 781.89408. Red colour represent the deamidation site, green indicates the presence of Hyp..... 124

FIGURE 3.8: CAD of tryptic digest peptide [coll_I (α 1) 397-414, coll_I(α 1)]²⁺ ion at m/z 793.88150. Red colour represent the deamidation site, green indicates the presence of Hyp..... 125

FIGURE 3.9: Exponential decay curves of the non-deamidated peptide at 62 °C. A. [coll_I (α 1) -1152-1167]²⁺ ion at m/z 781.89408, and B. [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150. The exponential coefficients represent the rate constants of deamidation reaction for each peptide. The intercept represents the extent of deamidation in each case..... 127

FIGURE 3.10: Exponential decay curves of the non-deamidated peptide [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 at A) 50 °C and B) 62 °C. The exponential coefficients represent the rate constants of deamidation reaction. The intercept represents the extent of deamidation in each case 130

FIGURE 3.11: Exponential decay curves of the non-deamidated peptide at 62 °C. A, B. Effect of digestion conditions on [coll_I (α 1) -1152-1167]²⁺ ion at m/z 781.89408, and [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 respectively. C, and D. Effect of acidic treatment on [coll_I (α 1) -1152-1167]²⁺ ion at 781.89408 m/z , and [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 respectively 131

FIGURE 4.1: Samples labeled with number 8 (left) and number 9 (right). Both are bovine phalanx from Cooldrinagh, Dublin. Age according to radio carbon date is (915±50) and (1235±30) years respectively 151

FIGURE 4.2: Samples labeled with number 10 (left) and number 11 (right). Both samples are bone from bovine, Ness of Brodgar, Scotland. Age according to radio carbon date is (3829±27) and (3878±26) years old respectively 151

FIGURE 4.3: Samples labeled with number 12 (left) and number 13 (right). Both are bone from Bos Taurus, Magura. Age according to radio carbon date is (6354±37) and (6484±37) years old respectively 151

FIGURE 4.4: Samples labeled with number 14 (left) and number 15 (right). Both are bone from Bovini, La Ferrassie, France. Age according to radio carbon date is (31400±400) and (32100±450) years old respectively 152

FIGURE 4.5: Samples labeled with number 16 (left) and number 17 (right). Both are bone from Bovini, La Ferrassie, and La Chauverie, France. Age according to radio carbon date is (32550±450) and (50000±3900) years old respectively 152

FIGURE 4.6: Powder bone obtained from the grinding process 152

FIGURE 4.7: a) to i) Enlargement of the m/z [793.88153]²⁺ ion present in the spectrum, which corresponds to a deamidated peptide of collagen extracted from archaeological bones. The age of the bones used in the experiment are shown in the figure 155

FIGURE 4.8: a) to f): Enlargement of the m/z [829.74176]³⁺ ion present in the spectrum, which corresponds to a deamidated peptide of collagen extracted from cattle bone. The figure shows the age of the bone used 156

FIGURE 4.9: a) to h): Enlargement of the m/z [927.78113]³⁺ ion present in the spectrum, which correspond to a deamidated peptide of collagen extracted from bone, which ages are shown in the figure 157

FIGURE 4.10: Mass spectrum of tryptic digest of collagen extracted from 3,000 year-old cow bones. **N/D** represents deamidation site. **P** represents Hyp. The inserts show three possible deamidated peptides corresponding to m/z [793.88153]²⁺, m/z [927.78113]³⁺, and m/z [829.74176]³⁺. a) and b) are real and simulated isotopic distributions respectively 158

FIGURE 4.11: a) Narrowband spectrum corresponding to m/z [793.88153]²⁺ ion in a mixture of the non-deamidated and the deamidated form. The insert represents the zoom of the peak where the two isotopes should appear. b) Narrowband spectrum of the deamidated form of the peptide m/z [793.88153]²⁺ 160

FIGURE 4.12: Up: ECD and Down: CAD of tryptic digest peptide [coll_I (α1) 397-414] ²⁺ ion at m/z 793.88150. The deamidation site is located in position 3 (N-terminal), in the peptide (N). Diagnostic ion corresponding to the deamidation at position 3 is labelled as z_{16-57}^{\bullet} . P in green represent the modification of hydroxylation 161

FIGURE 4.13: Up: ECD and Down: CAD of tryptic digest peptide [coll_I (α1) 1141-1167] ³⁺ ion at m/z 829.74184. The deamidation site is located in position 15 (N-terminal) labelled as (N), in the peptide. Diagnostic ion corresponding to the deamidation site is labelled as c_{14+58}^{\bullet} in the ECD spectrum. P in green represent the modification of hydroxylation..... 162

FIGURE 4.14: Up: ECD and Down: CAD of tryptic digest peptide [coll_I (α2) 674-706]³⁺ ion at m/z 927.78145. The deamidation site is located in position 15 (N-terminal), labelled as (N) in the peptide. P highlighted in green represent the modification of hydroxylation..... 163

List of Tables

TABLE 2.1: Tryptic peptides of Collagen I. Confirmed using CAD	76
TABLE 2.2: Deamidation rates and extent of deamidation of collagen peptides at 62 °C	90
TABLE 2.3: Peak list generated from the externally calibrated CAD spectrum of [GHNGLQGL <u>P</u> GLAGHHGDQGA <u>P</u> GAVGPAGPR+4H] ⁴⁺ <i>m/z</i> 697.09326.....	100
TABLE 2.4: Peak list generated from the ECD spectrum of [GHNGLQGL <u>P</u> GLAGHHGDQGA <u>P</u> GAVGPAGPR+4H] ⁴⁺ <i>m/z</i> 697.09326	102
TABLE 2.5: CAD peak table [DGLNGL <u>P</u> GP <u>I</u> G <u>P</u> PGPR+2H] ²⁺ <i>m/z</i> 781.89408... 105	
TABLE 2.6: Peak-list generated from the ECD spectrum of [DGLNGL <u>P</u> GP <u>I</u> G <u>P</u> PGPR+2H] ²⁺ <i>m/z</i> 781.89408.....	107
TABLE 2.7: Peak list generated from the externally calibrated CAD spectrum of [GANGA <u>P</u> GIAG <u>A</u> PGF <u>P</u> GAR] ²⁺ <i>m/z</i> 793.88150	109
TABLE 2.8: Peak list generated from the ECD spectrum of [GANGA <u>P</u> GIAG <u>A</u> PGF <u>P</u> GAR] ²⁺ <i>m/z</i> 793.88150	110
TABLE 3.1: Assignments of peptides in archeological bone material.....	121
TABLE 3.2: Deamidation rates and grade of deamidation of collagen peptides from young collagen and ancient collagen at 62°C	127
TABLE 3.3: CAD peak table [DGLNGL <u>P</u> GP <u>I</u> G <u>P</u> PGPR+2H] ²⁺ <i>m/z</i> 781.89408... 137	
TABLE 3.4: Peak list generated from the ECD spectrum of [DGLNGL <u>P</u> GP <u>I</u> G <u>P</u> PGPR+2H] ²⁺ <i>m/z</i> 781.89408.....	138
TABLE 3.5: Peak list generated from CAD spectrum of [GANGA <u>P</u> GIAG <u>A</u> PGF <u>P</u> GAR] ²⁺ <i>m/z</i> 793.88150	139

TABLE 3.6: Peak list generated from the ECD spectrum of [GANGA <u>P</u> GIAGA <u>P</u> GF <u>P</u> GAR] ²⁺ <i>m/z</i> 793.88150	140
TABLE 3.7: Internal calibrants used in the MS mode.....	142
TABLE 4.1: Molecular formulae and ion masses of the peptides assigned across the range of samples analyzed	168
TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples	170
TABLE 4.3: List of assigned ions from the ECD spectrum of <i>m/z</i> [793.88150] ²⁺ .	176
TABLE 4.4: List of assigned ions from the CAD spectrum of <i>m/z</i> [793.88150] ²⁺ .	177
TABLE 4.5: List of assigned ions from the ECD spectrum of <i>m/z</i> [829.74184] ³⁺ .	178
TABLE 4.6: List of assigned ions from the CAD spectrum of <i>m/z</i> [829.74184] ³⁺ .	179
TABLE 4.7: List of assigned ions from the ECD spectrum of <i>m/z</i> [927.78145] ³⁺ .	181
TABLE 4.8: List of assigned ions from the CAD-MS ² spectrum of <i>m/z</i> [927.78145] ³⁺	182

Acknowledgements

I would like to especially thank my supervisor Prof. Peter B. O'Connor for providing me with the opportunity to undertake this PhD, for his great interest and assistance in the pursuit of this research and in the preparation of this thesis. Thank you to the EPSRC for providing financial support for this work.

I wish to express my gratitude to Dr. Mark Barrow, who has been patient, collaborative and an excellent guide throughout my journey in the lab. I would also like to express my gratitude to Dr. David Kilgour for his guidance over the past 2 years of collaborative work. Additionally, I would like to thank my previous committee members: Dr. Giovanni Costantini and Dr. Alison Rodger for their valuable advice during my PhD preparation. My thanks go to Julia Smith from Bruker for her help and support during my studies. Big thanks to Jason Noone for all his IT support during the past 3 years. No thesis or papers would have been possible without his help.

Especial thanks to Yuko Lam, and Adam Mbugi for their help in conducting vital experiments for the development of this thesis.

I am indebted to my student colleagues and friends for providing a stimulating and fun environment in which to learn and grow. I am especially grateful to my dear friend Andrea Clavijo-Lopez, Rebecca Wills, Yulin Qi (who always makes me

laugh), Juan Wei, Andrew Soulby, Maria Van Agthoven, and Chris Wootton, for their friendship and support during my time in the office.

My appreciation also goes to the Department of Chemistry, as well as the University of Warwick for allowing me to study here and for their incredible support in helping me develop professional and personal skills.

I extend my thanks to all my close friends who have been supportive and made my life happy over my time in Coventry. A special thanks to Maria Tareen, Justina Walls, Kevin Walls, and Carlos Duque-Daza. Thank you all for being such nice company throughout my journey.

I am grateful to people who has helped me settle in this country and supported me during my studies. Thank you Pearl for your support and love. Very special thanks to my boyfriend (Mick), who has provided me with infinite support, has always believed in my abilities to achieve my dreams and has coped with my ups and downs in pursuing my PhD. Finally my biggest thank you goes to my mother Janette, my aunt Amina, my sister Victoria and family who have always supported me and encouraged me over the years.

I especially want to dedicate this thesis to my grandparents (Francisco Hurtado and Mercedes Hurtado). Nothing would have been possible without the early support that you two have provided me since I was a baby. Thank you for being there for me all the time. I love you both with all my heart.

Declaration

I hereby declare that this thesis, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy and titled “FT-ICR-MS Approach to Monitor Asparagine Deamidation and its Isomers Products in Collagen from Ancient Bones”, represents my own work and has not been previously submitted to this or any other institution for a degree, diploma or other qualification. The thesis includes published work that was edited from the following articles:

Perez, P., O'Connor, P. B., Differentiation of Isomeric Amino Acid Residues in Proteins and Peptides Using Mass Spectrometry. *Mass Spectrometry Reviews*. **2012**, *31*. 609-625.

Perez, P., O'Connor, P. B., Deamidation of Collagen. *Analytical Chemistry*. **2012**, *84*. 3017-3025.

Perez P., Lam P., Kilgour D., Bristow A., McBride E., O'Connor P., Use of High Resolution Mass Spectrometry for the Analysis of Polymeric Excipients in Drug Formulations. *Analytical Chemistry*. **2012**. *84* (20), 8579–8586

Also some of the work presented in the thesis has been undertaken by the collaborators, thus is explicitly stated where appropriate.

Pilar Perez Hurtado

July 2013

Abstract

This project investigates the use of ultrahigh resolution mass spectrometry along with fragmentation techniques such as electron capture dissociation (ECD) and collisionally activated dissociation (CAD) to study deamidation of ancient bone, by using potential deamidation markers present in bovine collagen standards. With the application of these techniques, several marker peptides present in the digested protein standard of bovine collagen were successfully assigned. The sequences of these peptides correlated well with the reported sequences for bovine collagen in the literature.

FT-ICR-MS was used to monitor deamidation of collagen by following a shift of +0.948 Da in the spectrum, resulting in a mass difference of 19 mDa from the ^{13}C of the non-deamidated form and the ^{12}C of the deamidated form, which can be difficult to assign due to overlap with the ^{13}C isotopic distribution in peptides. The rate constants for the deamidation reaction were calculated, and the extent of deamidation before sample handling was determined. The methodology developed was then applied to collagen extracted from real bone samples, both modern and ancient, proving to be a useful method for monitoring asparagine deamidation before sample preparation. Differentiation of the isomers products of deamidation (aspartic and isoaspartic acid) were successfully assigned (where possible) using the diagnostic ions originated from their ECD spectra.

Abbreviations

A β	β -Amyloid
ABC	Ammonium bicarbonate
AC	Alternating current
ACN	Acetonitrile
AD	Alzheimer's disease
AdoMet	S-adenosyl-L-methionine
Ala	Alanine
Asn	Asparagine
Asp	Aspartic Acid
CAD	Collisionally Activated Dissociation
CID	Collisionally Induced Dissociation
C $_{\alpha}$	α -Carbon
C $_{\beta}$	β -Carbon
D	Aspartic Acid
D-	Dextrorotatory
Da	Dalton
DHB	2,5-Dihydroxy Benzoic Acid
E	Glutamic Acid
ECD	Electron Capture Dissociation
EI	Electron Ionisation
ESI	Electrospray Ionisation
ETD	Electron Transfer Dissociation
eV	Electron volts

FAB	Fast Atom Bombardment
FT	Fourier Transform
FWHM	Full width at half maximum
Glu	Glutamic Acid
Gln	Glutamine
HECD	Hot Electron Capture Dissociation
HE CAD	High-energy Collisionally Activated Dissociation
HP-LC	High Performance Liquid Chromatography
His	Histidine
Hyp	Hydroxyproline
LCMS	Liquid Chromatography Mass Spectrometry
ICR	Ion Cyclotron Resonance
Ile	Isoleucine
IRMPD	Infrared Multiphoton Dissociation
IsoAsp	Isoaspartic Acid
L-	Levorotatory
LC	Liquid Chromatography
Leu	Leucine
LiAlH ₄	Lithium aluminum hydride
M	Molecular mass
[M+H] ⁺	Protonated molecular ion
MeOH	Methanol
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MALDI	Matrix-Assisted Laser Desorption Ionisation

<i>m/z</i>	Mass-to-charge ratio
PIMT	Protein L-isoaspartyl- <i>O</i> -methyltransferase
Pro	Proline
PTMs	Post-translational Modifications
Q	Glutamine
RF	Radio Frequency
RP	Reversed Phase
S/N	Signal-to-noise ratio
SWIFT	Stored Waveform Inverse Fourier Transform
T	Tesla
TOF	Time of Flight
TOF-MS	Time of Flight Mass Spectrometry
Trp	Tryptophan
Thr	Threonine
Tyr	Tyrosine
UV	Ultraviolet Radiation
VUV	Vacuum Ultraviolet Radiation
UW	Utah-Washington mechanism

Chapter 1: Introduction

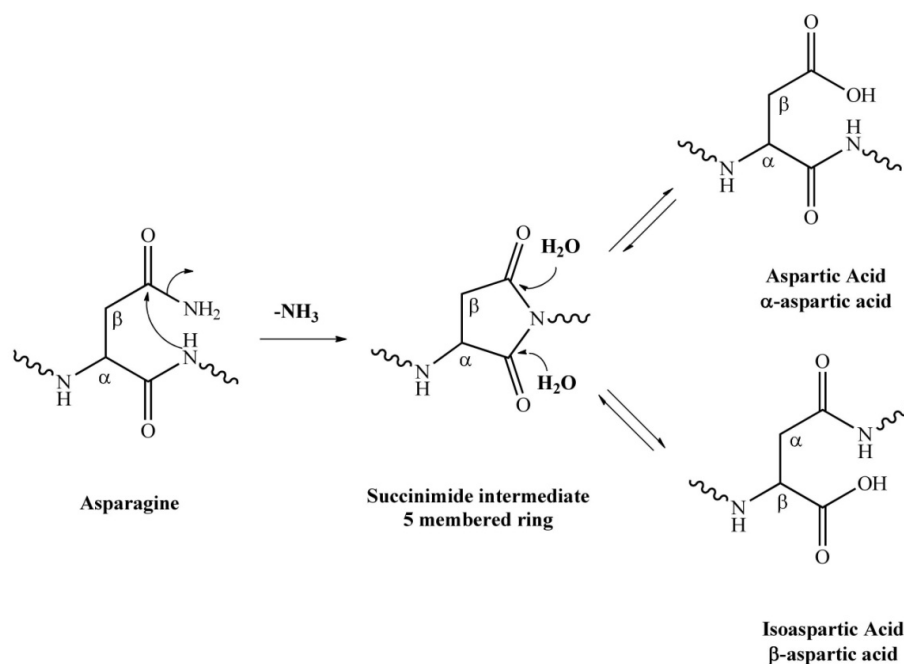
¹This chapter has been partially reproduced from
Perez, P., O'Connor, P. B., Differentiation of Isomeric Amino Acid Residues in Proteins
and Peptides Using Mass Spectrometry. *Mass Spectrometry Reviews*. **2012**, *31*. 609-625.

1. 1. Archaeological Importance of Deamidation

Deamidation of asparagine (Asn) and glutamine (Gln) in proteins and peptides has been the focus of extensive biomedical research for the last 20 years.¹ Research carried out by groups such as Robinson's and Smith's has significantly contributed towards understanding the role of deamidation in aging proteins.¹⁻⁹ For example the accumulation of Asn and Gln deamidation was showed to be the major modification in proteins such as crystallins.^{6-8, 10-12} This modification has been reported to increase with age.² As a result, studies involving Asn and Gln deamidation are closely related to the process of aging. However, the study of deamidation of proteins originating from ancient materials has recently increased interest in the field. Recent reports have shown the importance of deamidation for assessing the age of ancient materials.¹³⁻¹⁶ Particularly, in studies by Collins and co-workers that have pointed out that deamidation is remarkably increased in collagen from archaeological mammal bones. These studies emphasise the use of Gln and Asn deamidation as a potential technique for assessing the age of the materials.¹⁷ Collagen has been extensively used for radio carbon dating¹⁸ but there is no evidence of deamidation of collagen being used for age assessment in archaeological bone. The observation of Collins and co-workers, from the archaeological point of view, introduces a new possibility to determine the relative age of ancient materials. Thus, the importance of developing a method to monitor Asn deamidation in collagen is highlighted in this thesis.

To monitor deamidation of Asn in collagen is a challenge. This is mainly because deamidation of Asn could be artificially introduced into the system by sample handling. A different case is for Gln deamidation which has been predicted to be 10-

fold slower than Asn deamidation.¹⁹ For this reason the study presented in this thesis only focuses on Asn deamidation and not Gln deamidation. The present work will provide a solution to the issue of introducing artificial deamidation in the sample and will be discussed in depth in the following chapters. The background for the work presented in this thesis will be described in the following sections.

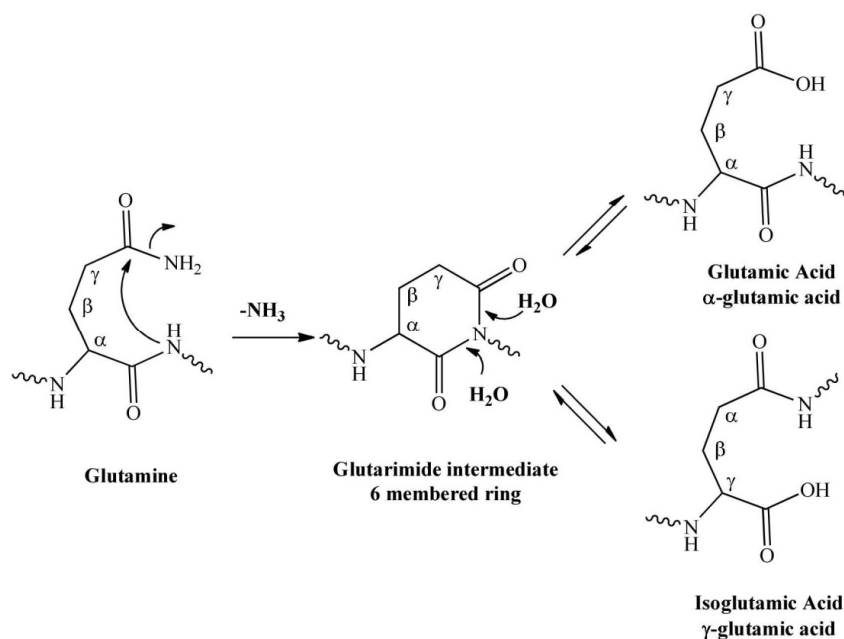


SCHEME 1.1: Isomerisation of aspartic acid (α -aspartic acid) and deamidation of Asn to isoaspartic acid (β -aspartic acid) and aspartic acid via a succinimide intermediate. Adapted from Cournoyer JJ, Pittman JL, Ivleva VB, Fallows E, Waskell L, Costello CE, O'Connor PB. 2005. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Science* 14:452-463.

1.2. Introduction to Peptide and Protein Deamidation

Deamidation of Asn and isomerisation of aspartic acid (Asp) residues in proteins are common, spontaneous, non-enzymatic post-translational modifications (PTMs), and result in a mixture of isomers of Asp and isoaspartic acid (IsoAsp) or, in the case of Gln deamidation glutamic acid (α -Glu) and isoglutamic acid (γ -Glu), as shown in Schemes 1.1 and 1.2 respectively.¹ Asn deamidation is faster due to the formation of a stable 5-membered ring intermediate, by contrast the reaction in Gln goes through a

less stable 6-membered ring intermediate as shown in Scheme 1.2. The modification is also associated with many diseases, particularly those which involve problems with protein folding. Due to the difficulty of distinguishing Asp and IsoAsp, insufficient work has been done to understand their formation or impact on diseases at a molecular level. The first reaction (Asn deamidation) results in a +0.984 Da mass shift making it relatively straightforward to detect. The second one, Asp isomerisation results in a mixture of Asp and IsoAsp (see Scheme 1.1) that can be differentiated by mass spectrometry using several strategies. Since the differentiation of both isomers using Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is part of the scope of the thesis; the strategies used for their analysis will be discussed further in the chapter.



SCHEME 1.2: Gln/Glu can undergo a similar reaction to Asn/Asp but the intermediate is a 6-membered succinimide which is hydrolyzed to a mixture of glutamic and γ -Glu.

1. 2.1. Mechanism of Deamidation in Peptides and Proteins

The most accepted mechanism describes that the reaction occurs spontaneously under physiological conditions through a succinimide intermediate whose rate is affected by both: its amino acid sequence and three-dimensional structure.²⁰⁻²² A nucleophilic attack of the peptide nitrogen occurs on the γ -carbonyl of the Asn losing ammonia to form the succinimide intermediate, as previously shown in Scheme 1.1. The reaction is highly affected by the neighbouring amino acids around the Asn residue. Consequently, the rate of the reaction will be dependent upon the bulkiness of the side chain.²³ Hydrolysis of the succinimide intermediate results in a mixture of products including L-IsoAsp, D-IsoAsp, L-Asp, and D-Asp. From this mixture, IsoAsp and Asp can be differentiated using mass spectrometry (MS), but MS can rarely differentiate stereoisomers.^{24, 25}

The most direct mechanism of deamidation involves the nucleophilic attack of the carbonyl group, which is a unimolecular reaction or a first order reaction. Consequently, the extent of deamidation in the system can be calculated by measuring the intensity ratios between the non-deamidated and the deamidated species at different points of the reaction. Non-deamidated and deamidated species behave similarly under ESI conditions,²⁶ thus allowing a relationship to be established between the peak intensities of each species. Deamidation at time zero can be determined by using the first order equation and extrapolating back to time zero.

$$A = A_0 e^{-kt} \quad (1.1)$$

where A = concentration of species A ($N_{(t)}$, intensity of the non-deamidated monoisotopic peak), A_0 = initial concentration of species A ($N_{(0)}+D_{(0)}$, where D is the

intensity of the monoisotopic peak of the deamidated species), k = rate constant, t = time. The exponential coefficient in equation 1.1 represents the rate constant for the deamidation reaction.

1. 2. 2. Mass Spectrometric Methods for Studying Deamidation

The most common methods used for studying deamidation using mass spectrometry are isotopic deconvolution^{15, 19, 23, 27} and mass defect.^{28, 29} The first method uses the +0.984 Da shift in the mass spectrum related to the change from NH₂ to –OH, which is highlighted by the atypical combination of isotopic patterns of the two species, as shown in Figure 1.1.A.

By assuming that the intensities of the non-deamidated species (N) and the deamidated species (D) are additive, the pattern can be deconvolved to its two separate forms, as seen in Figure 1.B and 1.C. To achieve deconvolution, the theoretical isotopic pattern of the two separate forms N and D needs to be fitted with the overall experimental pattern. This method is a quick way of monitoring deamidation and can be used to interpret data from instruments with low resolution such as ion traps mass spectrometers. Previous data processing using this method has been applied by Robinson and co-workers. They carried out an extensive analysis of the rates of Asn and Gln deamidation in over 700 hundred peptides using the data from a quadrupole mass spectrometer.^{19, 23, 27} Isotopic deconvolution has also been carried out using data from MALDI-TOF^{15, 30-32} and FT-ICR-MS.³³⁻³⁵

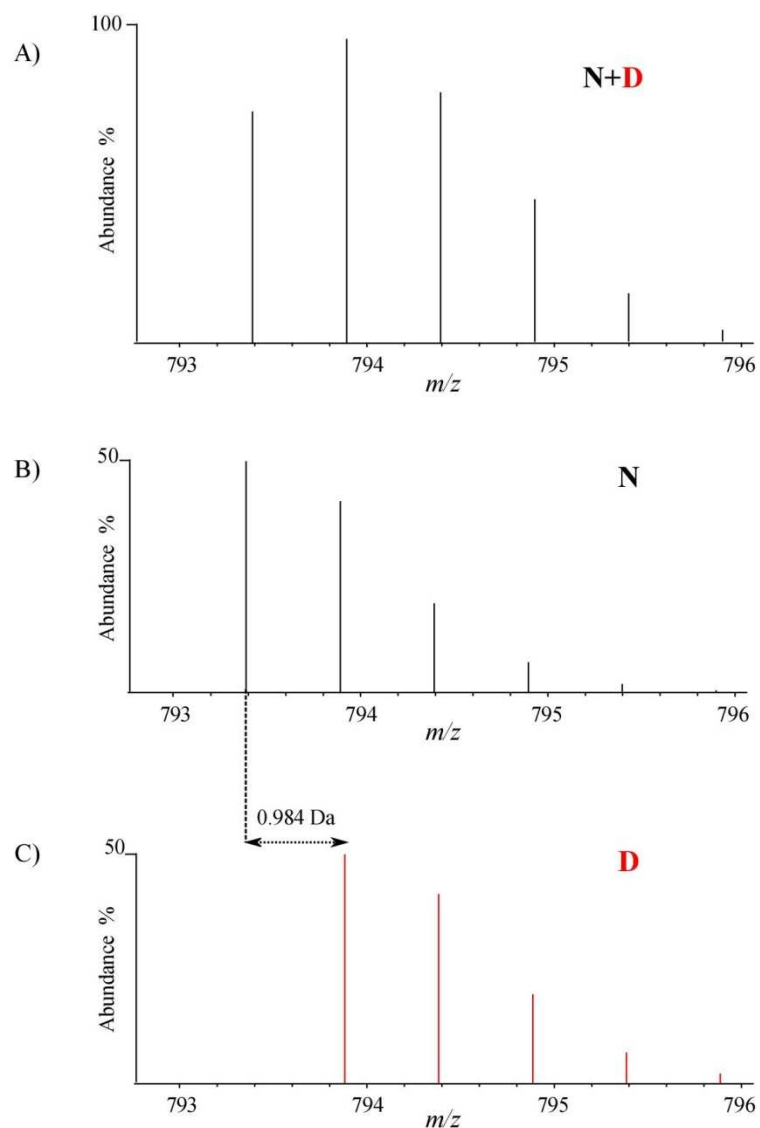


FIGURE 1.1: A) Isotopic distribution of the combination of non-deamidated (N) and deamidated (D) species. B) Deconvoluted isotopic distribution of the N. C) Deconvoluted isotopic distribution of D, showing the shift of 0.984 Da.

The mass-defect method uses the high resolving power of FT-ICR-MS to resolve the 19 mDa mass difference between the $A+1$ or ^{13}C isotope of a non-deamidated peptide and the monoisotopic peak (A) or ^{12}C of the deamidated form, as seen in Fig 1.2, which can be obtained in broadband mode (full spectrum) at around 500,000 resolving power.^{28, 36} Once the peaks are resolved, the relative intensity of the two peaks will show the extent of deamidation assuming the same ionisation and detection conditions for both forms.²⁶ Due to the speed in the analysis using the

mass defect method and the access to high resolution mass spectrometry, in this case FT-ICR-MS, the mass defect was the method of choice to carry out the experiments presented in this thesis. For this reason the following sections will describe the instrumentation and techniques used for the work presented here.

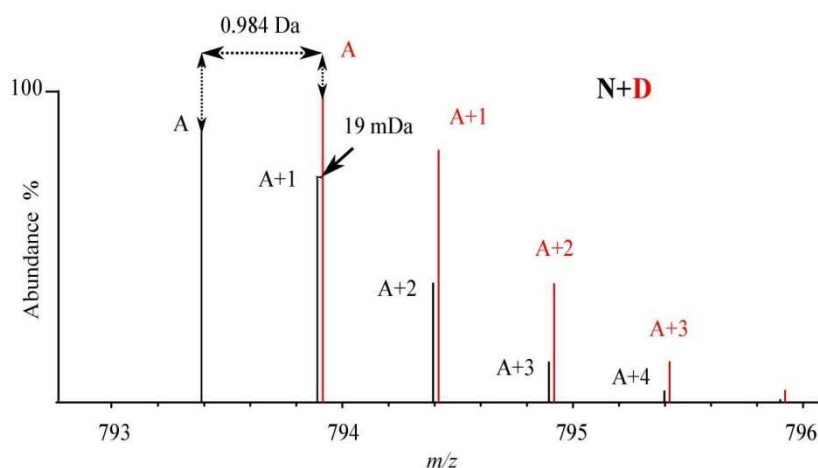


FIGURE 1.2: Simulation of the high resolution isotopic patterns corresponding to non-deamidated (black) and deamidated species (red).

1.3. Mass Spectrometry

Mass spectrometry (MS) is a powerful tool used to determine the mass-to-charge ratio of ions from a wide range of sample types. The technique has been implemented in numerous disciplines such as pharma industry, biology, medicine, archaeology, etc. The highlighted advantages of mass spectrometry rest on its universal applicability; nearly all molecules show a mass spectrum. High selectivity and sensitivity allow selection of single components amongst complex mixtures, also allowing detection levels as low as 1 femtomole. The principal layout of the instrument consists of four distinct regions that can be combined differently to obtain the desired instrument for specific analysis. Example of the distribution of these regions is shown in Fig. 1.3 in the FT-ICR-MS arrangement. In this case the ionisation source is electrospray ionisation (ESI); fragmentation methods such as

collisionally activated dissociation (CAD), electron transfer dissociation (ETD), and electron capture dissociation (ECD) can be used; and analyser-ion detection occurs in the ion cyclotron resonance cell (ICR cell).

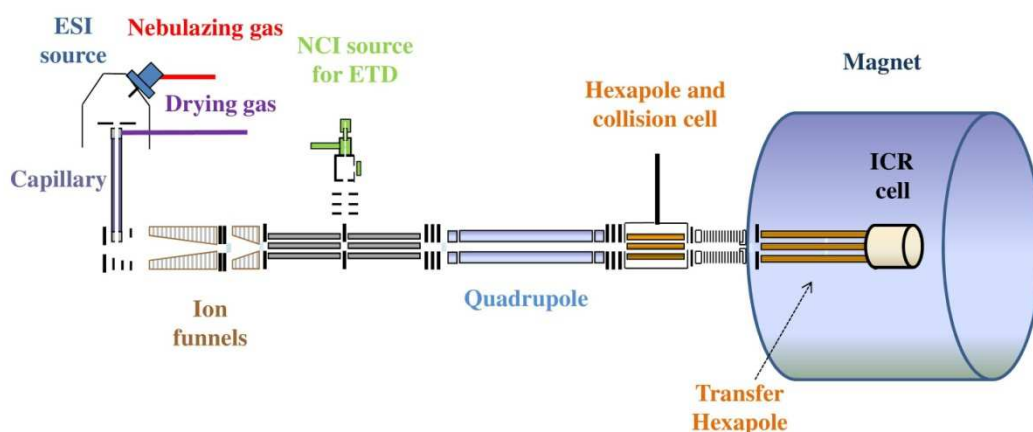


FIGURE 1.3: Schematic representation of the 12 T Bruker Solarix system. Adapted with permission from Bruker Daltonik GmbH.

1.3.1. Ionisation sources

Mass spectrometers separate and detect ionised molecules using electric and in some cases magnetic fields. For the mass spectrometer to measure mass-to-charge ratio, the sample must be ionised. This is needed because most molecules are electrically neutral. The ionisation step can be achieved by adding or removing one or more charged particles, which will generate positive or negative charges. Different methods are used to achieve ionisation of molecules; they are classified into two groups: “classical” and “soft” ionisation methods. The “classical” methods refers primarily to electron ionisation (EI), which is the primary ionisation/fragmentation method used for structural analysis of low molecular weight compounds (<500 Da). In EI, the analyte precursors are ionised and fragmented by interaction with a beam of energetic electrons (70 eV).^{37, 38} These are emitted from an incandescent filament and travel through the ion chamber to an anode (electron trap) on the opposite side as

shown in Fig. 1.4. A stream of vaporized sample molecules interacts with the beam of electrons to produce different products (precursor ions, metastable intermediates, and fragments). These are pulsed out of the source and accelerated into the mass analyser device. EI produces radical rearrangements to generate significant fragmentation. EI requires the analyte to be vaporized before ionisation; as a consequence it is typically limited to stable molecules with molecular weights below 500 Da.

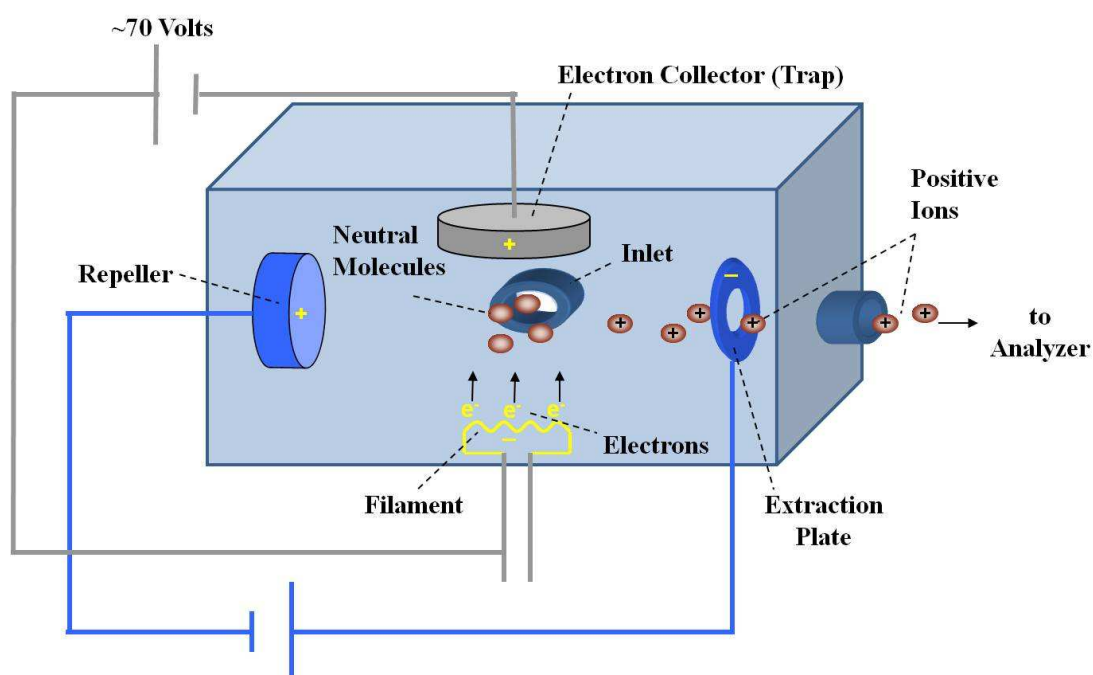


FIGURE 1.4: Electron ionisation source.

The main disadvantage in the analysis of an EI spectrum is emphasized when no information about the precursor ion is available. All the vaporized molecules in the sample will be ionised contributing to the mass spectrum, which take the spectrum to a high level of complexity, making the interpretation very challenging. Also the extensive fragmentation of the precursor ion often complicates the spectrum. The need for new ionisation techniques to address these issues and to extend ionisation to non-volatile molecules such as proteins, peptides, and other macromolecules drove

the creation of other ionisation methods. Amongst these are fast atom bombardment (FAB),³⁹ matrix-assisted laser desorption/ionisation (MALDI),⁴⁰ and electrospray ionisation (ESI).⁴¹

Fast Atom Bombardment (FAB)

This ionisation technique uses a high energy beam (~ 8 keV) of neutral atoms (Cs, Xe, or Ar)³⁹ to vaporize the analyte immersed in a matrix such as glycerol, which is commonly used for this purpose.

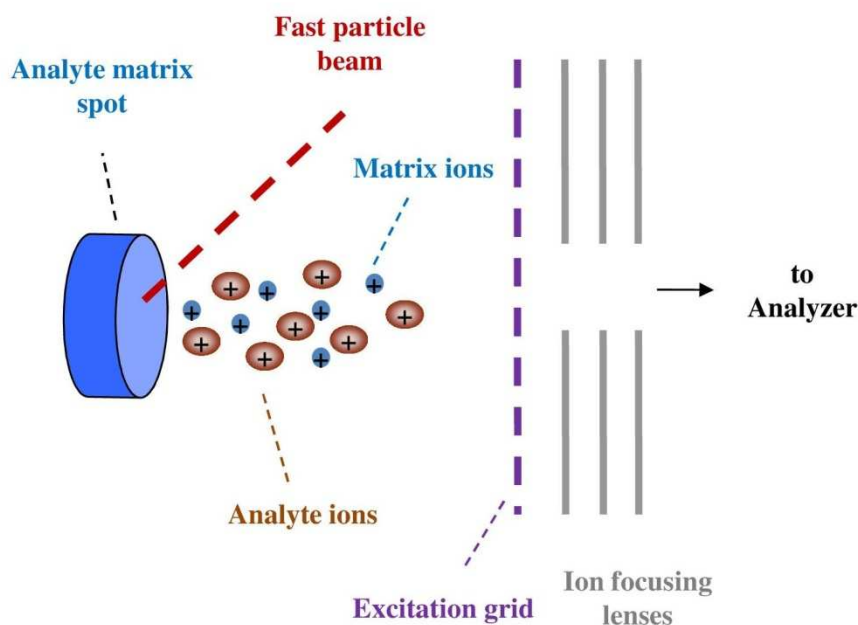


FIGURE 1.5: Process of ionisation in FAB.

In the bombardment process, high energy atoms or ions desorb positive ions, negative ions, and neutrals from the surface of the matrix as seen in Fig. 1.5. These charged species are then transferred to the mass analyser by applying an electric field. Later on, other ionisation techniques were implemented to compensate for the low signal-to-noise ratio and to facilitate sample delivery in the MS.

Matrix-assisted Laser Desorption/Ionisation (MALDI)

In laser desorption/ionisation (LDI) the analyte molecules absorb radiation from a laser at the working wavelength. If the analyte is not able to absorb radiation, a matrix is used to assist the ionisation process, which is the technique known as matrix-assisted laser desorption/ionisation (MALDI). Matrices that readily absorb the laser radiation are aromatic molecules such as 2,5-dihydroxy benzoic acid (DHB)⁴² and alpha cyano hydroxycinnamic acid.⁴³ (See Fig 1.6 for structures).

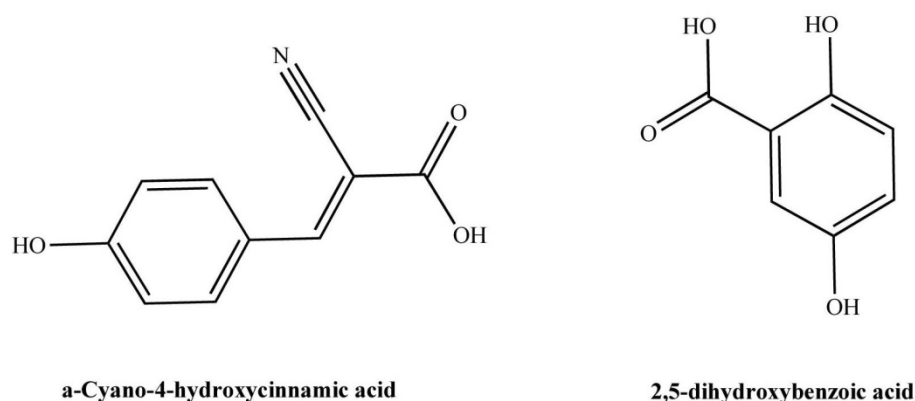


FIGURE 1.6: Chemical structure for alpha cyano hydroxycinnamic acid and 2,5-dihydroxy benzoic acid (DHB).

In most cases the analyte is mixed with the matrix and dried down on a surface. The exact mechanism of the MALDI ionisation process is not totally understood, but there are several theories.⁴⁴ The most accepted theory is that protons are transferred from the matrix molecules to the analyte molecules, either during or after the desorption process. In this process the analyte becomes charged, as seen in Fig 1.7. The technique was first implemented by Karas and co-workers in 1987 to ionise proteins from a crystalline nicotinic acid matrix mixture using a neodymium-yttrium aluminium garnet laser Nd:YAG (4th harmonic, 266 nm) as the radiation source.⁴⁰

The function of the matrix is to be an intermediate in the energy transfer between the laser and analyte. Later on the technique was used in 1989 by Beavis and co-workers where they observed the influence of the wavelength on laser desorption of ions from organic solids for various amino acids and dipeptides, some absorbing at 266 nm and all nonabsorbing at 355 nm.⁴⁵ The exact mechanism of this process still remains unknown but is hypothesized to be a process where laser absorption and excitation of the matrix boost vaporization of the sample. As a consequence, a mixture of ionised matrix and neutral analyte is produced followed by proton transfer that result in an ionised analyte.⁴⁴

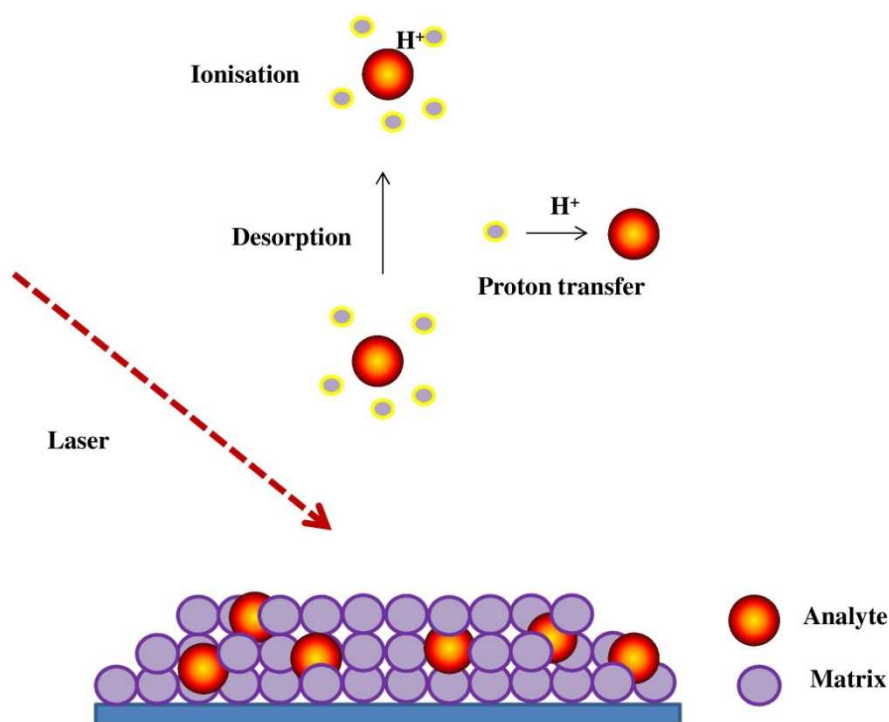


FIGURE 1.7: Representation of the MALDI ionisation process.

MALDI typically generates singly charged species as opposed to ESI which often generates multiply charged ions (depending on the number of available ionisation

sites). The method is used for detection and characterization of biomolecules, such as proteins, peptides, oligosaccharides and oligonucleotides, with molecular masses between 400 and 350,000 Da.^{46,47}

Electrospray Ionisation (ESI)

Electrospray ionisation has been around for a while, but it was not until 1989 when John Fenn pioneered its use for ionisation of high mass biomolecules that were analysed by mass spectrometry.⁴¹ His significant achievement was recognized by the awarding of the Nobel Prize in Chemistry in 2002. ESI serves as an interface to ease the transition of molecules from solution to the gas phase. Usually, the analyte is dissolved in a volatile solution such as 50:50 methanol: water, but the solvent will depend on the properties of the analyte and whether ionisation is in positive or negative mode. If ionisation in positive mode is desired, a small amount of acid is added to boost the ionisation of the analyte. The solution passes through an electrospray needle that has a high potential difference with respect to the entrance of the mass spectrometer (normally in the range of 2.5 to 4 kV). This can be done for either charge, positive or negative, by changing the polarity of the potentials applied to the needle. Charge accumulates at the meniscus of the solution, at the tip of the capillary as shown in Fig. 1.8. This charge is induced by the electrical potential. A cone is formed at the end of the capillary and highly-charged droplets are expelled from the capillary needle.⁴⁸ The process is assisted by a nebulising gas (usually nitrogen), which flows towards the inlet of the mass spectrometer (Fig. 1.8).

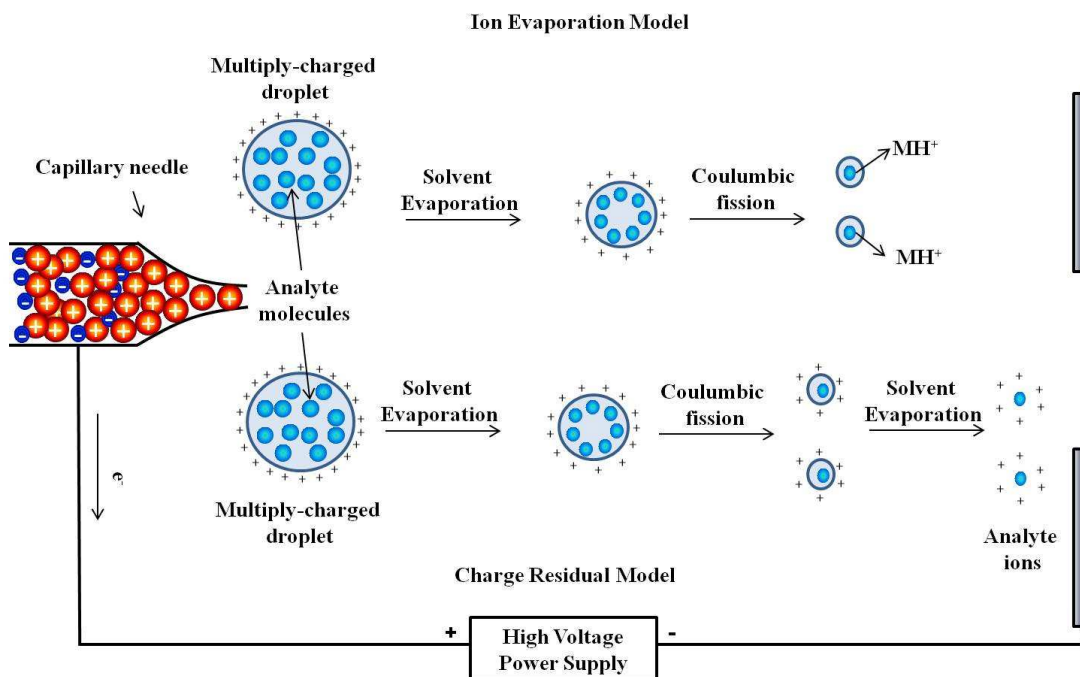


FIGURE 1.8: Electro spray ionisation process. Figure adapted from Kebarle, P. A Brief Overview of the Present Status of the Mechanisms Involved in Electro spray Mass Spectrometry. *Journal of Mass Spectrometry*. **2000**, 35, 804-817.

The process of ionisation is believed to occur through two mechanisms; the ion evaporation model⁴⁹ (IEM) and the charge residual model (CRM).⁵⁰ During these processes solvent evaporates from the droplets and their size decreases down to the point when coulombic repulsion exactly balances the surface tension and the droplet explodes, this process is best known as “Coulombic Explosion.” The point at which this happens is known as the Rayleigh Limit. The IEM suggests that when the droplet approaches certain radius desorption of the analyte ions is assisted by the field strength at the surface of the droplet.⁴⁹ The CRM model proposes that the highly-charged droplets undergo repeated cycles of solvent evaporation and Coulombic explosions, leading eventually to droplets that only contain one analyte ion as seen in Fig. 1.8.⁵⁰ In general is thought that ions with a small number of charges are preferably formed by the IEM model, and multiply-charged ions are

formed by the CRM model.⁴⁸ Multiply-charged ions are normally produced in ESI, which takes the spectrum to a higher level of complexity.

1.3.2. Mass Analysers

There is a variety of mass analysers available on the market. They vary according to, separation principles, mass range limit, resolving power, mass accuracy, dynamic range, sensitivity, and price. They often use electric and/or magnetic field to sort the ions by their mass-to-charge ratio. In a mass spectrometer, the mass analyser is located at low pressures, which prevent collisions with background gases, to avoid damping of ions during trajectory.

Quadrupole Mass Analyser

Quadrupole have been used since the 1950's and are still one of the most popular mass analysers.⁵¹ Quadrupole are commonly used because they can operate at a relatively poor vacuum, and also due to their low cost compared to other mass analysers.

The quadrupole mass analyser consists of four identical rods in a parallel arrangement. Each rod is connected electrically opposite to each other and they allow ions with specific m/z values to pass through the middle of the arrangement until the ions reach the detector. This is achieved by applying a direct current (DC) radio frequency (RF) to the rods.^{37, 38} Other ions will clash toward the rods and get neutralized as seen in Fig. 1.9 for the red ions. This allows filtering ions with a specific m/z by scanning through different voltages at the particular m/z range.

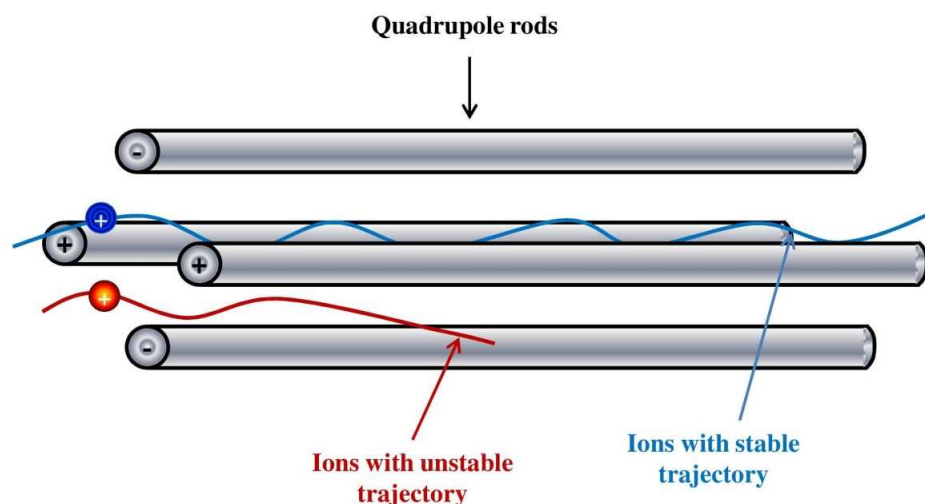


FIGURE 1.9: Representation of a quadrupole mass analyser.

The mass accuracy and resolution of quadrupoles are lower compared to other mass analysers such as time of flight and FT-ICR. Quadrupoles can be used for tandem mass spectrometry analyses, where three quadrupoles are placed in sequence. This arrangement of quadrupoles is known as triple quadrupole mass spectrometry.

Time of flight (TOF) Mass Analyser

TOF mass analyser is the simplest of all mass analysers in terms of theory. Ions are given a particular kinetic energy and they are allowed to travel through a field-free region to a detector (normally the length of the region is from 0.5 to several meters). The time it takes for the ions to reach the detector is measured and converted to the mass-to-charge ratio with the following equation:

$$v = \left(\frac{2zeV}{m} \right)^{1/2} \quad (1.2)$$

Where v is the velocity of the ion, z is the charge, e is the magnitude of the electronic charge, V is the accelerating potential and m is the mass of the ion. Usually the ion packet is accelerated to a defined kinetic energy and then separated in the field-free

region according to their mass-to-charge ratio, as seen in Fig. 1.10. During this travel the pressures are kept to around 10^{-7} mbar to avoid collisions with background gas. The time the ion takes to reach the detector is associated with equation 1.2 and the mass is calculated. The main advantages of TOF instruments are reflected in the wide range of m/z that can be measured with good sensitivity, high speed, moderate resolving powers (5,000 to 20,000), and cost (£60 K to £400 K). The resolution of TOF instruments is limited by the length of the tube, propagation delay in detector, and the kinetic energy distribution of the ions. Sensitivity is also limited due to ion stability and ion transfer efficiency.

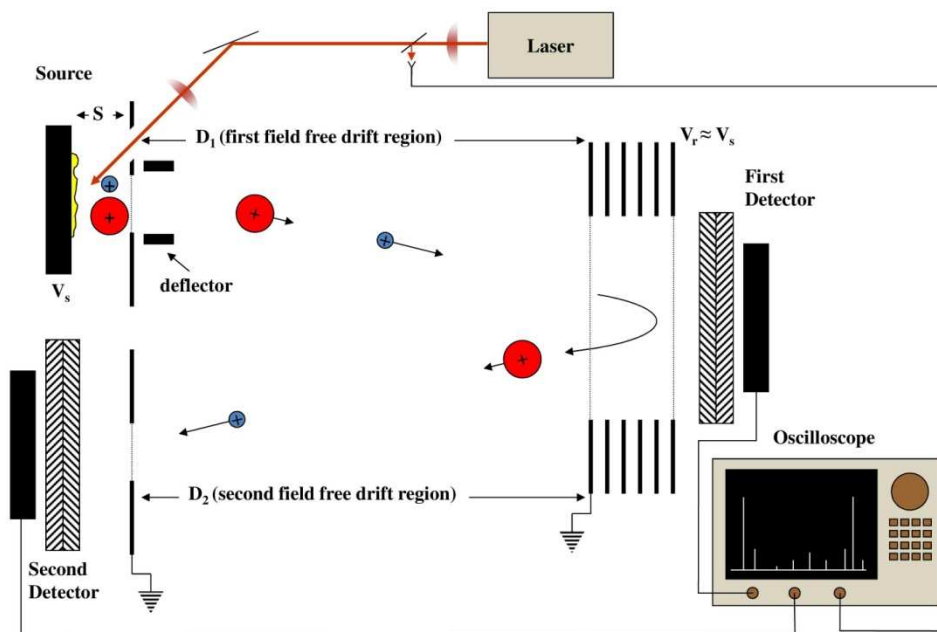


FIGURE 1.10: Combined Linear/Reflectron MALDI-TOF mass spectrometer. Reprinted with permission from O'Connor.

1.3.3. Fourier Transform–Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS)

First introduced by Marshall and coworkers, FT-ICR mass spectrometry^{52, 53} is used to determine the mass of ions that are trapped in a “cell,” which is typically cubical

or cylindrical in geometry and it is located in a magnetic field. Ions are generated externally in a separate ion source and then injected in packets into the cell where they are trapped within an electric and magnetic field, see Fig. 1.3. Since ions are charged particles, they will rotate around the center of the magnetic field axis, resulting in an orbit. This motion is known as the “cyclotron motion.” Ions are excited to their particular cyclotron orbit by applying an RF potential to the excitation plates that are opposite to each other in the cell (see Fig. 1.11).

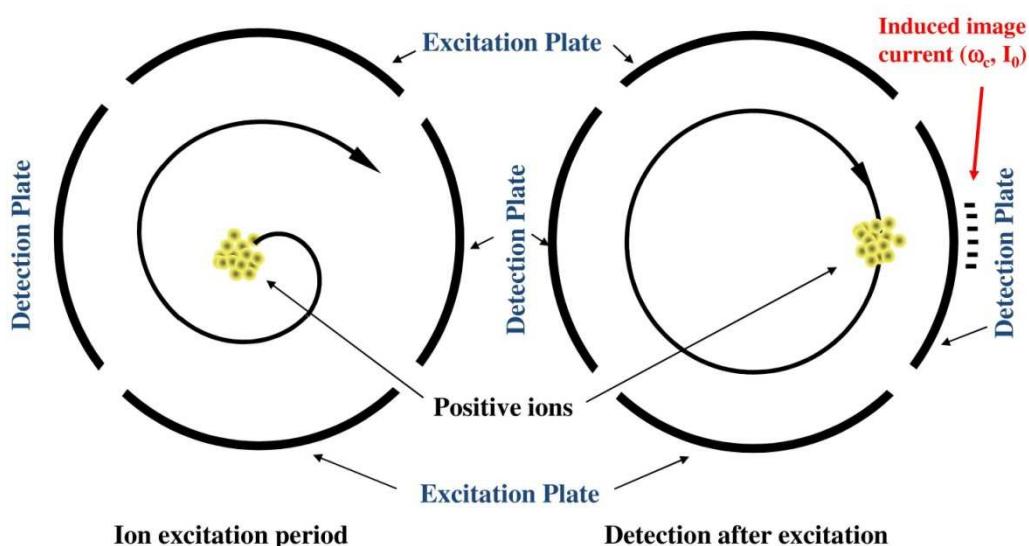


FIGURE 1.11: Representation of the process of excitation and detection of the ions in the ion cyclotron resonance (ICR) cell. Reproduced with permission of Bruker Daltonik GmbH.

The detection of the ions is performed when they pass the two detector plates located opposite to each other. The potential (voltage) change between the detection plates can be measured as a function of time and it is from here that the raw data (known as a transient, time-domain data, or free induction decay (FID)) is obtained. This raw data represents the detection of all the ions at the same time, with their different cyclotron frequencies. As a consequence it is necessary to extract the data from the different ion packets. This is achieved by using a Fourier transformation (FT)⁵⁴

where the time-domain signal spectrum is transformed into a frequency-domain spectrum.

The major advantage of this technique relies on the fact that once excited, the ions can rotate thousands of times (kHz range) for several seconds, which is measured constantly. The exceptional characteristic of numerous rotations allows high precision measurements that generate high-quality mass spectra with a resolving power that often reaches 10^6 , and it is used for experiments that require high resolution and mass accuracy, which is the case of the experiments presented in this thesis.

Ion Cyclotron Resonance Cell (ICR cell)

One of the key components of the FT-ICR MS is the trapped-ion cell (ICR cell, ICR trap, or Penning trap), which is located in the homogenous region of a strong magnetic field (see Fig. 1.3), and where all the measurements of the mass-to-charge ratio (m/z) of an ion take place. Each measurement cycle requires ion injection, excitation, and acquisition of the induced ion signal (transient detection).

Over many years, significant improvements have been made to the ICR cells. These improvements are based on achieving a more uniform electromagnetic field inside the cell. To address this issue a new type of cell called the “infinite cell” was introduced in 1991. The approach of this cell was to eliminate the z-ejection of ions, which is caused by the z component of the excitation electric field, without affecting the trapping and sensitivity properties of the cell.⁵⁵ This is the cell currently used by

Bruker Daltonics Ltd. in the 12 T Bruker solariX (Fig. 1.12), which is the instrument used for the work reported in this thesis.

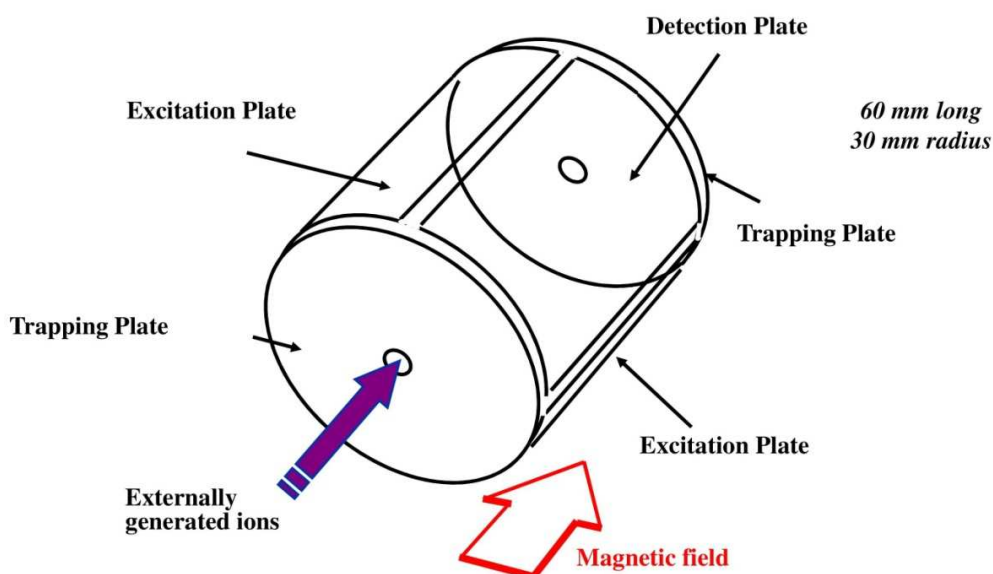


FIGURE 1.12: Diagram of the “Infinite cell”. Red arrow shows the direction of the magnetic field. Figure was reproduced with the permission of Bruker Daltonik GmbH.

Ion Motion in the ICR Cell

In general, charged particles moving in the presence of a magnetic field will experience a force known as the Lorentz Force. The axes of the force, motion of the positively charged particle, and magnetic field are mutually perpendicular, as directed by right hand rule, see Fig. 1.13.

The Lorentz Force acting on a charged particle in the presence of a magnetic field alone is given by the following equation 1.3:

$$\vec{F} = q\vec{v} \otimes \vec{B} \quad (1.3)$$

Where q is the charge of the particle in coulombs (C), v is the velocity of the charged particle (ms^{-1}), B is the magnetic field strength (tesla, T). The vector cross product in equation 1.3 means that the direction of the magnetic component of the Lorentz force is perpendicular to the plane determined by v and B as shown in Fig. 1.13. If the

trajectory of the ion, moving in the direction of the magnetic field, is stable at constant speed, the centripetal force (the inward force) and the centrifugal force (the outward force) need to be the same, as shown in equation 1.4:

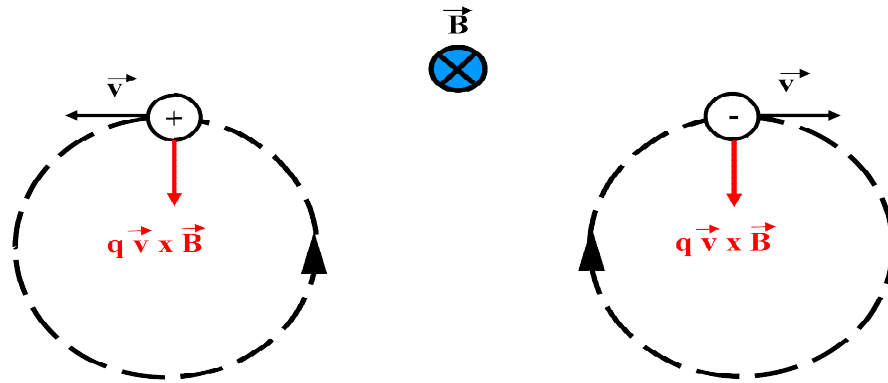


FIGURE 1.13: Direction of the Lorentz force due to the presence of a magnetic field perpendicular to the plane of the paper.

Equation 1.4 can be expressed in the magnitude form as:

$$\frac{mv^2}{r} = qvB \sin \theta \quad (1.5)$$

In this equation θ is the angle between the axes of the ion motion and the magnetic field.

To simplify equation 1.5 it can be assumed that the ion is moving only perpendicularly to the axis of the magnetic field, then equation 1.5 would be:

$$\frac{mv^2}{r} = qvB \quad (1.6)$$

If the angular velocity is defined by:

$$\omega = \frac{v}{r} \quad (1.7)$$

Rearrangement of equation 1.6 yields:

$$\frac{v}{r} = \frac{qB}{m} \quad (1.8) \quad \text{or} \quad \omega = \frac{qB}{m} \quad (1.9)$$

Equation 1.9 is called the cyclotron equation, where m is the mass of the ion (kg), and ω (rad s^{-1}), is the cyclotron frequency. According to equation 1.9, all ions of a given mass-to-charge ratio have the same ICR frequency, which is independent of their velocity. From this last equation it is clear that ions with lower m/z have higher cyclotron frequencies than ions with higher m/z .

Ions are trapped in the cell where they oscillate along the magnetic field lines, which is possible due to the potential applied to the trapping plates shown in Fig. 1.12. The potential is usually $\sim 1\text{V}$ of the same polarity as the ion charge. The ions will then oscillate along the z -axis of the ICR cell. As consequence of this trapping potential an electric field is induced in the xy -plane causing the ions to be pushed towards the detection and excitation plates. The presence of this electric field introduces a third motion of the ions called magnetron motion, which causes the centre of the ion cyclotron orbit to oscillate around the ICR cell axis (see Fig. 1.14).

So far, the derivation up to equation 1.9 has assumed the presence of no electric field. However, as explained above there is a radial electric field induced by the trapping potentials in the ICR cell, creating a force which opposes (and, thus, has opposite sign with respect to the Lorentz force). The force is called radial force, and is defined by the following equation:

$$\vec{F}_r = q\vec{E}_r = \frac{(qV_{trap}\omega r)}{a^2} \quad (1.10)$$

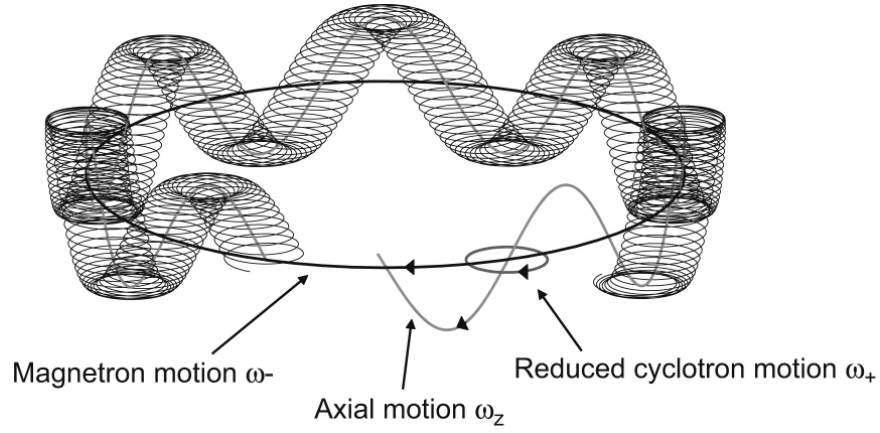


FIGURE 1.14: Ion motion in the ICR cell. The axis of the magnetic field is parallel to the plane of the paper. Reprinted from Kolhinen, V.S., et al. Commissioning of the double Penning trap system MLLTRAP. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2009, 600, 391-397, with permission from Elsevier.

Where α is a constant that depends on the geometry of the cell and a is the distance between the trapping plates. Expressing equation 1.6 in terms of ω and combining with equation 1.10:

$$m\omega^2 r = qB\omega r - \frac{(qV_{trap}\alpha r)}{a^2} \quad (1.11)$$

Rearranging equation 1.11:

$$\omega^2 - \frac{qB\omega}{m} + \frac{qV_{trap}\alpha}{ma^2} = 0 \quad (1.12)$$

Equation 1.12 is a quadratic equation, which provides two solutions.

$$\omega_+ = \frac{\omega_c}{2} + \left(\left(\frac{\omega_c}{2} \right)^2 - \frac{\omega_z^2}{2} \right)^{1/2} \quad (1.13)$$

$$\omega_- = \frac{\omega_c}{2} - \left(\left(\frac{\omega_c}{2} \right)^2 - \frac{\omega_z^2}{2} \right)^{1/2} \quad (1.14)$$

Where ω_+ and ω_- are the reduced cyclotron and magnetron frequencies, in which ω_c and ω_z are the unperturbed cyclotron frequency and the axial oscillation frequency respectively. In equation 1.14 ω_z is defined by:

$$\omega_z = \left(\frac{2qV_{trap}\alpha}{m\alpha^2} \right)^{1/2} \quad (1.15)$$

As seen in equation 1.15 the trapping oscillation frequency is directly proportional to the trapping potentials, which is the reason why they are kept low ~ 1 V. The magnetron and the trapping frequencies are usually a lot less than the cyclotron frequency, and they are not commonly detected.

Ion Excitation and Detection in the ICR Cell

Once the ions are in the cell, their rotation radius becomes highly important because it must be large enough to be measured by the detection plates but not so large that the ions strike the plates and are lost. Furthermore, field inhomogeneities worsen as the ion radius increases, thus the optimal radius is usually around 50 % of the cell radius. In order to detect the ions, excitation is necessary. This is achieved by applying a radio frequency (RF) potential to one set of opposed excitation plates. The excited ions enter an orbit that is resonant with the applied oscillating excitation. As the ion packet passes close to the detection plates, it induces accumulation of charges on the plate with opposite sign to the ion packet. As a result an alternating image current will be generated on the two detection plates. The image current is converted into an alternating voltage, which is amplified to generate a time domain signal (a transient). The time domain signal is then Fourier transformed into a frequency-domain spectrum, which is calibrated in terms of m/z ratio using the cyclotron equation (see Fig.1.15).

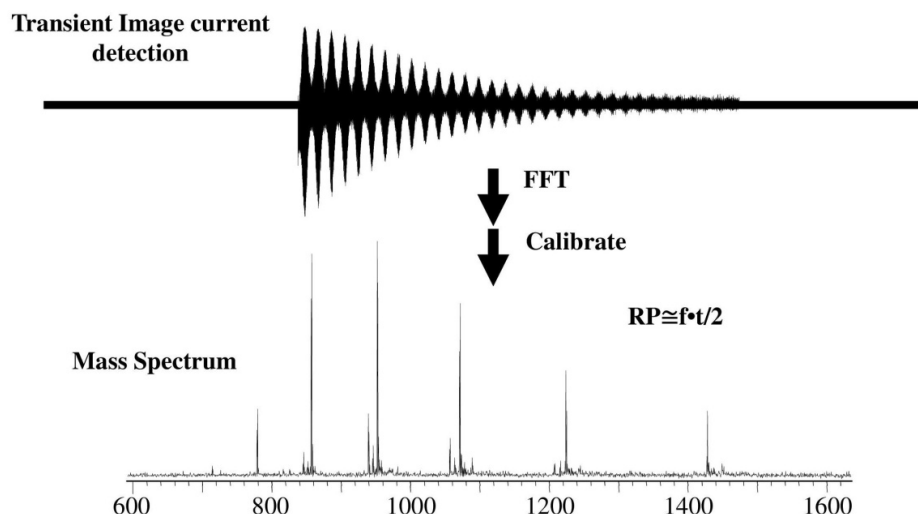


FIGURE 1.15: Fourier transformation of the time domain signal into frequency domain signal. Adapted from P. B. O'Connor, 2010

After the ions have been excited into a larger orbit, they will lose their energy and go back down to the centre of the cell, which causes the transient to decay throughout time. The decay in the transient will directly affect the FT-ICR resolving power,⁵⁶ which is often represented by the “rule of thumb” in equation 1.16:

$$R \sim \frac{fT}{2} \quad (1.16)$$

where R is resolution, f is the cyclotron frequency, and T is the transient length.

1.3.4. Fragmentation Methods Used for Structural Analysis of Proteins and Peptides in Mass Spectrometry

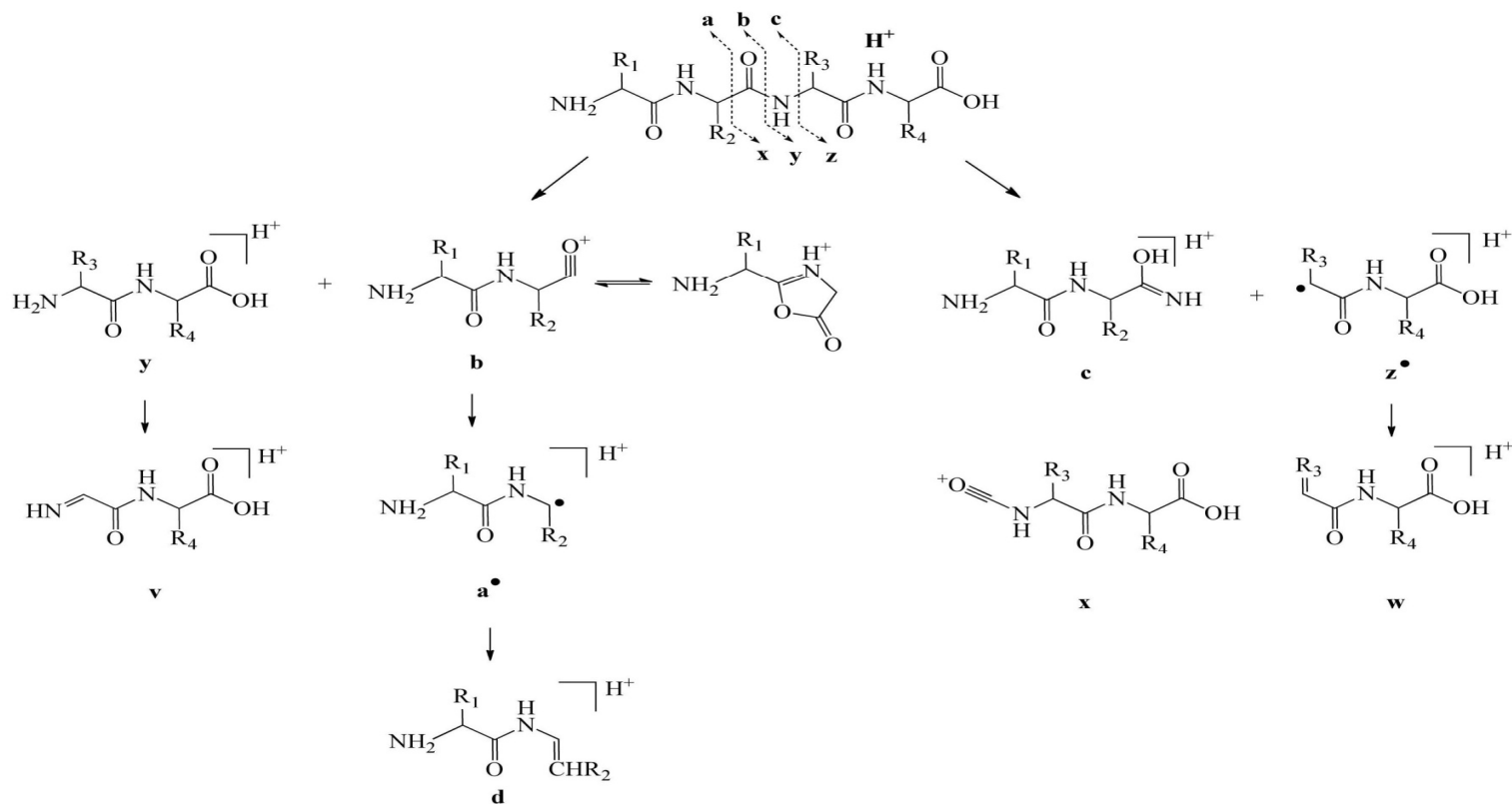
Collisionally Activated Dissociation (CAD)

CAD is the fragmentation of ions as a result of collisions with neutral gases. In this process, a percentage of the kinetic energy of an ion is converted to internal energy by collision with gas molecules. CAD was discovered by Keith Jennings, although he named the method collisionally induced dissociation (CID), a nomenclature

confusion which persists to this day.^{57, 58} He realized that by increasing the gas pressure in the flight tube of the spectrometer, precursor ions could be fragmented, and that these product fragments carried chemical structure information with them. The technique was rapidly implemented elsewhere by McLafferty.⁵⁹

Low-energy CAD is considered a “slow-heating” fragmentation method because ions are activated and deactivated via low energy collisions. Collision with gas molecules results in vibrational excitation allowing the energy transmitted from the collision to be randomly distributed throughout the molecule, in a process known as intramolecular vibrational redistribution; as a consequence, fragment ions are primarily formed from the lowest energy bonds or rearrangements available in the molecule. For peptides and proteins, the amide bond usually dissociates first, via a proton transfer rearrangement,⁶⁰ thus generating b and y ions^{61, 62} (see scheme 1.3)

The formation of b and y ions in the gas phase is described by the “Mobile Proton” model.⁶³⁻⁷² This model states that fragmentation is driven by the charge-bearing proton and its location. The fragmentation of protonated peptides would require the presence of a proton at the cleavage site. When the amino acid is attached to the proton, the energy is used to move the proton from the side chain to the backbone where dissociation happens. Thus, the general fragmentation mechanism of peptides following low energy (<100 eV collisions) CAD is believed to occur via nucleophilic attack of the carbonyl oxygen of one amide group to a nearby carbonyl carbon of another amide group, which has been rendered electrophilic by protonation.⁷³



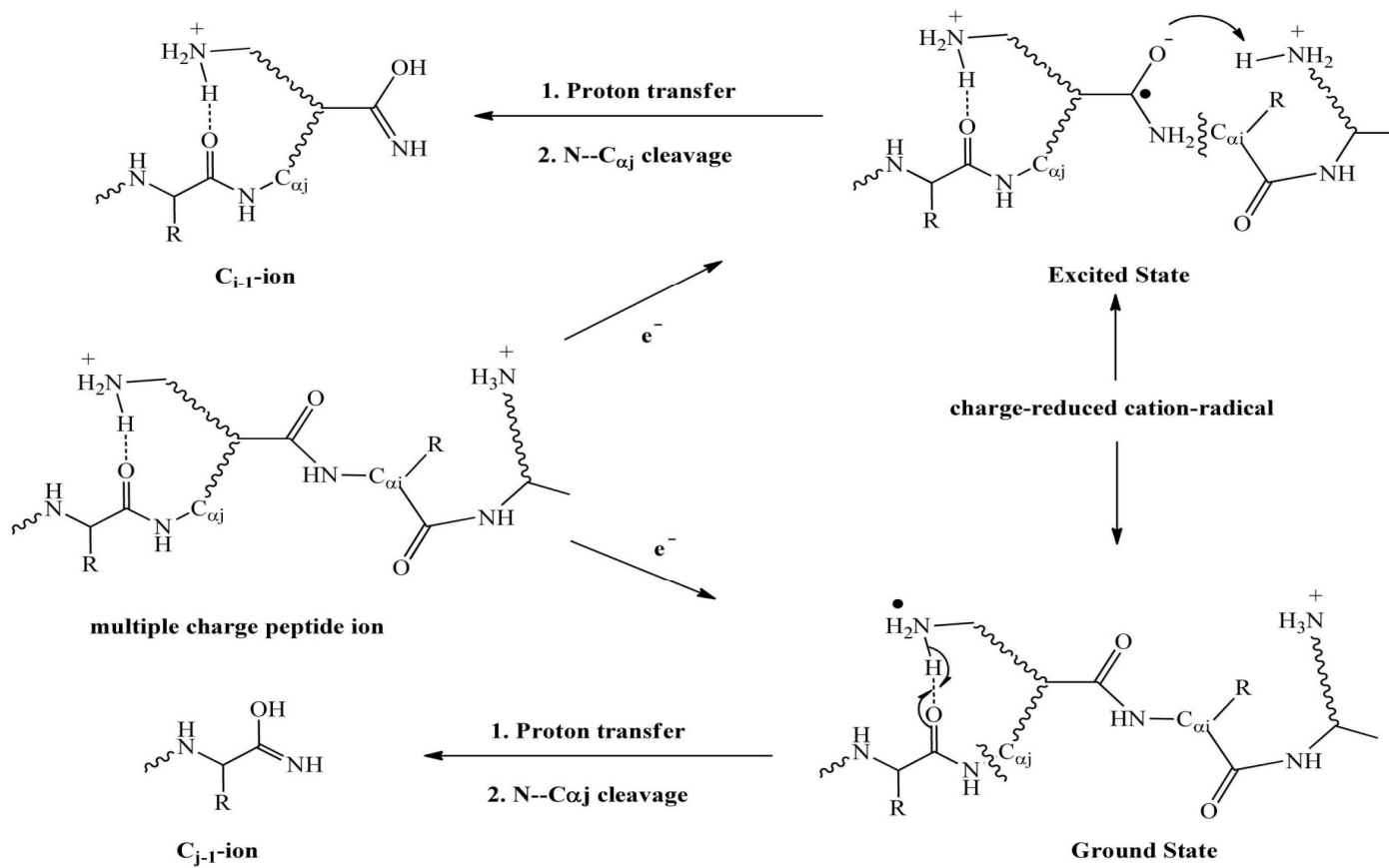
SCHEME 1.3: Product ion fragmentation of peptides. Adapted from: Roepstorff, P. and Fohlman, J. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomedical Mass Spectrometry*. **1984**, *11*, 601-50. Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., and Watson, J.T. CID of Peptides. *Analytical Chemistry*. **1987**, *59*, 2621-2625. And Wysocki, V.H., Tsapralis, G., Smith, L.L., and Brecci, L.A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *Journal of Mass Spectrometry*. **2000**, *35*, 1399-1406.

In addition to backbone cleavages, other labile bonds in peptides and proteins also dissociate easily; this includes those involving PTMs, such as glycosylation of the side chain amide nitrogen of Asn (N-glycosylation) and the side chain oxygens of serine (Ser) and threonine (Thr) (O-glycosylation), as well as phosphorylation on the side chain groups of Ser, Thr, and tyrosine (Tyr). In some cases, loss of such labile groups represents a considerable limitation of the method since the localisation of PTMs at a certain residues is, in some cases, required for study of the function of a modified protein.

Electron Capture Dissociation (ECD)

Electron capture dissociation is a fragmentation technique used in tandem mass spectrometry experiments where low energy electrons (<0.5 eV) are reacted with multiply charged positive ions. These ions are charge reduced (for example, an $[M+nH]^{n+}$ becomes $[M+nH]^{(n-1)+\bullet}$) forming a radical cation.

The mechanism by which this fragmentation takes place is the topic of extensive discussion in the field. Scheme 1.4 shows the currently most accepted mechanism of the fragmentation pathway corresponding to ECD. This mechanism has been proposed independently by the groups of Simons⁷⁴⁻⁷⁷ and Turecek⁷⁸ and is consequently referred to as the Utah-Washington mechanism or UW for short. The UW mechanism considers ECD to occur by a cascade process⁷⁹ where a number of electronic states of the incipient charge-reduced ion are sampled consecutively.



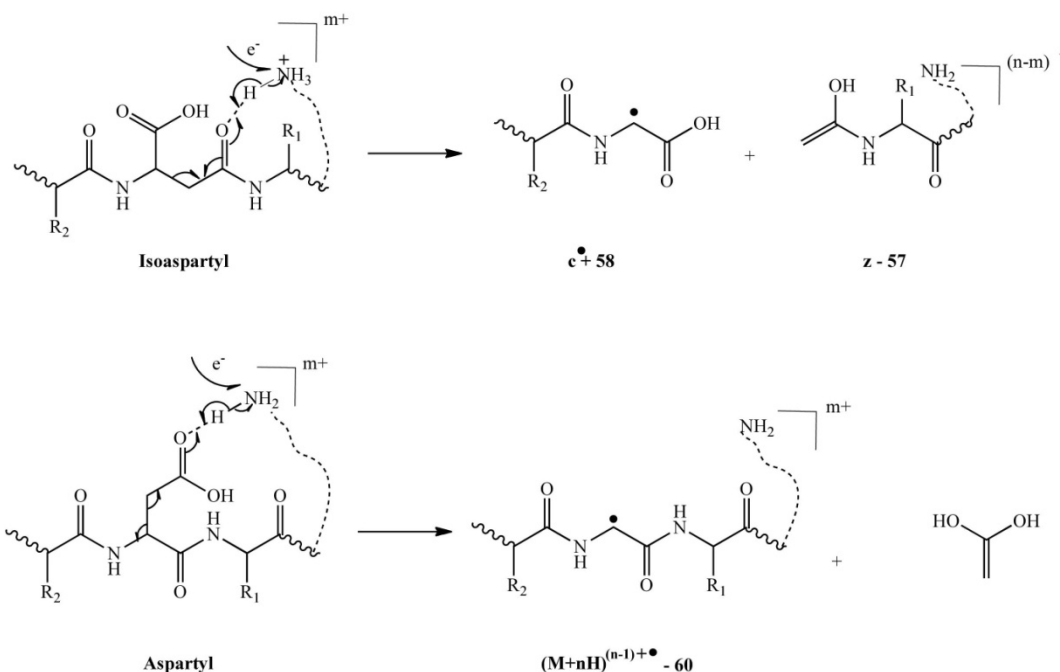
SCHEME 1.4: General mechanism of Electron Capture Dissociation of peptides. Ground and excited electronic states of peptide ions. Reprinted from Syrstad EA, Turecek F. 2005. Toward a general mechanism of electron capture dissociation. *Journal of the American Society for Mass Spectrometry* 16:208-224. Copyright © 2005, with permission from ELSEVIER.

If the ground state of an ammonium radical (lysine or N-terminus) is reached, a hydrogen atom can be transferred from the ammonium group onto a proximate amide carbonyl. Because of the competition between H-atom loss and transfer, the latter is favoured when occurring at an internally solvated amide carbonyl. According to the UW mechanism the capture of the low energy electron (~ 0.2 eV) is believed to occur at the π orbital of an amide group in the presence of a remote charge in the molecule, producing a charged reduced species. Subsequently, the N-C α bond cleaves producing c and z \bullet fragment ions.

Even though ECD generally provides a better range of masses for sequencing peptides than does CAD, it has the limitation that it does not cleave at proline (Pro) because the N-C α bond is inside the Pro ring which thus requires two cleavages to see a fragment (note, this Pro effect can be taken advantage of to simplify sequencing of Pro rich peptides).⁸⁰ ECD has proved in many applications to be highly complementary to CAD⁸¹ and has to date been particularly effective in defining the position and structure of labile PTMs in proteins.⁸²⁻⁸⁶

ECD has also a demonstrated ability to differentiate IsoAsp from Asp in synthetic peptides.⁸⁷ The primary fragmentation during this process is cleavage of the N-C α bond which forms c/z \bullet or c \bullet /z cleavage pairs (Scheme 1.5). Because ECD cleaves the N-C α bond, and IsoAsp does not have an N-C α bond, this allows the differentiation of the two species, due to the different fragments produced, see scheme 1.5.

IsoAsp residues show fragment ions corresponding to cleavages on either side of the residue as well as the $c^{\bullet}+58/z-57$. Overall, ECD represents the fragmentation method of choice for sequencing peptides and proteins.



SCHEME 1.5: ECD mechanism for the formation of aspartyl $((M+nH)^{(n-1)+\bullet}$ and isoaspartyl ($c^{\bullet}+58$ and $z-57$) diagnostic fragment ions. Reprinted with permission from Sargaeva NP, Lin C, O'Connor PB. 2009. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid beta Peptides, Including A beta 1-42, Using Electron-Ion Reactions. *Analytical Chemistry* 81:9778-9786. Copyright © 2009 American Chemical Society.

ECD requires the precursor sample ions to be immersed in a dense population of low kinetic energy electrons, restricting its use to FTMS instruments. Some exceptions to this were reported in a couple of studies, where ECD has been achieved on an ion trap mass spectrometer.^{88, 89} Another limitation of ECD is its charge state restrictions, for example single charge would get neutralised leaving no more charges to be detected.

Six years after the appearance of ECD, researchers were looking for alternative ways of achieving ECD in instruments other than FT-ICR-MS; thus, in 2004 electron transfer dissociation (ETD) was introduced by Joshua Coon and John Syka.^{90, 91} ETD is similar to ECD except that small-molecule odd-electron (OE) anions transfer an electron to multiply charged cations to induce dissociation. ETD can be performed in instruments such as ion traps due to their ability to trap negative and positive ions simultaneously, to promote their interaction. The key to successful ETD relies on the electrophilic characteristics of the anion used for electron donation. Molecules like anthracene, fluorobenzene, etc., are selected based on their facility to donate an electron rather than extracting a proton from the cation. This fragmentation method has also allowed new types of isomer differentiation on peptides.^{92, 93}

The main disadvantage of this fragmentation technique is that it is generally performed on an ion trap, which has a poor mass resolution and accuracy compared to other instruments such as Orbitrap and FT-ICR-MS. The main advantage is the low cost of this type of instrumentation.

Due to the lack of mass resolution in the ion trap, the resultant ETD fragments are often also reacted with small gaseous molecules that promote proton transfer to reduce their charge state. Reducing the charge state makes the separation between the peaks of an isotopic cluster broader, which at the same time makes the spectrum less complicated to interpret and reduces the need for high resolution.

Infrared Multiphoton Dissociation (IRMPD)

IRMPD is a fragmentation technique similar to CAD. Ions are heated with a CO₂ laser (10.6 μm) until they dissociate. The fragmentation is originated due to energy redistribution through the molecule, where lowest energy bonds are the preferential cleavage site. Like CAD IRMPD generates mostly b and y ions. IRMPD is commonly performed on a FT-ICR-MS,^{64, 94} or ion traps.⁶⁵ The mechanism of fragmentation involves the absorption of infrared photons by ions trapped in the cell. In this process the molecule heats up until it dissociates, resulting in the cleavage of an amide bond. When the dissociation is performed on a FT-ICR-MS, the IR laser is aligned with the centre of the trap, which is the place where the ion cloud is normally located. After radiation, the fragments generated stay in the centre of the trap, where they are efficiently excited and then detected. The evident advantage of IRMPD is that no gas is needed, because no collision is involved, which keeps the pressure constant during dissociation. The use of IRMPD along with ECD is advantageous because each technique will complement the sequence coverage of the peptide in study. This is due to the facility of performing both fragmentation techniques simultaneously, which is not the case for CAD.

1.4. Differentiation of Isomeric Amino Acids Using Mass Spectrometry

Determination of isomeric products in biological environments using mass spectrometry represents a considerable challenge. For more than 30 years, researchers have tried numerous ways of performing analyses of this kind. Until recently, this has only been routinely possible using high-energy collisionally-activated dissociation (HE-CAD) on magnetic sector instruments.^{95, 96} The first evidence of isomer differentiation in peptides using this type of instrument was

reported in the 1980s.⁹⁶⁻⁹⁸ Since then many other attempts have been made to distinguish isomeric forms in peptides with varying success. Some of these attempts used fast atom bombardment, and entirely relied on signature immonium ions.⁹⁷⁻¹⁰¹ Other attempts were based on abundance ratio of b/y fragments.^{102, 103} Additional studies showed that b+H₂O peaks were observed with the IsoAsp residue and not in the Asp residue.^{102, 104}

High energy CAD has assisted in the differentiation of isomers using mass spectrometry since 1980. The fragmentation using high energy CAD results in b and y ions, but also c, z, d, a, b, v, and w ions, due to the excess of energy that results in secondary fragmentation, see Scheme 1.3 for structure of these ions. Thus, the technique allows differentiating between isomeric forms such as Ile/Leu in proteins and peptides.^{96, 105-107} However, commercial access to this type of fragmentation was limited due to the high cost, complexity, and low sensitivity of the instrumentation involved (magnetic sector instruments). More modern mass spectrometers were developed in such a way that only low-energy CAD was achievable (< 100 eV).^{105, 108} This range of energies is commonly used in mass spectrometers such as triple quadrupoles and ion traps. Typical gases used for collisions are argon, helium, and nitrogen.

Research carried out in the last decade has opened the door to new fragmentation methods with the capability to reliably discriminate between isomers according to the fragmentation patterns of peptides. The pioneering fragmentation method used to achieve these new types of isomer differentiation on peptides was ECD.¹⁰⁹ A considerable number of publications have been published since then using this

technique in proteomic studies. Its early application in isomer differentiation was explored by Zubarev and co-workers,¹¹⁰ when the energy of the electrons in the external cathode electron gun of a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer was increased from 0.2 eV (conventional ECD experiments) to 3-13 eV (hot ECD or HECD) producing secondary fragmentation that allowed structural discrimination between isomeric forms of Isoleucine/Leucine (Ile/Leu).^{110, 111} Recently Cooper and co-workers utilized the application of hot ECD in the field of medicine, to differentiate Ile from Leu and identified a new hemoglobin variant in blood samples.¹¹²

Other isomers such as Asp and IsoAsp play a significant role in the development of diseases like Alzheimer's, amyloidosis, etc.¹¹³ Currently, numerous papers are published annually in biological journals regarding Asp/IsoAsp formation in proteins and one systematic study has been published regarding the ratios of the rates of formation of these two products over time.^{114, 115} This study was a considerable effort which involved high resolution HPLC separation of peptides containing the two Asp isomers, as well as synthesis of each individual peptide in both isoforms to determine their peak elution times. O'Connor and co-workers reported the differentiation of Asp from IsoAsp in peptides^{87, 92} and proteins¹¹⁶ using conventional ECD experiments. The methodology was further applied to differentiation of Asp and IsoAsp in β -peptides^{117, 118} and also in the differentiation of α -glutamic (α -glu) and γ -Glu.¹¹⁹ Zubarev in 2009 also demonstrated the valuable application of ECD in proteome-scale identification of Asp isomerisation and Asn deamidation.^{120, 121} The application of ECD has also been extended to other isomers such as glu and γ -glu,¹¹⁹ ϵ -peptides,¹²² and stereoisomers.^{24, 25}

An extensive body of research in mass spectrometry of isomers using ECD has focused on the structural assignment of oligosaccharides.^{85, 123} Although the analysis of oligosaccharides, such as glycoprotein glycans, is an example of isomeric differentiation using mass spectrometry, a full approach of this topic is beyond the scope of this thesis. Joseph Zaia has compiled extensive summaries of this area.^{124,}
125

ETD provided an alternative technique to dissociate multiply charged peptides and proteins into the homologous c and z[•] type ion series fragments using reactions of multiply charged peptide/protein cations with negatively charged radical anions.^{90, 91} The ability to produce similar fragmentation to ECD with a relatively low cost instrument was a significant step forward in proteomic studies, particularly in the characterization of PTMs. ETD has also been applied to differentiation of isomers of Asp using the same diagnostic peaks as those observed with ECD.^{92, 93}

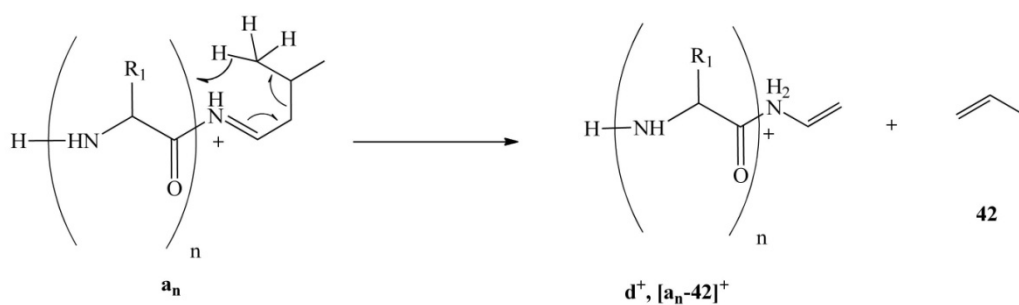
“Hot” ECD was introduced by Zubarev, in 2002.^{110, 111} In these studies it was found that the rate of dissociative capture was also significant using hot electrons (3-13 eV). In terms of energy, this reaction was well separated from the conventional ECD reaction typically carried out at around 0.2 eV, where significantly less fragmentation is observed. Besides the N-C_α bond cleavage, HECD also promotes other fragmentation pathways, cleaving more bonds than ECD and producing abundant secondary fragmentation. The secondary fragments are typically even-electron ions originated from excited z[•], as well as b and y ions, due to the excess of vibrational energy imparted by the higher energy electrons. These secondary ions are extremely useful and have been used to differentiate isomers such as Leu from Ile in

peptides.^{111, 112} A comprehensive approach of the role of ECD in proteomics has been reviewed by Cooper and co-workers¹²⁶, and can also be found in a review chapter by Zubarev and co-workers.¹²⁷

1.4.1. High-energy CAD and hot ECD applied in the differentiation of Ile from Leu

Several high-energy dissociation techniques have been developed that potentially allow Leu/Ile differentiation. Amongst these are high-energy CAD,^{96, 97, 128} hot electron capture dissociation (HECD),¹¹⁰⁻¹¹² VUV-photodissociation,¹²⁹ and, most recently, studies have reported the use of MALDI in-source decay followed by collisional cooling, that allows the differentiation of these isomers.^{130, 131}

Johnson showed that secondary fragmentation was achieved as a consequence of the high-energy collision (HE-CAD) with gas molecules.⁹⁶ The presence of a characteristic peak at 54 Da higher than that corresponding to an ion of type y, was present in the spectra of high energy CAD of peptides, using magnetic sector instruments.⁹⁶ This new secondary fragment type was assigned by Biemann¹⁰⁵ with the notation w. It was not frequently observed and it mainly occurred when the N-terminal amino acid of the z_n+1 ion was of the type that cannot undergo cleavage of the $C_\beta-C_\gamma$ bond, such as aromatic amino acids and when the next amino acid is one that readily cleaves at the $C_\beta-C_\gamma$ bond, such as Leu or Glu. The isoforms could be differentiated based on the principle that Leu can only form one type of w and d ion by losing propyl from the side chain, while Ile does not. See Scheme 1.6.



SCHEME 1.6: Proposed fragmentation for the elucidation of Leu and IsoLeu using high-energy CAD. Adapted from Biemann, K. and S.A. Martin. Mass spectrometric determination of the amino acid sequence of peptides and proteins. *Mass Spectrometry Reviews*. **1987**, 6, 1-75.

An earlier study reported by Aubagnac in 1985 using FAB high energy CAD showed that Leu/Ile could be differentiated using the MS/MS spectra of the m/z 86 immonium ions corresponding to the Leu and Ile residues. The limitation of this approach was for peptides containing only one form of either Ile or Leu, because, in the FAB spectra of peptides containing both, the ion at m/z 86 is usually a mixture of isomeric immonium ions formed from both residues.⁹⁷

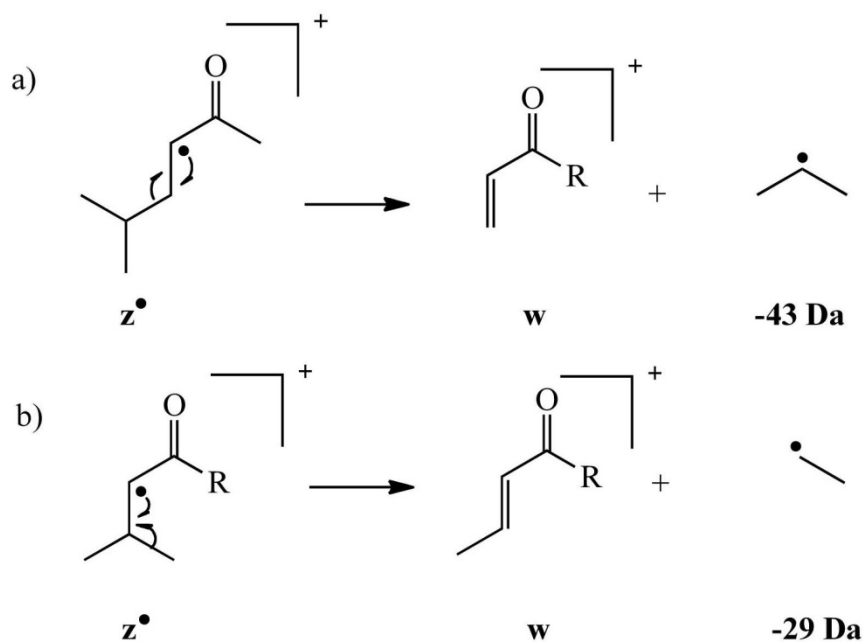
Later approaches to address this issue were proposed by Biemann's group in 1988. They used the nominal loss of an olefin from the side chain of the C-terminal amino acid of an a_n ion, the $(a_n-42)^+$ ion in the case of Leu and the $(a_m-28)^+$ ion in the case of Ile. These ions, well known as d-series ions, were present in the MS/MS spectrum of protonated molecules $[\text{MH}]^+$ of Substance P. This approach was limited to cases where both the a-series fragment and the diagnostic d-series were sufficiently intense to be recognized.¹⁰⁶

In 2002, Zubarev and co-workers reported the use of more energetic electrons in electron capture dissociation (HECD) that allowed them to induce secondary

fragmentation of amino acid side chains of a^{\bullet} and z^{\bullet} ions, thereby producing d and w ions. The loss of $\bullet\text{CH}(\text{CH}_3)_2$ (-43 Da) or $\bullet\text{CH}_2\text{CH}_3$ (-29 Da) from a^{\bullet} and z^{\bullet} ions containing N-terminal Leu or Ile, respectively enables the two amino acids to be distinguished, Scheme 1.7.^{110, 111}

Macht and co-workers¹²⁸ reported the presence of w -fragment ions as well as d -fragment ions corresponding to Leu and Ile in a MALDI-TOF/TOF with HE-CAD. Since Leu can only form one type of w and d ions, by losing propyl from the side chain, Leu was differentiated unambiguously from Ile. More recently in 2009, Zhang and co-workers also observed the presence of w ions corresponding to Leu by using photodissociation with 157 nm light in a commercial tandem TOF-MS.¹²⁹ In the spectra obtained for the peptide VNVDEVGGEALGR, the presence of x_3 and y_3 indicated the presence of either Ile/Leu; alternatively, the presence of w_3 certainly indicated the presence of Leu.

In 2009 Cooper and co-workers reported a new variant of hemoglobin that was very difficult to precisely identify using previous ESI-MS procedures such as CAD triple quadrupole (tripleQ), where only 95 % of the variants encountered could be assigned. In their study, the use of HECD in a FT-ICR allowed them to successfully distinguish between Leu and Ile residues in human hemoglobinopathies without the need of performing DNA analysis.¹¹² This remarkable study shows that HECD represents a useful tool in the analysis of a considerable number of hemoglobin variants that involve amino acid exchanges governed by single mutations in the nucleotide codon to either Leu or Ile.



SCHEME 1.7: Secondary fragmentation of a z ion to produce the w ion from a peptide containing (a) a Leu residue and (b) an Ile residue. Reprinted with permission from Kjeldsen F, Haselmann KF, Sorensen ES, Zubarev RA. 2003. Distinguishing of Ile/Leu amino acid residues in the PP3 protein by (hot) electron capture dissociation in Fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* 75:1267-1274. Copyright © 2003 American Chemical Society.

In 2010, Soltwisch and Dreisewerd showed how making use of in-source decay fragmentation (ISD) in MALDI-TOF is also a valuable tool for differentiating the two isomers.¹³⁰ Here the species were generated in a fast charge-remote dissociation process which is similar to electron capture (ECD) and electron transfer dissociation (ETD). However, because it is initiated by the transfer of neutral hydrogen rather than a charged particle, it can be used to analyze singly protonated as well as negatively charged ions. This process produces a small percentage of diagnostic fragment d and w ions, but they show a consistency and reliability for assignment of either Leu or Ile.¹³¹

Another amino acid, which differs in mass by only 0.03638 Da, from Leu/Ile is hydroxyproline (Hyp). This amino acid is produced by the hydroxylation of Pro during the post-translational modification of proteins by prolyl hydroxylase. Hyp is found only rarely in most proteins, but in collagen it constitutes about 10% of all amino acids.¹³² In 1990 Biemann and co-workers reported the differentiation amongst Leu/Ile and Hyp, the masses of which differ by 36 mDa, using high-energy CAD. They observed that differentiation was only possible if an N-terminal basic moiety was present in the peptide. This is due to the ability of this basic moiety to provide ions that retain their charge at the N-terminus.¹⁰⁷ This was the only way of distinguishing between these amino acids until the development of modern high resolution mass spectrometers, such as the FT-ICR.

1.4.2. Use of high-energy CAD, ECD, and ETD to differentiate Asp from IsoAsp in proteins and peptides

Deamidation is, perhaps, the most common post-translational modification in proteins and results in a mixture of isomers of Asp and IsoAsp. It is associated with many diseases, particularly those which involve problems with protein folding. Due to the difficulty of distinguishing Asp and IsoAsp in proteins, insufficient research has been performed regarding their formation or impact on diseases at a molecular level, although there is extensive correlation and there is much speculation.¹¹³

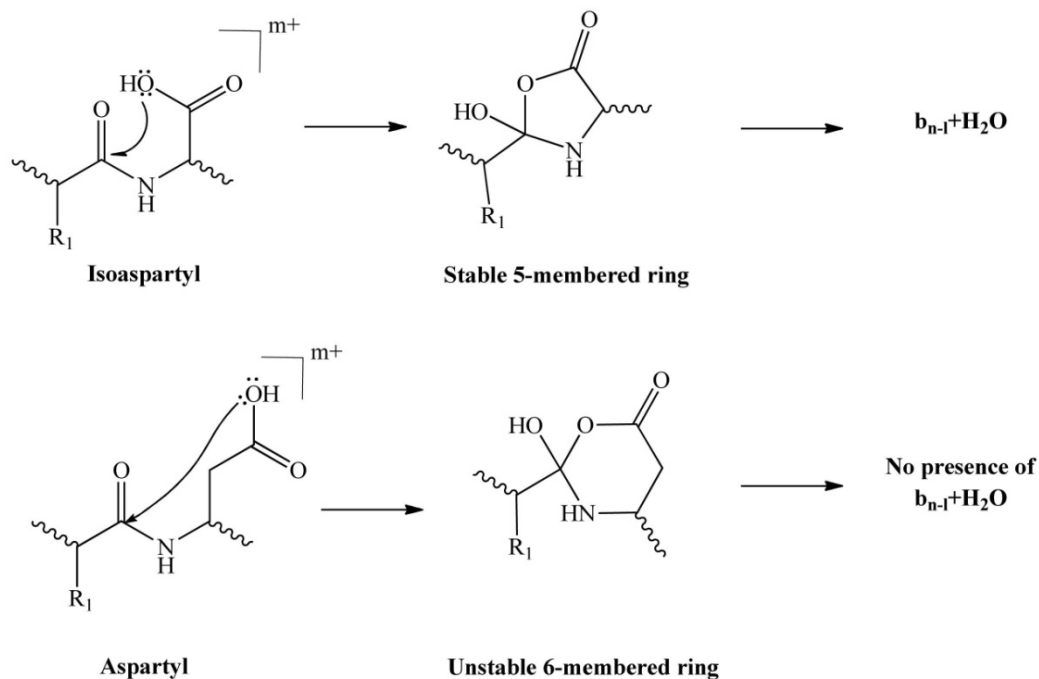
Early methods were developed in order to differentiate the isomers products of deamidation. In 1988 Lloyd and co-workers used fragment ion abundance ratios to differentiate Asp from isoAsp residues in the MS/MS spectra of peptides using a double focusing magnetic sector mass spectrometer.¹⁰⁰ Their studies were based on

the experimental observations of Okada and Kawase in 1977, where the intensities of certain fragments corresponding to α -Glu and γ -Glu, (see Scheme 1.2 for structures), were a function of the type of linkage present in the protein.¹³³ The most relevant observation was the facile loss of CO from the b type fragment ion in the α -glu case and the absence of any CO loss from the b type fragment ion formed in the γ -glu case. The conclusions to these observations were based on the stability of the carbocation formed in the two cases: a less stable primary carbocation in the γ -glu case and a secondary cation in the case of α -glu. The same principle was applied by Lloyd and co-workers in Asp/IsoAsp differentiation. They observed the loss of CO from b_{n+1} to form a_{n+1} while the same loss was much less abundant for the isoAsp versions ($b_{n+1}/a_{n+1}(\text{Asp}) > b_{n+1}/a_{n+1}(\text{isoAsp})$); these results are also explained by the same carbocation stability theory. The structure proposed by Lloyd for the a_1 ion in the MS/MS spectrum of the peptide DIRKF-NH₂, which was missing in the spectrum of the isoAsp form, was actually the immonium ion for an Asp residue ($m/z = 88$). Immonium ions are fragments containing single side chains or backbone chain formed by a combination of *a* type and *y* type cleavage.

Later, Papayannopoulos and Carr^{134, 135} reported the use of MS to differentiate between Asp and IsoAsp. The differentiation was achieved by the presence of the fragment corresponding to the isoAsp $[b_{n-1} + \text{H}_2\text{O}]^+$ in the spectrum, where *n* is the position of the Asp/IsoAsp; this fragment was not found in the spectrum corresponding to Asp form of the same peptide. The same fragment was also used to differentiate the isoAsp from Asp forms of a deamidated proteolytic peptide, from hirudin, an anticoagulant peptide.¹³⁶ The formation of the fragment $[b_{n-1} + \text{H}_2\text{O}]^+$ could be explained using the mechanism proposed by Papayannopoulos,¹³⁴ where

nucleophilic attack of the carboxyl OH group of β -Asp at the adjacent carbonyl can be assumed to form an oxazolidone intermediate, as seen on Scheme 1.8. This intermediate should easily decompose to the observed $[b_{n-1} + H_2O]^+$ ion, along with CO and a C-terminal aldimine-type moiety that represent a pair of highly stable neutral species. Since the isoAsp side chain resembles the C-terminus, generation of $[b_{n-1} + H_2O]^+$ for isoAsp residues is suggested to resemble the fragmentation channel shown to occur in the low-energy MS/MS spectra of peptides where a C-terminal hydroxyl rearrangement generates $[b_{n-1} + H_2O]^+$ fragment ions. Schindler and co-workers managed to prove the mechanism proposed by Papayannopoulos with respect to the hydroxyl transfer. This was achieved by using ^{18}O labelling of the peptide and then performing MS/MS. The MS/MS of the ion $[b_{n-1} + H_2O]^+$ showed a 1:1 ratio of $^{16}O:^{18}O$, which proves that there is a migration of $-OH$ from the side chain to the diagnostic fragment.¹³⁶

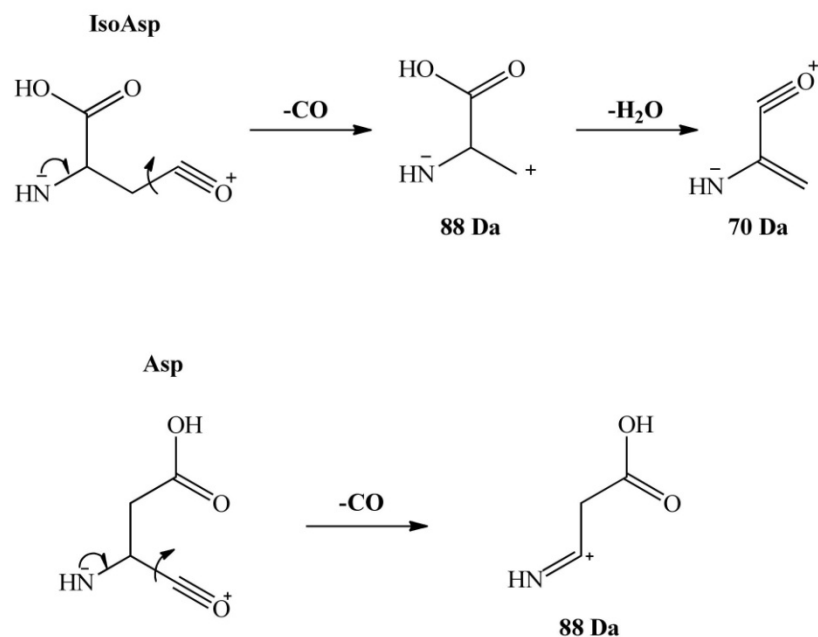
In 2000, Lehmann and co-workers indirectly determined the presence of isoaspartyl residues in peptides. This was achieved based on the relative abundance of fragment ions.¹³⁷ In general, the product ion spectra of peptides with isoAsp sites exhibit different characteristics than their Asp-containing analogues. A significant difference was observed in the spectra of peptides containing Asp and IsoAsp. As a general trend, the presence of isoAsp enhanced y-ion formation and reduced b-ion formation at the cleavage site.



SCHEME 1.8: Mechanism proposed for the attack of $-OH$ from the IsoAsp/Asp side chain to the carbonyl backbone. Adapted from Papayannopoulos, I.A. and K. Biemann. Fast atom bombardment and tandem mass-spectrometry of synthetic peptides and by-products. *Peptide Research*. **1992**, 5, 83-90.

This tendency is observed due to a competition between forming the oxazolone containing the N-terminus (b ion) and direct cleavage to form a terminal amine containing the C-terminus (y ion). In the case of peptides containing Asp, the oxazolone is formed as a five-membered ring due to nucleophilic attack at an Asp side chain, while in a peptide containing isoAsp the result is a less favorable six-membered ring, resulting in reduced b-ion abundance. Immonium ion intensities were also investigated as candidates to differentiate the isoforms.¹³⁸ On average the immonium ion corresponding to Asp was 5.5 times higher than the one for the isoAsp form.

In 2000, a study by Gonzalez and co-workers showed that the $[b_{n-1}+H_2O]^+$ fragment ion and its complement, the $[y_{1-n}-46]^+$ fragment ion, can be used to differentiate IsoAsp from Asp residues in sets of synthetic peptides, including D7 and D23 of β -Amyloid peptide analogues.¹⁰² The data showed that the $[b_n+H_2O]^+$ ion intensity was higher if a basic amino acid or the N-terminus was on the N-terminal side of the isoAsp residue. Thus the resulting fragment ions were highly controlled by the position of the basic amino acid. Otherwise, the rearrangements to generate the diagnostic ions were not feasible, due to the lack of intermolecular interaction between the side chain of the basic amino acid and the isoAsp residues. In addition to this limitation, it was also noted that the diagnostic ion, $[b_n+H_2O]^+$ from the tryptic peptides, was frequently undetected or at low abundance. This is because the ESI-MS/MS of tryptic peptides are dominant for C-terminal ions (mainly “y” ions), while the N-terminal ions usually appear with much lower intensity. In these cases immonium ions were used to differentiate the isomers. The IsoAsp spectrum also showed a signal at 70 Da, which was not detected for the Asp isoform. A mechanism for the formation of this immonium ion was proposed via a rearrangement of a primary carbocation (unstable structure). The loss of water and a further rearrangement yields a more stable acylium ion ($C_3H_4NO^+$), detected at m/z 70.02928 (see Scheme 1.9). Regardless, these immonium ions could not be used to predict the presence of an isoAsp residue because Pro also yields an intense immonium ion ($C_4H_8N^+$) at m/z 70.06567.



SCHEME 1.9: Proposed structures for the formation of the IsoAsp/Asp immonium ions. Adapted from Gonzalez, L.J., T. Shimizu, Y. Satomi, L. Betancourt, V. Besada, G. Padron, R. Orlando, T. Shirasawa, Y. Shimonishi, and T. Takao. Differentiating alpha- and beta-aspartic acids by electrospray ionization and low-energy tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*. **2000**, *14*, 2092-2102.

Another way of identifying the isomers in proteins and peptides was proposed by O'Connor and co-workers.^{87, 116, 139} The first report showed the development of a methodology to reliably identify the isomers in synthetic peptides by using ECD as the fragmentation method.⁸⁷ Besides the conventional ECD N-C_α bond cleavage leading to complementary c and z[•] fragments, the authors also detected a diagnostic cleavage giving c_n[•] + 58.0054 (C₂H₂O₂) and z_n-56.9976 (C₂HO₂) (see Fig. 1.16 and 1.17), where n is the position of the isoaspartyl residue. Figures 1.16 and 1.17 show that IsoAsp generated peaks corresponding to c[•]+ 58 and z-57 that were observed for BUSM₂ (RAAGAD_βGD_βGAGAD_βAR) and not for BUSM₁ (RAAAGAD_αGD_αGAGAD_αAR). The c[•]+ 58 and z-57 diagnostic fragments are usually less abundant than the adjacent conventional c_n and z_n species. The same fragments were observed with Electron Transfer Dissociation (ETD).^{93, 140, 141} The

generation of $c^{\bullet}+58$ and $z-57$ indicates the presence and position of isoAsp, and the generation of $[M+nH]^{(n-1)+\bullet}-60$ indicates the presence (but not the position) of Asp. A more comprehensive study was performed by the same group in 2006 when they proved that the methodology could be used to differentiate the isomers in tryptic digested proteins.¹¹⁶ In 2009 Sargaeva and co-workers demonstrated the use of the diagnostic fragments generated by ECD to identify the isomers in the whole amyloid β 1-42 fragment, which corresponds to the most abundant component in plaque deposits in patients with Alzheimer's.

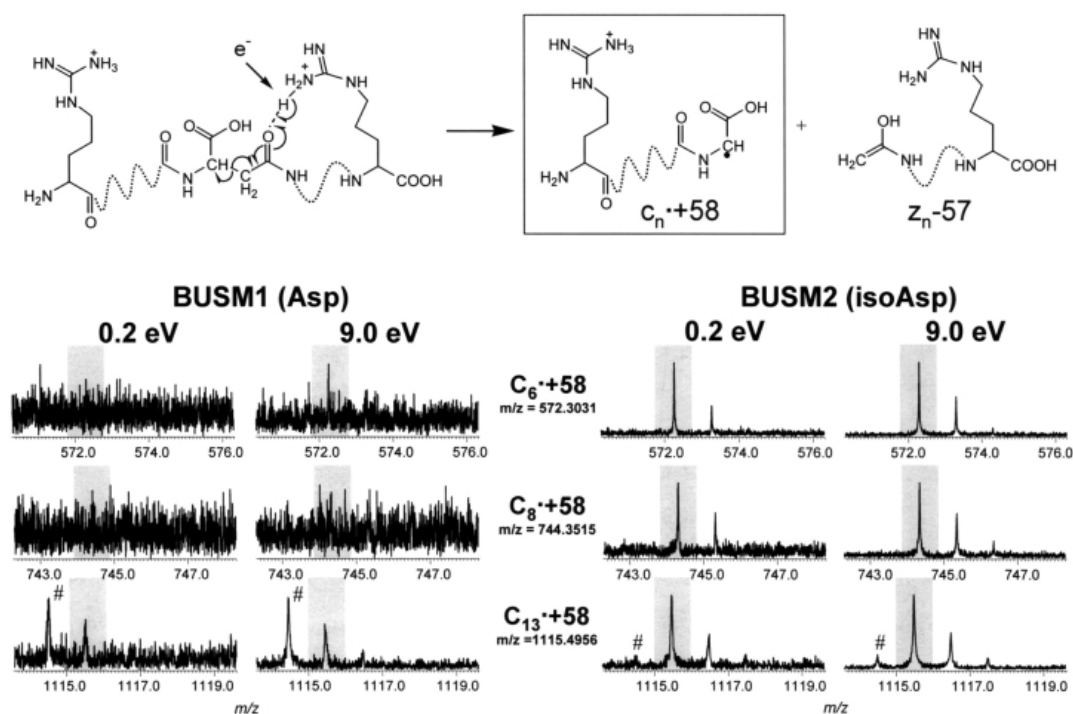


FIGURE 1.16: ECD spectrum of two peptides BUSM1 and BUSM2, the shaded regions represent the $c^{\bullet}+58$ corresponding to the presence of isoAsp. Reprinted with permission from Cournoyer JJ, Pittman JL, Ivleva VB, Fallows E, Waskell L, Costello CE, O'Connor PB. 2005. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Science* 14:452-463. Copyright © 2005 John Wiley and Sons.

Little has yet been done in terms of quantitation of Asp/IsoAsp isomers using ECD. A report by Cournoyer in 2007 showed the use of a two point calibration curve to achieve relative quantitation of the isomers, which also proved the linearity of the

plots independent of the sequence of the small synthetic peptides. Construction of this calibration curve was partially based on the evidence that Asn deamidation in linear peptides produces a mixture of IsoAsp/Asp at a ratio of 3:1.^{20, 104, 142}

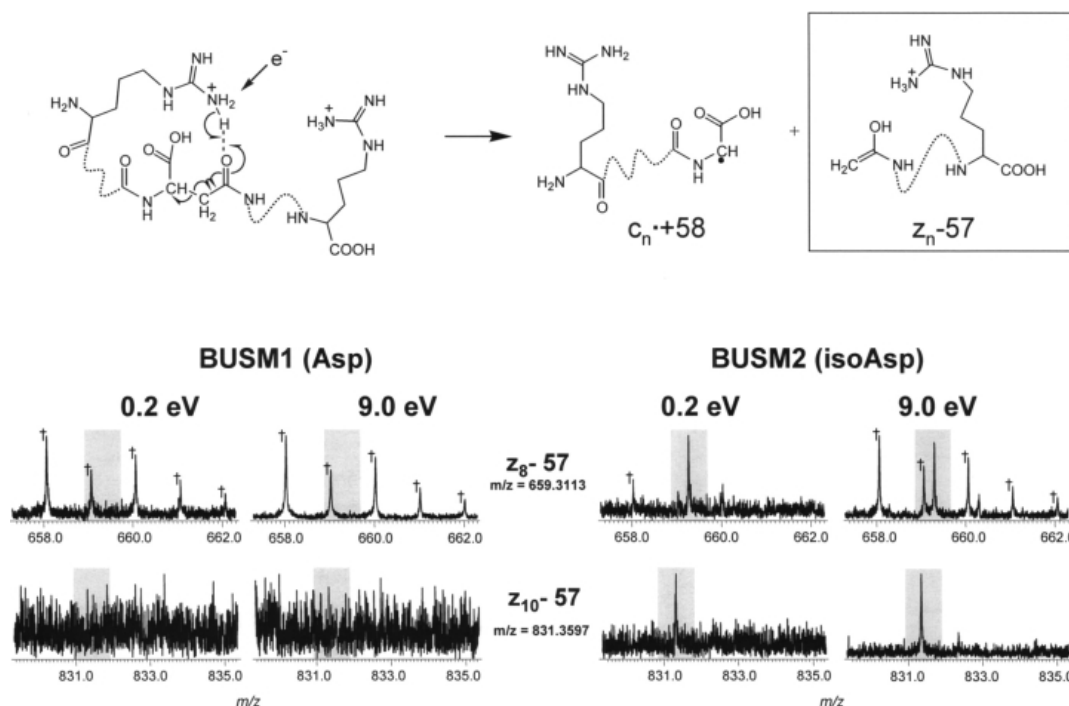


FIGURE 1.17: ECD spectrum of two peptides BUSM1 and BUSM2, the shaded regions represent the $z_8 - 57$ corresponding to the presence of isoAsp. Reprinted with permission from Cournoyer JJ, Pittman JL, Ivleva VB, Fallows E, Waskell L, Costello CE, O'Connor PB. 2005. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation. *Protein Science* 14:452-463. Copyright © 2005 John Wiley and Sons.

The authors used this evidence to suggest that rapid aging of a peptide with an asparaginyl residue released from a protein by enzyme digestion should produce the ~75% isoAsp mixture that can be used for rough calibration. The results proved this hypothesis, but the method was only applicable for quantification of IsoAsp produced by deamidation, not for Asp isomerisation. Also, it is not yet clear if this approach could be applied to other peptides, but it is suggested to be a good approximation.

Despite the fact that ETD and ECD have shown to be reliable in the differentiation of Asp from IsoAsp in peptides and proteins, they are less commercially available than low-energy CAD. Thus researchers are developing alternative methodologies that can allow them to distinguish between these isomers using more common fragmentation methods.¹⁴³ Also, other studies available in the literature have attempted to differentiate the two isomers using negative ion fragmentation,¹⁴⁴ however the results from this study showed that the ion abundance corresponding to α and β backbone cleavage were remarkably low and occasionally absent.

1.4.3. Differentiation of α -Glu and γ -Glu residues in peptides

Like Asn, Gln (Q) deamidates in proteins and peptides by direct hydrolysis, under acidic conditions and proceeds via a glutarimide intermediate at neutral or alkaline conditions.¹ The reaction is much slower than Asn deamidation²³ with the exception of when Gln is located at the N-terminus, under which condition it deamidates more rapidly than Asn to form pyroglutamic acid.¹⁴⁵ The deamidation rates of pentapeptides containing Asn and Gln have been theoretically calculated in previous studies²³ and the general trend shows that the half-lives of the sequence dependent Asn peptide deamidation at neutral pH and physiological temperature range from 0.5 to 500 days, while those of Gln range from 600 to 20,000 days.

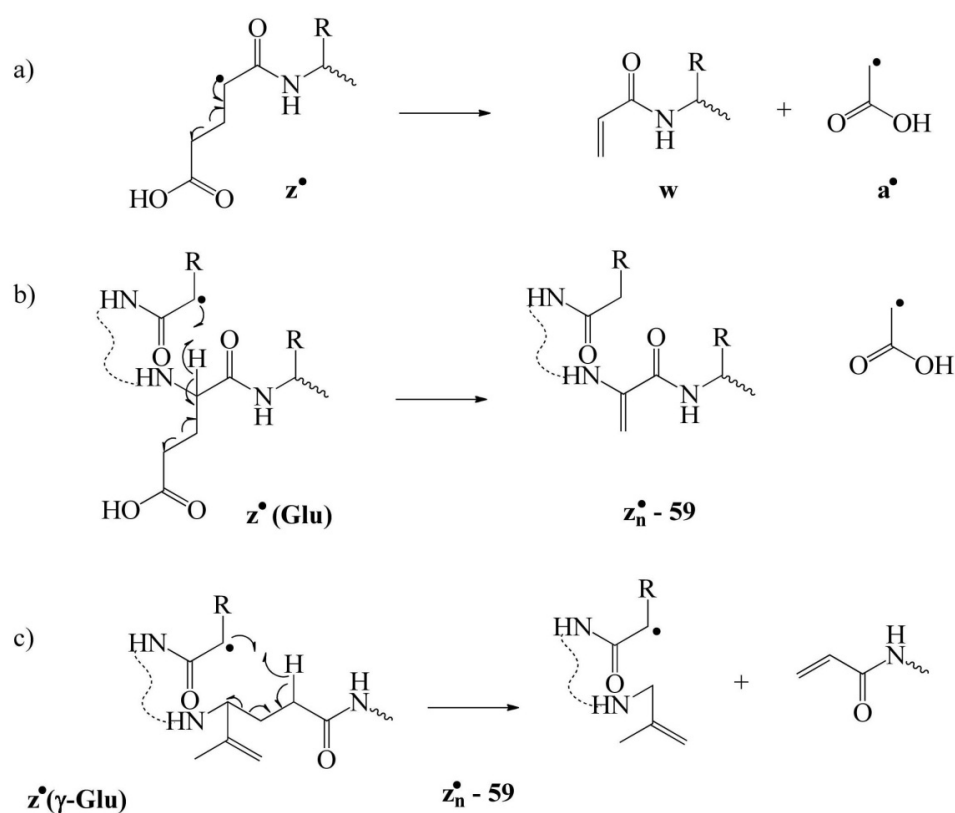
Although Asn deamidation has been extensively studied, there are a limited number of Gln studies in the literature.^{19, 23, 145} The only general results that emerged from these studies were the lower deamidation rate at the Gln site, but no indication of a reaction pathway had been reported, until 1991, when Capasso and co-workers reported the first evidence of the formation of a glutarimide intermediate and its deamidation products, α - and γ -glutamate (Glu, E).¹⁴⁶ However, the identification of

the Gln deamidation sites is straightforward; similar to Asn, it can be located by a 1 Da mass shift per site.

The differentiation of the deamidation products (α -Glu and γ -Glu) using mass spectrometry has remained largely unexplored. In 1977 Okada and Kawase were the first to find that the intensities of certain fragments corresponding to α - and γ -glutamyl were a function of the type of linkage present in the protein.¹³³ Again, much like Asn, the most relevant observation was the facile loss of CO from the b type fragment ion in the α -glutamyl case and the absence of CO loss from the b type fragment ion formed in the γ -glutamyl case. The conclusions to these observations were based on the stability of the carbocation formed in both cases; a less stable primary carbocation in the γ -glutamyl case, and a secondary cation in the case of α -glutamyl. Later, Nagasawa and Magnan used chemical ionisation mass spectrometry to identify the isomers.¹⁴⁷ In their studies the separation of the isomers was achieved by ion exchange chromatography followed by chemical ionisation. At the time this was the only possible way of differentiating between the two isomers. Unfortunately all these methods were limited to small peptides with less than five amino acids, and most of them can only be applied to peptides with a glutamyl residue located at the N-terminus when it can generate the cyclic structure for identification.

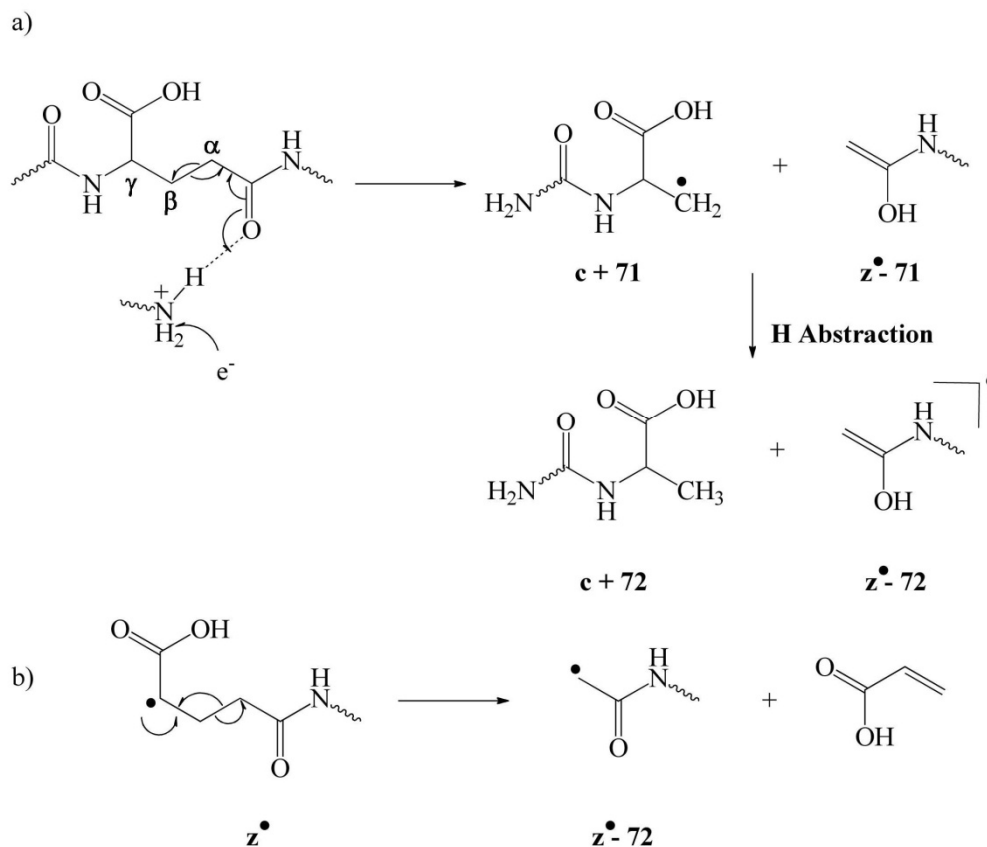
Recently, studies reported by Li and co-workers in 2010 show the differentiation of the isomers using ECD.¹¹⁹ In their studies, an ECD methodology was developed to directly differentiate the α - from the γ -glu residues in three human eye lens crystalline peptides in their potentially deamidated forms. The methodology was entirely based on the Asp/IsoAsp differentiation method also developed by the same

group.⁸⁷ They found that the presence of $z^{\bullet}-72$ ions can be used to both identify the existence and locate the position of the γ -Glu residues.



SCHEME 1.10: a) Direct partial chain loss (w Ion formation), b) through-space radical rearrangement, and c) mechanism for the formation of site specific $z_n^{\bullet}-59$ ions in ECD of the γ Glu. Reprinted with permission from Li XJ, Lin C, O'Connor PB. 2010. Glutamine Deamidation: Differentiation of Glu and gamma-Glu in Peptides by Electron Capture Dissociation. *Anal Chem* 82:3606-3615. Copyright © 2010 American Chemical Society.

They also reported that when the peptide contains a charge carrier near its N-terminus, the $c+57$ and $c+59$ ions may also be generated at the γ -Glu residue. The presence of Glu residue or (α -Glu) was inferred from the observation of a series of $z^{\bullet}-59$ ions. Scheme 1.10 and 1.11 show the mechanisms proposed by this group for the formation of these diagnostic ions.



SCHEME 1.11. Proposed mechanism for the formation of site specific $\text{z}^\bullet - 72$ ions in ECD of the γ Glu containing peptide via a) primary backbone cleavage and b) secondary fragmentation. Reprinted with permission from Li XJ, Lin C, O'Connor PB. 2010. Glutamine Deamidation: Differentiation of Glu and gamma-Glu in Peptides by Electron Capture Dissociation. *Anal Chem* 82:3606-3615. Copyright © 2010 American Chemical Society.

1.5. Aims and Objectives

The work presented in this thesis involved the use of high resolution mass spectrometry (FT-ICR-MS) to study deamidation of Asn from bovine collagen I, and also the differentiation of the deamidation products (IsoAsp and Asp) by using ECD.

The use of FT-ICR-MS to study deamidation of collagen is a relatively a new approach. The interest in this particular protein originated from the research carried out by Matthew Collins and co-workers at the University of York. They observed that deamidation of Asn and Gln was remarkably increased in collagen from ancient mammal bones. Since deamidation of Asn is easily induced by sample preparation, researchers have stayed away from its use for potentially assessing the age of ancient bone. To address this concern a methodology is needed to reliably measure the extent of deamidation before sample preparation. The methodology is introduced in this thesis and will be discussed in depth in Chapters 2 and 3. Since deamidation produces IsoAsp and Asp, the differentiation of the isomers in ancient bone will be presented in Chapter 5.

The aims of this project were to:

1. Find particular peptides within the sequence of collagen that can serve as markers for deamidation, and that can also be reliably identified in bovine collagen I from different sources.
2. Develop a methodology to measure the extent of deamidation before sample preparation in bovine collagen fragments.
3. Apply the methodology to real samples, ancient bones in this case.

4. Differentiate the isomers products of deamidation using ECD that could be use as alternative methodology to measure the extent of deamidation of Asn.

1.6. Research Papers

The research presented in this thesis has resulted in three publications

- **Perez P.**, Lam P., Kilgour D., Bristow A., McBride E., O'Connor P., Use of High Resolution Mass Spectrometry for the Analysis of Polymeric Excipients in Drug Formulations. *Analytical Chemistry*, **2012**.
- **Perez P.** , O'Connor P. Deamidation of Collagen. *Analytical Chemistry*. **2012**, 84 (6), 3017–3025.
- **Perez P.** , O'Connor P. Differentiation of Isomeric Amino Acid Residues in Proteins and Peptides Using Mass Spectrometry. *Mass Spectrometry Reviews*. Article first published online: 9 Feb **2012**.

1.7. References

1. Robinson, N.E. and Robinson, A.B., *Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*. Althouse Press: Cave Junction, OR, **2004**.
2. Miesbauer, L.R., Zhou, X., Yang, Z., Sun, Y., Smith, D.L., and Smith, J.B. Post-Translational Modifications of Water-Soluble Human Lens Crystallins from Young Adults. *Journal of Biological Chemistry*. **1994**, 269, 12494-502.
3. David, L.L., Lampi, K.J., Lund, A.L., and Smith, J.B. The Sequence of Human B1-Crystallin Cdna Allows Mass Spectrometric Detection of B1 Protein Missing Portions of Its N-Terminal Extension. *Journal of Biological Chemistry*. **1996**, 271, 4273-4279.
4. Lampi, K.J., Ma, Z., Shih, M., Shearer, T.R., Smith, J.B., Smith, D.L., and David, L.L. Sequence Analysis of Ba3, Bb3, and Ba4 Crystallins Completes the Identification of the Major Proteins in Young Human Lens. *Journal of Biological Chemistry*. **1997**, 272, 2268-2275.
5. Smith, J.B., Miesbauer, L.R., and Sun, Y.P. Post Translational Modification of Crystallins Isolated from Human Lenses. *Pure and Applied Chemistry*. **1994**, 66, 95-100.
6. Hanson, S.R.A., Smith, D.L., and Smith, J.B. Deamidation and Disulfide Bonding in Human Lens Γ -Crystallins. *Experimental Eye Research*. **1998**, 67, 301-312.
7. Lund, A.L., Smith, J.B., and Smith, D.L. Modifications of the Water-Insoluble Human Lens A-Crystallins. *Experimental Eye Research*. **1996**, 63, 661-672.
8. Zhang, Z.L., Smith, D.L., and Smith, J.B. Human Beta-Crystallins Modified by Backbone Cleavage, Deamidation and Oxidation Are Prone to Associate. *Experimental Eye Research*. **2003**, 77, 259-272.
9. Hanson, S.R.A., Hasan, A., Smith, D.L., and Smith, J.B. The Major in Vivo Modifications of the Human Water-Insoluble Lens Crystallins Are Disulfide Bonds, Deamidation, Methionine Oxidation and Backbone Cleavage. *Experimental Eye Research*. **2000**, 71, 195-207.
10. Bloemendal, H., De Jong, W., Jaenicke, R., Lubsen, N.H., Slingsby, C., and Tardieu, A. Ageing and Vision: Structure, Stability and Function of Lens Crystallins. *Progress in Biophysics & Molecular Biology*. **2004**, 86, 407-485.
11. Wistow, G.J. and Piatigorsky, J. Lens Crystallins - the Evolution and Expression of Proteins for a Highly Specialized Tissue. *Annual Review of Biochemistry*. **1988**, 57, 479-504.
12. Wilmarth, P.A., Tanner, S., Dasari, S., Nagalla, S.R., Riviere, M.A., Bafna, V., Pevzner, P.A., and David, L.L. Age-Related Changes in Human Crystallins Determined from Comparative Analysis of Post-Translational Modifications in Young and Aged Lens: Does Deamidation Contribute to Crystallin Insolubility? *Journal of Proteome Research*. **2006**, 5, 2554-2566.
13. Leo, G., Andreotti, A., Marino, G., Pucci, P., Colombini M. P., and Birolo, L. Deamidation at Asparagine and Glutamine as a Major Modification Upon Deterioration/Aging of Proteinaceous Binders in Mural Paintings. *Analytical Chemistry*. **2011**, 83, 2056-2064.
14. Araki, N. and Moini, M. Age Estimation of Museum Wool Textiles from Ovis Aries Using Deamidation Rates Utilizing Matrix-Assisted Laser

- Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2011**, 25, 3396-3400.
15. Wilson, J., Van Doorn, N.L., and Collins, M.J. Assessing the Extent of Bone Degradation Using Glutamine Deamidation in Collagen. *Analytical Chemistry*. **2012**, 84, 9041-9048.
 16. Van Doorn, N.L., Wilson, J., Hollund, H., Soressi, M., and Collins, M.J. Site-Specific Deamidation of Glutamine: A New Marker of Bone Collagen Deterioration. *Rapid Communications in Mass Spectrometry*. **2012**, 26, 2319-2327.
 17. Buckley, M., Walker, A., Ho, S.Y.W., Yang, Y., Smith, C., Ashton, P., Thomas-Oates, J., Cappellini, E., Koon, H., Penkman, K., Elsworth, B., Ashford, D., Solazzo, C., Andrews, P., Strahler, J., Shapiro, B., Ostrom, P., Gandhi, H., Miller, W., Raney, B., Zylber, M.I., Gilbert, M.T.P., Prigodich, R.V., Ryan, M., Rijdsdijk, K.F., Janoo, A., and Collins, M.J. Comment on "Protein Sequences from Mastodon and Tyrannosaurus Rex Revealed by Mass Spectrometry". *Science*. **2008**, 319.
 18. Brock, F., Wood, R., Higham, T.F.G., Ditchfield, P., Bayliss, A., and Ramsey, C.B. Reliability of Nitrogen Content (%N) and Carbon:Nitrogen Atomic Ratios (C:N) as Indicators of Collagen Preservation Suitable for Radiocarbon Dating. *Radiocarbon*. **2012**, 54, 879-886.
 19. Robinson, N.E., Robinson, Z.W., Robinson, B.R., Robinson, A.L., Robinson, J.A., Robinson, M.L., and Robinson, A.B. Structure-Dependent Nonenzymatic Deamidation of Glutamyl and Asparagyl Pentapeptides. *Journal of Peptide Research*. **2004**, 63, 426-436.
 20. Geiger, T. and Clarke, S. Deamidation, Isomerization, and Racemization at Asparagyl and Aspartyl Residues in Peptides. *Journal of Biological Chemistry*. **1987**, 262, 785-794.
 21. Aswad, D.W., Paranandi, M.V., and Schurter, B.T. Isoaspartate in Peptides and Proteins: Formation, Significance, and Analysis. *Journal of Pharmaceutical and Biomedical Analysis*. **2000**, 21, 1129-1136.
 22. Radkiewicz, J.L., Zipse, H., Clarke, S., and Houk, K.N. Accelerated Racemization of Aspartic Acid and Asparagine Residues Via Succinimide Intermediates: An Ab Initio Theoretical Exploration of Mechanism. *Journal of the American Chemical Society*. **1996**, 118, 9148-9155.
 23. Robinson, N.E. and Robinson, A.B. Molecular Clocks. *Proceedings of the National Academy of Sciences of the United States of America*. **2001**, 98, 944-949.
 24. Adams, C.M., Kjeldsen, F., Zubarev, R.A., Budnik, B.A., and Haselmann, K.F. Electron Capture Dissociation Distinguishes a Single D-Amino Acid in a Protein and Probes the Tertiary Structure. *Journal of the American Society for Mass Spectrometry*. **2004**, 15, 1087-1098.
 25. Adams, C.M.Z., R. A. Distinguishing and Quantifying Peptides and Proteins Containing D-Amino Acids by Tandem Mass Spectrometry. *Analytical Chemistry*. **2005**, 77, 4571-4580.
 26. Stroop, S.D. A Modified Peptide Mapping Strategy for Quantifying Site-Specific Deamidation by Electrospray Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2007**, 21, 830-836.
 27. Robinson, N.E., Robinson, A.B., and Merrifield, R.B. Mass Spectrometric Evaluation of Synthetic Peptides as Primary Structure Models for Peptide and Protein Deamidation. *Journal of Peptide Research*. **2001**, 57, 483-493.

28. Schmid, D.G., Von Der Mulbe, F., Fleckenstein, B., Weinschenk, T., and Jung, G. Broadband Detection Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry to Reveal Enzymatically and Chemically Induced Deamidation Reactions within Peptides. *Analytical Chemistry*. **2001**, *73*, 6008-6013.
29. Perez Hurtado, P. and O'connor, P.B. Deamidation of Collagen. *Analytical Chemistry*. **2012**, *84*, 3017-3025.
30. Zomber, G., Reuveny, S., Garti, N., Shafferman, A., and Elhanany, E. Effects of Spontaneous Deamidation on the Cytotoxic Activity of the Bacillus Anthracis Protective Antigen. *Journal of Biological Chemistry*. **2005**, *280*, 39897-39906.
31. Solstad, T., Carvalho, R.N., Andersen, O.A., Waidelich, D., and Flatmark, T. Deamidation of Labile Asparagine Residues in the Autoregulatory Sequence of Human Phenylalanine Hydroxylase. *European Journal of Biochemistry*. **2003**, *270*, 929-938.
32. Carvalho, R.N., Solstad, T., Bjørge, E., Barroso, J.F., and Flatmark, T. Deamidations in Recombinant Human Phenylalanine Hydroxylase: Identification of Labile Asparagine Residues and Functional Characterization of Asn → Asp Mutant Forms. *Journal of Biological Chemistry*. **2003**, *278*, 15142-15152.
33. Robinson, N.E., Zabrouskov, V., Zhang, J., Lampi, K.J., and Robinson, A.B. Measurement of Deamidation of Intact Proteins by Isotopic Envelope and Mass Defect with Ion Cyclotron Resonance Fourier Transform Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2006**, *20*, 3535-3541.
34. Robinson, N.E., Lampi, K.J., Mciver, R.T., Williams, R.H., Muster, W.C., Kruppa, G., and Robinson, A.B. Quantitative Measurement of Deamidation in Lens Beta B2-Crystallin and Peptides by Direct Electrospray Injection and Fragmentation in a Fourier Transform Mass Spectrometer. *Molecular Vision*. **2005**, *11*, 1211-1219.
35. Zabrouskov, V., Han, X., Welker, E., Zhai, H., Lin, C., Van Wijk, K.J., Scheraga, H.A., and McLafferty, F.W. Stepwise Deamidation of Ribonuclease a at Five Sites Determined by Top Down Mass Spectrometry†. *Biochemistry*. **2005**, *45*, 987-992.
36. Noah E. Robinson, M.L.R., Stephanie E. S. Schulze, Bert T. Lai, and Harry B. Gray. Deamidation of α -Synuclein. *Protein Science*. **2010**, *18*, 1766-1773.
37. Watson, J.T., *Introduction to Mass Spectrometry*. Raven Press: New York, **1985**.
38. Chapman, J.R., *Practical Organic Mass Spectrometry*. 2nd ed.; Wiley: West Sussex, England, **1997**.
39. Tomer, K.B. The Development of Fast Atom Bombardment Combined with Tandem Mass Spectrometry for the Determination of Biomolecules. *Mass Spectrometry Reviews*. **1989**, *8*, 445-482.
40. Karas, M.I., Bachmann, D., Bahr, U., and Hillenkamp, F. Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. *International Journal of Mass Spectrometry and Ion Processes*. **1987**, *78*, 53-68.
41. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*. **1989**, *246*, 64-71.

42. Strupat, K., Karas, M., and Hillenkamp, F. 2,5-Dihydroxybenzoic Acid: A New Matrix for Laser Desorption-Ionization Mass Spectrometry. *International Journal of Mass Spectrometry and Ion Processes*. **1991**, *111*, 89-102.
43. Beavis, R.C. and Chait, B.T. Cinnamic Acid Derivatives as Matrices for Ultraviolet Laser Desorption Mass Spectrometry of Proteins. *Rapid Commun Mass Spectrom*. **1989**, *3*, 432-435.
44. Zenobi, R. and Knochenmuss, R. Ion Formation in Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry [Review]. *Mass Spectrometry Reviews*. **1998**, *17*, 337-366.
45. Beavis, R.C., Chait, B.T., and Standing, K.G. Matrix-Assisted Laser-Desorption Mass Spectrometry Using 355 Nm Radiation. *Rapid Communications in Mass Spectrometry*. **1989**, *3*, 436-439.
46. Marvin, L.F., Roberts, M.A., and Fay, L.B. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry in Clinical Chemistry. *Clinica Chimica Acta*. **2003**, *337*, 11-21.
47. Laremore, T.N., Murugesan, S., Park, T.J., Avci, F.Y., Zagorevski, D.V., and Linhardt, R.J. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometric Analysis of Uncomplexed Highly Sulfated Oligosaccharides Using Ionic Liquid Matrices. *Analytical Chemistry*. **2006**, *78*, 1774-1779.
48. Kebarle, P. A Brief Overview of the Present Status of the Mechanisms Involved in Electrospray Mass Spectrometry. *Journal of Mass Spectrometry*. **2000**, *35*, 804-817.
49. Iribarne, J.B. and Thomson, B.A. On the Evaporation of Small Ions from Charged Droplets. *Journal of Chemical Physics*. **1976**, *64*, 2287-2294.
50. Mack, L.L., Kralik, P., Rheude, A., and Dole, M. Molecular Beams of Macroions. Ii. *Journal of Chemical Physics*. **1970**, *52*, 4977-4986.
51. De Hoffmann, E.a.S.V., *Mass Spectrometry Principles and Applications*. John Wiley and Sons, Ltd.: Chichester, **2002**.
52. Comisaro, M.B. and Marshall, A.G. Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chemical Physics Letters*. **1974**, *25*, 282-283.
53. Comisaro, M.B. and Marshall, A.G. Frequency-Sweep Fourier-Transform Ion-Cyclotron Resonance Spectroscopy. *Chemical Physics Letters*. **1974**, *26*, 489-490.
54. Marshall, A.G. and Comisarow, M.B. Fourier and Hadamard Transform Methods in Spectroscopy. *Analytical Chemistry*. **1975**, *47*, 491A.
55. Caravatti, P. and Allemann, M. The 'Infinity Cell': A New Trapped-Ion Cell with Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Organic Mass Spectrometry*. **1991**, *26*, 514-518.
56. Amster, I.J. Fourier Transform Mass Spectrometry. *Journal of Mass Spectrometry*. **1996**, *31*, 1325-1337.
57. Jennings, K.R. Metastable Transitions In the Mass Spectrum of Benzene. *Journal of Chemical Physics*. **1965**, *43*, 4176-4177.
58. Jennings, K.R. Collision-Induced Decompositions of Aromatic Molecular Ions. *International Journal of Mass Spectrometry and Ion Physics*. **1968**, *1*, 227-235.
59. Haddon, W.F. and Mclafferty, F.W. Metastable Ion Characteristics. Vii. Collision-Induced Metastables. *Journal of the American Chemical Society*. **1968**, *90*, 4745-4746.

60. Tsaprailis, G., Somogyi, A., Nikolaev, E.N., and Wysocki, V.H. Refining the Model for Selective Cleavage at Acidic Residues in Arginine-Containing Protonated Peptides. *International Journal of Mass Spectrometry*. **2000**, *196*, 467-479.
61. Roepstorff, P. and Fohlman, J. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomedical Mass Spectrometry*. **1984**, *11*, 601.
62. Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., and Watson, J.T. Cid of Peptides. *Analytical Chemistry*. **1987**, *59*, 2621-2625.
63. Wysocki, V.H., Tsaprailis, G., Smith, L.L., and Brechi, L.A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *Journal of Mass Spectrometry*. **2000**, *35*, 1399-1406.
64. Summerfield, S.G., Whiting, A., and Gaskell, S.J. Intra-Ionic Interactions in Electrosprayed Peptide Ions. *International Journal of Mass Spectrometry and Ion Processes*. **1997**, *162*, 149-161.
65. Poulter, L. and Taylor, L.C.E. A Comparison of Low and High-Energy Collisionally Activated Decomposition Ms-Ms for Peptide Sequencing. *International Journal of Mass Spectrometry and Ion Processes*. **1989**, *91*, 183-197.
66. Nold, M.J., Wesdemiotis, C., Yalcin, T., and Harrison, A.G. Amide Bond Dissociation in Protonated Peptides. Structures of the N-Terminal Ionic and Neutral Fragments. *International Journal of Mass Spectrometry and Ion Processes*. **1997**, *164*, 137-153.
67. Harrison, A.G. and Yalcin, T. Proton Mobility in Protonated Amino Acids and Peptides. *International Journal of Mass Spectrometry*. **1997**, *165*, 339-347.
68. Vaisar, T. and Urban, J. Low-Energy Collision Induced Dissociation of Protonated Peptides. Importance of an Oxazolone Formation for a Peptide Bond Cleavage. *European Mass Spectrometry*. **1998**, *4*, 359-364.
69. Vaisar, T. and Urban, J. Gas-Phase Fragmentation of Protonated Mono-N-Methylated Peptides. Analogy with Solution-Phase Acid-Catalyzed Hydrolysis. *Journal of Mass Spectrometry*. **1998**, *33*, 505-524.
70. Summerfield, S.G. and Gaskell, S.J. Fragmentation Efficiencies of Peptide Ions Following Low Energy Collisional Activation. *International Journal of Mass Spectrometry*. **1997**, *165*, 509-521.
71. Harrison, A.G. and Tu, Y.P. Ion Chemistry of Protonated Aspartic Acid Derivatives. *Journal of Mass Spectrometry*. **1998**, *33*, 532-542.
72. Paizs, B., Lendvay, G., Vekey, K., and Suhai, S. Formation of B₂⁺ Ions from Protonated Peptides: An Ab Initio Study. *Rapid Communications in Mass Spectrometry*. **1999**, *13*, 525-533.
73. O'hair, R.a.J. Special Feature: Commentary - the Role of Nucleophile-Electrophile Interactions in the Unimolecular and Bimolecular Gas-Phase Ion Chemistry of Peptides and Related Systems. *Journal of Mass Spectrometry*. **2000**, *35*, 1377-1381.
74. Sobczyk, M., Anusiewicz, W., Berdys-Kochanska, J., Sawicka, A., Skurski, P., and Simons, J. Coulomb-Assisted Dissociative Electron Attachment: Application to a Model Peptide. *Journal of Physical Chemistry A*. **2005**, *109*, 250-258.

75. Anusiewicz, W., Berdys-Kochanska, J., and Simons, J. Electron Attachment Step in Electron Capture Dissociation (ECD) and Electron Transfer Dissociation (Etd). *Journal of Physical Chemistry A*. **2005**, *109*, 5801-5813.
76. Anusiewicz, I., Berdys-Kochanska, J., Skurski, P., and Simons, J. Simulating Electron Transfer Attachment to a Positively Charged Model Peptide. *Journal of Physical Chemistry A*. **2006**, *110*, 1261-1266.
77. Sobczyk, M. and Simons, J. The Role of Excited Rydberg States in Electron Transfer Dissociation. *Journal of Physical Chemistry B*. **2006**, *110*, 7519-7527.
78. Syrstad, E.A. and Turecek, F. Toward a General Mechanism of Electron Capture Dissociation. *Journal of the American Society for Mass Spectrometry*. **2005**, *16*, 208-224.
79. Sawicka, A., Skurski, P., Hudgins, R.R., and Simons, J. Model Calculations Relevant to Disulfide Bond Cleavage Via Electron Capture Influenced by Positively Charged Groups. *Journal of Physical Chemistry B*. **2003**, *107*, 13505-13511.
80. Leymarie, N., Berg, E.A., McComb, M.E., O'connor, P.B., Grogan, J., Oppenheim, F.G., and Costello, C.E. Tandem Mass Spectrometry for Structural Characterization of Proline-Rich Proteins: Application to Salivary Prp-3. *Analytical Chemistry*. **2002**, *74*, 4124-4132.
81. Zubarev, R.A., Zubarev, A.R., and Savitski, M.M. Electron Capture/Transfer Versus Collisionally Activated/Induced Dissociations: Solo or Duet? *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 753-761.
82. Mirgorodskaya, E., Roepstorff, P., and Zubarev, R.A. Localization of O-Glycosylation Sites in Peptides by Electron Capture Dissociation in a Fourier Transform Mass Spectrometer. *Analytical Chemistry*. **1999**, *71*, 4431-4436.
83. Kelleher, N.L., Zubarev, R.A., Bush, K., Furie, B., Furie, B.C., Mclafferty, F.W., and Walsh, C.T. Localization of Labile Posttranslational Modifications by Electron Capture Dissociation: The Case of Gamma-Carboxyglutamic Acid. *Analytical Chemistry*. **1999**, *71*, 4250-3.
84. Stensballe, A., Jensen, O.N., Olsen, J.V., Haselmann, K.F., and Zubarev, R.A. Electron Capture Dissociation of Singly and Multiply Phosphorylated Peptides. *Rapid Communications in Mass Spectrometry*. **2000**, *14*, 1793-1800.
85. Hakansson, K., Cooper, H.J., Emmett, M.R., Costello, C.E., Marshall, A.G., and Nilsson, C.L. Electron Capture Dissociation and Infrared Multiphoton Dissociation Ms/Ms of an N-Glycosylated Tryptic Peptide to Yield Complementary Sequence Information. *Analytical Chemistry*. **2001**, *73*, 4530-4536.
86. Shi, S.D.H., Hemling, M.E., Carr, S.A., Horn, D.M., Lindh, I., and Mclafferty, F.W. Phosphopeptide/Phosphoprotein Mapping by Electron Capture Dissociation Mass Spectrometry. *Analytical Chemistry*. **2001**, *73*, 19-22.
87. Cournoyer, J.J., Pittman, J.L., Ivleva, V.B., Fallows, E., Waskell, L., Costello, C.E., and O'connor, P.B. Deamidation: Differentiation of Aspartyl from Isoaspartyl Products in Peptides by Electron Capture Dissociation. *Protein Science*. **2005**, *14*, 452-463.
88. Baba, T., Hashimoto, Y., Hasegawa, H., Hirabayashi, A., and Waki, I. Electron Capture Dissociation in a Radio Frequency Ion Trap. *Analytical Chemistry*. **2004**, *76*, 4263-6.

89. Silivra, O.A., Kjeldsen, F., Ivonin, I.A., and Zubarev, R.A. Electron Capture Dissociation of Polypeptides in a Three-Dimensional Quadrupole Ion Trap: Implementation and First Results. *Journal of the American Society for Mass Spectrometry*. **2005**, *16*, 22-27.
90. Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., and Hunt, D.F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*. **2004**, *101*, 9528-33.
91. Coon, J.J., Syka, J.E.P., Schwartz, J.C., Shabanowitz, J., and Hunt, D.F. Anion Dependence in the Partitioning between Proton and Electron Transfer in Ion/Ion Reactions. *International Journal of Mass Spectrometry*. **2004**, *236*, 33-42.
92. O'connor, P.B., Cournoyer, J.J., Pitteri, S.J., Chrisman, P.A., and Mcluckey, S.A. Differentiation of Aspartic and Isoaspartic Acids Using Electron Transfer Dissociation. *Journal of the American Society for Mass Spectrometry*. **2006**, *17*, 15-19.
93. Chan, W.Y.K., Chan, T.W.D., and O'connor, P.B. Electron Transfer Dissociation with Supplemental Activation to Differentiate Aspartic and Isoaspartic Residues in Doubly Charged Peptide Cations. *Journal of the American Society for Mass Spectrometry*. **2010**, *21*, 1012-1015.
94. Little, D.P., Speir, J.P., Senko, M.W., O'connor, P.B., and Mclafferty, F.W. Infrared Multiphoton Dissociation of Large Multiply-Charged Ions for Biomolecule Sequencing. *Analytical Chemistry*. **1994**, *66*, 2809-2815.
95. Medzihradzky, K.F. and Burlingame, A.L. The Advantages and Versatility of a High-Energy Collision-Induced Dissociation-Based Strategy for the Sequence and Structural Determination of Proteins. *Methods (Orlando)*. **1994**, *6*, 284-303.
96. Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., and Watson, J.T. Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass-Spectrometer Differentiation of Leucine and Isoleucine. *Analytical Chemistry*. **1987**, *59*, 2621-2625.
97. Aubagnac, J.L., El Amrani, B., Devienne, F.M., and Combarieu, R. Characterization of Leucine and Isoleucine by Use of Fab Ionization Method in Tandem Mass Spectrometry. *Organic Mass Spectrometry*. **1985**, *20*, 428-429.
98. Heerma, W. and Bathelt, E.R. Identification of Leucine, Isoleucine and Norleucine in the Positive-Ion Fab Spectra of Peptides. *Biomedical and Environmental Mass Spectrometry*. **1986**, *13*, 205-208.
99. Castet, S., Enjalbal, C., Fulcrand, P., Guichou, J.F., Martinez, J., and Aubagnac, J.L. Characterization of Aspartic Acid and Beta-Aspartic Acid in Peptides by Fast-Atom Bombardment Mass Spectrometry and Tandem Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **1996**, *10*, 1934-1938.
100. Lloyd, J.R., Cotter, M.L., Otori, D., and Doyle, D.L. Distinction of Alpha-Aspartyl and Beta-Aspartyl and Alpha-Glutamyl and Gamma-Glutamyl-Transferase Peptides by Fast Atom Bombardment Tandem Mass Spectrometry. *Biomedical and Environmental Mass Spectrometry*. **1988**, *15*, 399-402.
101. Kusmierz, J.J. and Abramson, F.P. Tracing ¹⁵n with Chemical Reaction Interface Mass Spectrometry: A Demonstration Using ¹⁵n-Labeled Glutamine

- and Asparagine Substrates in Cell Culture. *Biological Mass Spectrometry*. **1994**, *23*, 756-763.
102. Gonzalez, L.J., Shimizu, T., Satomi, Y., Betancourt, L., Besada, V., Padron, G., Orlando, R., Shirasawa, T., Shimonishi, Y., and Takao, T. Differentiating Alpha- and Beta-Aspartic Acids by Electrospray Ionization and Low-Energy Tandem Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2000**, *14*, 2092-2102.
 103. Luu, N.C., Robinson, S., Zhao, R., Mckean, R., and Ridge, D.P. Mass Spectrometric Differentiation of Alpha- and Beta-Aspartic Acid in a Pseudo-Tetrapeptide Thrombosis Inhibitor and Its Isomer. *European Journal of Mass Spectrometry*. **2004**, *10*, 279-287.
 104. Chelius, D., Rehder, D.S., and Bondarenko, P.V. Identification and Characterization of Deamidation Sites in the Conserved Regions of Human Immunoglobulin Gamma Antibodies. *Analytical Chemistry*. **2005**, *77*, 6004-6011.
 105. Biemann, K. and Martin, S.A. Mass Spectrometric Determination of the Aminoacid Sequence of Peptides and Proteins. *Mass Spectrometry Reviews*. **1987**, *6*, 1-75.
 106. Johnson, R.S., Martin, S.A., and Biemann, K. Collision-Induced Fragmentation of (M+H)⁺ Ions of Peptides-Side-Chain Specific Sequence Ions. *International Journal of Mass Spectrometry and Ion Processes*. **1988**, *86*, 137-154.
 107. Kassel, D.B. and Biemann, K. Differentiation of Hydroxyproline Isomers and Isobars in Peptides by Tandem Mass Spectrometry. *Analytical Chemistry*. **1990**, *62*, 1691-1695.
 108. Hunt, D.F., Yates, J.R., Shabanowitz, J., Winston, S., and Hauer, C.R. Protein Sequencing by Tandem Mass Spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*. **1986**, *83*, 6233-6237.
 109. Zubarev, R.A., Kelleher, N.L., and McLafferty, F.W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society*. **1998**, *120*, 3265-3266.
 110. Kjeldsen, F., Haselmann, K.F., Budnik, B.A., Jensen, F., and Zubarev, R.A. Dissociative Capture of Hot (3-13 eV) Electrons by Polypeptide Polycations: An Efficient Process Accompanied by Secondary Fragmentation. *Chemical Physics Letters*. **2002**, *356*, 201-206.
 111. Kjeldsen, F., Haselmann, K.F., Sorensen, E.S., and Zubarev, R.A. Distinguishing of Ile/Leu Amino Acid Residues in the Pp3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry*. **2003**, *75*, 1267-1274.
 112. Williams, J.P., Creese, A.J., Roper, D.R., Green, B.N., and Cooper, H.J. Hot Electron Capture Dissociation Distinguishes Leucine from Isoleucine in a Novel Hemoglobin Variant, Hb Askew, Beta 54(D5)Val -≫ Ile. *Journal of the American Society for Mass Spectrometry*. **2009**, *20*, 1707-1713.
 113. Roher, A.E., Lowenson, J.D., Clarke, S., Wolkow, C., Wang, R., Cotter, R.J., Reardon, I.M., Zurchernecky, H.A., Heinrikson, R.L., Ball, M.J., and Greenberg, B.D. Structural Alterations in the Peptide Backbone of Beta-Amyloid Core Protein May Account for Its Deposition and Stability in Alzheimers-Disease. *Journal of Biological Chemistry*. **1993**, *268*, 3072-3083.
 114. Capasso, S. Estimation of the Deamidation Rate Of Asparagine Side Chains. *Journal of Peptide Research*. **2000**, *55*, 224-229.

115. Capasso, S. and Di Cerbo, P. Kinetic and Thermodynamic Control of the Relative Yield of the Deamidation of Asparagine and Isomerization of Aspartic Acid Residues. *Journal of Peptide Research*. **2000**, *56*, 382-387.
116. Cournoyer, J.J., Lin, C., and O'connor, P.B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. *Analytical Chemistry*. **2006**, *78*, 1264-1271.
117. Sargaeva, N.P., Lin, C., and O'connor, P.B. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid Beta Peptides, Including a Beta 1-42, Using Electron-Ion Reactions. *Analytical Chemistry*. **2009**, *81*, 9778-9786.
118. Sargaeva, N.P., Lin, C., and O'connor, P.B. Unusual Fragmentation of B-Linked Peptides by Exd Tandem Mass Spectrometry. *Journal of the American Chemical Society For Mass Spectrometry*. **2010**.
119. Li, X.J., Lin, C., and O'connor, P.B. Glutamine Deamidation: Differentiation of Glutamic Acid and Gamma-Glutamic Acid in Peptides by Electron Capture Dissociation. *Anal Chem*. **2010**, *82*, 3606-3615.
120. Yang, H., Fung, Y.M.E., Zubarev, A.R., and Zubarev, R.A. The Use of ECD for Proteomics-Wide Identification and Quantification of Iso-Asp Residues. *Molecular and Cellular Proteomics*. **2009**, S13-S13.
121. Yang, H.Q., Fung, E.Y.M., Zubarev, A.R., and Zubarev, R.A. Toward Proteome-Scale Identification and Quantification of Isoaspartyl Residues in Biological Samples. *Journal of Proteome Research*. **2009**, *8*, 4615-4621.
122. Cooper, H.J., Hudgins, R.R., and Marshall, A.G. Electron Capture Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Cyclodepsipeptides, Branched Peptides, and Epsilon-Peptides. *International Journal of Mass Spectrometry*. **2004**, *234*, 23-35.
123. O'connor, P.B. and Costello, C.E. Internal Calibration on Adjacent Samples (Incas) with Fourier Transform Mass Spectrometry. *Analytical Chemistry*. **2000**, *72*, 5881-5885.
124. Zaia, J. Mass Spectrometry and Glycomics. *Omics- a Journal of Integrative Biology*. **2010**, *14*, 401-418.
125. Zaia, J. Mass Spectrometry of Oligosaccharides. *Mass Spectrometry Reviews*. **2004**, *23*, 161-227.
126. Cooper, H.J., Hakansson, K., and Marshall, A.G. The Role of Electron Capture Dissociation in Biomolecular Analysis. *Mass Spectrometry Reviews*. **2005**, *24*, 201-222.
127. Zubarev, R., in *Principles of Mass Spectrometry Applied to Biomolecules*, ed. J. Laskin, Lifshitz, C. John Wiley & Sons, Inc.: New Jersey, 2006, vol. 1, pp 475-518.
128. Macht, M., Asperger, A., and Deininger, S.O. Comparison of Laser-Induced Dissociation and High-Energy Collision-Induced Dissociation Using Matrix-Assisted Laser Desorption/Ionization Tandem Time-of-Flight (Maldi-Tof/Tof) for Peptide and Protein Identification. *Rapid Communications in Mass Spectrometry*. **2004**, *18*, 2093-2105.
129. Zhang, L.Y. and Reilly, J.P. Peptide Photodissociation with 157 Nm Light in a Commercial Tandem Time-of-Flight Mass Spectrometer. *Analytical Chemistry*. **2009**, *81*, 7829-7838.
130. Soltwisch, J. and Dreisewerd, K. Discrimination of Isobaric Leucine and Isoleucine Residues and Analysis of Post-Translational Modifications in Peptides by Maldi in-Source Decay Mass Spectrometry Combined with Collisional Cooling. *Analytical Chemistry*. **2010**, *82*, 5628-5635.

131. Strupat K, K.V., Bui H, Viner R, Stafford R, Horning S. *Ap Maldi Produced Ions Inspected with an Exactive Mass Spectrometer*. in *Proceedings of the 59th Annual Conference on Mass Spectrometry and Allied Topics*. 2011. Denver, Colorado, June 5-9.
132. Woodhead, J. Collagen: The Anatomy of a Protein. *Studies in Biology No. 117*. **1980**.
133. Okada, K. and Kawase, M. Mass-Spectral Differentiation of Alpha-Linkages and Gamma-Linkages in Glutamyl Oligopeptides and Its Application for Structure Elucidation of Naturally Occurring Peptides. *Chemical & Pharmaceutical Bulletin*. **1977**, 25, 1497-1508.
134. Papayannopoulos, I.A. and Biemann, K. Fast-Atom-Bombardment and Tandem Mass-Spectrometry of Synthetic Peptides and by-Products. *Peptide Research*. **1992**, 5, 83-90.
135. Carr, S.A., Hemling, M.E., Bean, M.F., and Roberts, G.D. Integration of Mass Spectrometry in Analytical Biotechnology. *Analytical Chemistry*. **1991**, 63, 2802-2824.
136. Schindler, P., Muller, D., Marki, W., Grossenbacher, H., and Richter, W.J. Characterization of a Beta-Asp33 Isoform of Recombinant Hirudin Sequence Variant 1 by Low-Energy Collision-Induced Dissociation. *Journal of Mass Spectrometry*. **1996**, 31, 967-974.
137. Lehmann, W.D., Schlosser, A., Erben, G., Pipkorn, R., Bossemeyer, D., and Kinzel, V. Analysis of Isoaspartate in Peptides by Electrospray Tandem Mass Spectrometry. *Protein Science*. **2000**, 9, 2260-2268.
138. Schlosser, A. and Lehmann, W.D. Special Feature: Commentary - Five-Membered Ring Formation in Unimolecular Reactions of Peptides: A Key Structural Element Controlling Low-Energy Collision-Induced Dissociation of Peptides. *Journal of Mass Spectrometry*. **2000**, 35, 1382-1390.
139. Cournoyer, J.J., Lin, C., Bowman, M.J., and O'connor, P.B. Quantitating the Relative Abundance of Isoaspartyl Residues in Deamidated Proteins by Electron Capture Dissociation. *Journal of the American Society for Mass Spectrometry*. **2007**, 18, 48-56.
140. Sargaeva, N.P., Lin, C., and O'connor, P.B. Differentiating N-Terminal Aspartic and Isoaspartic Acid Residues in Peptides. *Analytical Chemistry*. **2011**, 6675-6682.
141. O'connor, P.B., Cournoyer, J.J., Pitteri, S.J., Chrisman, P.A., and Mcluckey, S.A. Differentiation of Aspartic and Isoaspartic Acids Using Electron Transfer Dissociation. *Journal of the American Society for Mass Spectrometry*. **2005**, 17, 15-19.
142. Potter, S.M., Henzel, W.J., and Aswad, D.W. In-Vitro Aging of Calmodulin Generates Isoaspartate at Multiple Asn-Gly and Asp-Gly Sites in Calcium-Binding Domain-Ii, Domain-Iii, and Domain-Iv. *Protein Science*. **1993**, 2, 1648-1663.
143. Wang, B., Shang, J., Qin, Y., Yan, B., and Guo, X. Differentiation of A- or B-Aspartic Isomers in the Heptapeptides by the Fragments of $[M + Na]^+$ Using Ion Trap Tandem Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **2011**, 22, 1453-1462.
144. Andreazza, H.J., Wang, T., Bagley, C.J., Hoffmann, P., and Bowie, J.H. Negative Ion Fragmentations of Deprotonated Peptides. The Unusual Case of Isoasp: A Joint Experimental and Theoretical Study. Comparison with

- Positive Ion Cleavages. *Rapid Communications in Mass Spectrometry*. **2009**, 23, 1993-2002.
145. Wright, H.T. Nonenzymatic Deamidation of Asparaginyl and Glutaminyl Residues in Proteins. *Critical Reviews in Biochemistry and Molecular Biology*. **1991**, 26, 1-52.
 146. Capasso, S., Mazzarella, L., Sica, F., and Zagari, A. 1st Evidence of Spontaneous Deamidation of Glutamine Residue Via Cyclic Imide to Alpha-Glutamic and Gamma-Glutamic Residue under Physiological Conditions. *Journal of the Chemical Society-Chemical Communications*. **1991**, 1667-1668.
 147. Nagasawa, H.T., Magnan, S.D.J., and Foltz, R.L. Differentiation of Alpha-Glutamyl from Gamma-Glutamyl-Transferase Dipeptides by Chemical Ionization Mass-Spectrometry. *Biomedical Mass Spectrometry*. **1982**, 9, 252-256.

Chapter 2: Development of a methodology to monitor deamidation of collagen I

¹This chapter has been partially/entirely reproduced from Hurtado, P. P., O'Connor, P. B., Deamidation of Collagen. *Analytical Chemistry*. **2012**, 84. 3017-3025.

2.1. INTRODUCTION

For the last three decades, extensive research into protein analysis has generated new tools that allow archaeologists to predict the age of ancient materials. So far, the most popular methodologies to achieve these are based on chronometric techniques such as radiocarbon dating, and amino acid dating (using Asp racemisation).^{1, 2} Radiocarbon dating has significant upper and lower limits. Due to the logarithmic decay rate of ^{14}C , the practical upper limit is about 50,000 years. This is because so little ^{14}C remains after almost 9 half-lives that it may be hard to detect and obtain an accurate reading, regardless of the size of the sample.³ The latter technique (amino acid racemisation) allows the study of a wider range of ages, up to 2 million years old, but it relies on the ability of certain amino acids such as Asp, Leu, and Glu to racemise.⁴⁻⁷

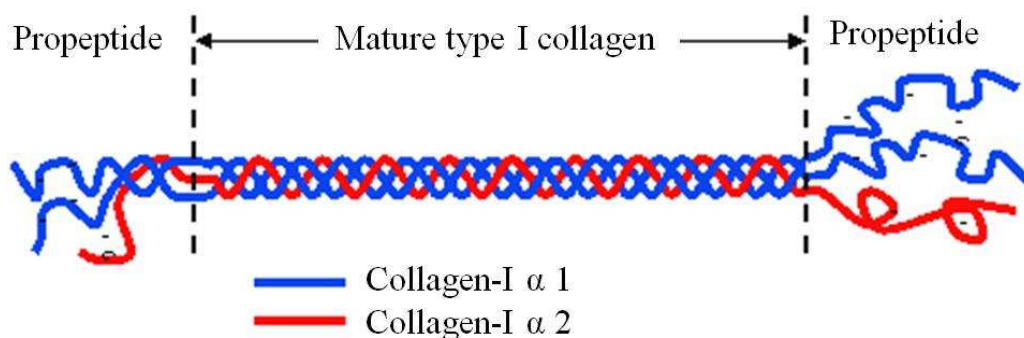


FIGURE 2.1: Representation of the structure of collagen type I. Adapted from <http://atlasgeneticsoncology.org/Genes/COL1A2ID411ch7q22.html>. Accessed on: May 17th 2013.

Collagen represents an ideal molecule for analysis in dating materials. Collagen is the major component of skin, tendons, ligaments, teeth and bones providing the framework that holds most multi-cellular animals together. Type I constitutes the major fibrillar collagen of bone. This protein is found, in at least small quantities, in most animal phyla and it is located in the connective tissues providing strength and conferring form while allowing flexibility.⁸ Due to its triple helix structure (see Fig.

2.1) and the solid components of the bone matrix, collagen persists with extraordinary longevity under extreme conditions,^{9, 10} but unfortunately the same strong structure that characterizes collagen also minimises Asp racemisation,¹¹⁻¹³ thus limiting its utility for current dating tool analyses. Recent reports have shown that certain protein modifications such as deamidation are remarkably increased in collagen from old bone in mammals.¹⁴ This observation could potentially contribute towards the possibility of using this diagenetic modification as a means of assessing the age of archaeological materials.

The relationship between aging and deamidation is a topic of continuous debate. One of the most relevant pieces of experimental evidence about deamidation and aging relates to *in vivo* accumulation of deamidated proteins of different kinds with age.¹⁵ It was hypothesized¹⁶ and then experimentally demonstrated¹⁷⁻¹⁹ that deamidation can serve as a biologically relevant molecular clock that regulates the timing of *in vivo* processes. Two of the pioneering groups that have extensively reported deamidation studies in proteins and peptides are the groups of Arthur Robinson and Noah Robinson. A significant library of sequence-determined Asn deamidation rates has been published by these groups.^{20, 21} Also a comprehensive book related to deamidation has been published by the same groups.²²

Smith and collaborators have significantly contributed towards deamidation research by their studies in PTMs in human lens crystallins using FAB.²³⁻²⁷ The correlation of developing cataracts with old age constituted the starting point in the crystalline research carried out by Smith's group. Crystallins are highly soluble and they constitute about 90 % of lens proteins. The accumulation of many PTMs in

crystallins has been demonstrated.²⁸⁻³⁰ Amongst these, deamidation of Asn and Gln is one of the major modifications,^{27, 31, 32} which affect lens transparency, decreases crystalline solubility, and is related to cataract formation.³³ Smith and collaborators have reported that the amount of deamidated proteins was increased with age in a range of samples from 20 to 27 years old.²³ In their experiments, deamidation half-life was estimated to be 41,000 days in the >30 kD fraction and 56,000 days in the 13 to 30 kD fraction.³⁴ This was screened in 50-65 years old subjects. Their findings also shown that a difference in deamidation as a function of molecular weight or aggregation was observed in a range of 6 Gln and 2 Asn deamidation in α A, α B, α B₁, and γ S-crystallins. Findings from Smith and collaborators support the basis of the potential use of Asn deamidation as a way of assessing the age of materials.

As previously described in Chapter 1, Asn deamidation results in a +0.984 Da mass shift and a charge shift making it relatively straightforward to detect.^{20, 35, 36} Several mass spectrometric methodologies have been utilized to monitor the shift of mass caused by peptide deamidation; amongst these are LC-MS^{37, 38} and MALDI-TOF-MS.³⁹⁻⁴² Asp isomerisation can be challenging to detect using mass spectrometry.⁴³⁻⁴⁵ Early methods were developed to differentiate the two isomeric forms^{46, 47} based on relative fragment ion abundance, but the introduction of new fragmentation techniques in mass spectrometry such as electron capture (ECD) and electron transfer dissociation (ETD) has improved the differentiation of aspartyl from isoaspartyl residues in protein modifications.^{43-45, 48}

Considering that deamidation follows first-order reaction kinetics, several methodologies can be applied to determine the half-life of the deamidating Asn.

One way is making use of ECD along with high resolution mass spectrometry, in which the peak intensity ratio of IsoAsp/Asp is monitored at different points of the reaction.^{44, 45, 49, 50} The drawback of this method is that it relies on the appearance of the diagnostic fragment ions for these two isomers. Other methods, such as isotopic deconvolution and the mass defect method overcome this issue but cannot distinguish the isomers of IsoAsp.

The other methods to measure the extent of deamidation (isotopic deconvolution and mass defect) have been previously discussed in Chapter 1. The present work aims to show a methodology to measure the extent of deamidation using the mass defect method in tryptic peptides of collagen I. The differentiation of Asp from IsoAsp using ECD was also explored as a potential alternative methodology, but was found so far to be unsuitable for kinetic measurements. The study also attempts to measure deamidation rates and half-lives to assess the feasibility of using Asn deamidation as an age dating method for collagen.

2.2. EXPERIMENTAL METHODOLOGY

2.2.1. Materials

Collagen type I, trypsin sequencing-grade, ammonium acetate ($\text{CH}_3\text{COONH}_4$), and ammonium bicarbonate (NH_4HCO_3) were purchased from Sigma Aldrich, Gillingham, UK. HPLC grade methanol and formic acid (HCOOH) were obtained from Fisher Scientific, Leicestershire, UK.

2.2.2. Digestion of collagen

Prior to digestion, 3 mg of collagen type I (triplicate) was weighed out and denatured in 0.5 mL of 50 mM ammonium bicarbonate (pH 7.5) at 62 °C, which was found to be the optimum temperature to get collagen into solution, after several trials using temperatures ranging from 40 to 75°C. This was repeated varying the denaturation time (40, 60, 80, 100, and 120 minutes). The samples were centrifuged at 11,298 g for 20 min at 21°C. The supernatant (350 µL) was collected and 12 µL of trypsin solution (2 µg/µL) was added using a ratio of 1:40 enzyme:protein. To minimise deamidation during tryptic digestion, the mixture was allowed to react for 4 hours at 37 °C,⁵¹ the reaction was halted by adding 5% formic acid, and the samples were dried in a SpeedVac system (Savant SPD121P Fisher Scientific, Leicestershire, UK) to evaporate volatiles at 35 °C, they were then stored at -80°C until analyses were carried out (\approx 1 day). The samples were reconstituted into a mixture of 1:1 MeOH:H₂O, and diluted into a mixture of 49.5:49.5:1 MeOH:H₂O:formic acid to a final solution of an approximate concentration of 2.5 µM, which is based on maximal quantity from collagen starting material.

2.2.3. FT-ICR-MS measurements

The experiments were carried out on a Bruker solarix 12 T FT-ICR-MS. In order to identify deamidation sites, samples of digested collagen were electrosprayed at approximately 2.5 µM concentration in 49.5:49.5:1 MeOH:H₂O:formic acid. For CAD experiments,⁵² the parent ions were isolated using the quadrupole, and were fragmented in the collision cell (q, typical collision energies ~10-30 eV), and then transmitted into the ion cyclotron resonance cell (ICR) for detection, see Fig. 1.3. For ECD experiments, the parent ions were isolated in the quadrupole and externally

accumulated in a hexapole collision cell (at injection energies <5 eV) for 1-10 s. After being transferred to and trapped in the ICR cell, ions were irradiated with 1.5 eV electrons from a heated hollow cathode (1.7 A) for 20 to 90 ms.⁵³ The mass spectrometer was externally calibrated using tuning mixture (HP mix). After performing CAD and ECD to characterize 30 tryptic peptides, the instrument was then internally calibrated with tryptic peptides from collagen I that have been previously identified by manually interpretation of their CAD and ECD spectra. The ECD spectra for the three peptides chosen to monitor deamidation were also internally calibrated using c and z• ions corresponding to the known peptides. The assigned mass tables for each spectrum are provided in section 2.6 of this chapter.

2.3. RESULTS AND DISCUSSION

2.3.1. Peptide Characterization

Collagen (I) is a heterotrimer composed of two $\alpha 1$ and one $\alpha 2$ chains and contains an uninterrupted triple helical domain (~1000 residues from each chain) of approximately 400 kDa molecular weight. The triple helix is stabilized by a high fractional content of the amino acids Pro ~12% and Hyp ~10%, by the close packing of the three chains, by hydrogen bonding between the chains, hydroxylysine crosslinks, and by an extensive hydration network. The folding of the three chains requires every third residue to be a Gly, generating a repeating (Gly-X-Y)_n sequence pattern, where X and Y are any other amino acid residue (see Fig. 2.2.).⁵⁴

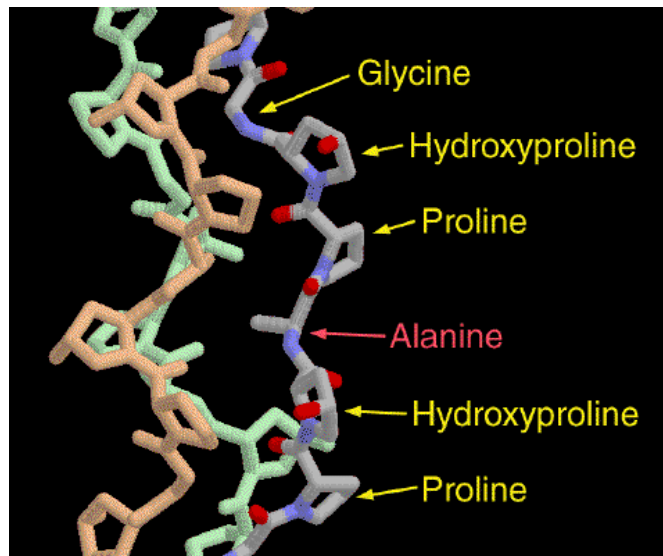


FIGURE 2.2: Sequence structure of collagen type 1. Figure adjusted from protein data bank: <http://www.rcsb.org/pdb/101/motm.do?momID=4>. Accessed on: May 17th 2013.

The most frequent tripeptide unit (triplet) is Gly-Pro-Hyp, which also contributes the maximal stability to the triple helix.⁵⁵ In general, Pro residues in collagen appear in positions 2 and 3 of the triplet, Hyp is confined exclusively to position 3 and very few Pro in that position are not hydroxylated (see table 2.1 for sequences and mass accuracy information.). This clearly suggests a steric effect in the enzymatic modification of prolines, thus increasing the challenge of locating the modification in a single peptide from collagen.⁵⁶ In order to locate and differentiate the Hyp residues (113.04768 Da) from Leu/Ile (113.08406 Da) in the peptides, high resolving power was necessary.

Thermal denaturation of collagen introduces a “statistical random coil” state when the molecule is in solution.⁸ The breakage of the three helices and subsequent tryptic digestion produces a significant number of peptides with similar mass, as shown in table 2.1 and the spectrum in Fig. 2.3. In order to characterize the peptides in collagen, the use of ECD and CAD was necessary; this is due to the high content of Pro and Hyp in these types of peptides. It is well known that ECD does not cleave the N-terminal to Pro,⁵⁷ thus the complementary information provided by CAD was essential in this case. The high resolving power of FT-ICR-MS along with its ability to fragment peptides by CAD and ECD, facilitated the reliable identification of 30 peptides using these fragmentation techniques, from which three were chosen for deamidation studies, because they were the only peptides that showed a shift of 0.984 Da in the spectrum as highlighted in red boxes in Fig 2.3.

TABLE 2.1: Tryptic peptides of Collagen I. Confirmed using CAD

Length coll-I(chain)	Identified peptides	MW	[M+H] ¹⁺	[M+2H] ²⁺	Exp.	Error ppm
1084-1092 (α 1)	GPAGPQGPR	835.43002	836.43730	418.72229	418.72251	-0.53
268-276 (α 1)	GFSGLDGAK	850.41845	851.42573	426.21651	426.21651	-0.01
994-1003 (α 2)	GPSGPQGIR	867.45624	868.46352	434.73540	434.73529	0.25
361-369 (α 1)	GSEGPQGVV	885.43042	886.43770	443.72249	443.72219	0.68
1052-1062 (α 2)	GPAGPSGPAGK	894.45589	895.46317	448.23523	448.23495	0.61
<u>958-966 (α1)</u>	<u>GVVGLPGQR</u>	<u>897.50319</u>	<u>898.51047</u>	<u>449.75888</u>	<u>449.75869</u>	<u>0.41</u>
412-421 (α 2)	AGVMGPAGSR	901.44396	902.45124	451.72926	451.72913	0.29
697-708 (α 1)	GANGA <u>P</u> GNDGAK	1045.43119	1046.43847	523.72288	523.72337	-0.95
<u>493-504 (α1)</u>	<u>GFP</u> GADGVAGPK	<u>1087.52978</u>	<u>1088.53706</u>	<u>544.77217</u>	<u>544.77229</u>	<u>-0.22</u>
1066-1078 (α 2)	IGQP <u>G</u> AVGPAGIR ¹	1191.67236	1192.67964	596.84346	596.84333	0.22
485-496 (α 2)	GE <u>P</u> GNIGF <u>P</u> GPK	1200.57747	1201.58475	601.29602	601.29594	0.12
1066-1078 (α 2)	IGQ <u>P</u> GAVGPAGIR ¹	1207.66728	1208.67456	604.84092	604.84092	0.00
538-551 (α 1)	GLTG <u>S</u> PG <u>S</u> PGPDGK	1257.58368	1258.59096	629.79912	629.79896	0.25
449-462 (α 2)	G <u>F</u> PG <u>S</u> PGNIGPAGK	1286.62548	1287.63276	644.32002	644.31970	0.50
970-983 (α 1)	G <u>F</u> P <u>G</u> L <u>P</u> GPSGEPGK	1327.64079	1328.64807	664.82768	664.82761	0.10
<u>415-429 (α1)</u>	<u>GPS</u> GPQGPS <u>G</u> PPGPK	<u>1331.64694</u>	<u>1332.65422</u>	<u>666.83075</u>	<u>666.83077</u>	<u>-0.03</u>
970-983 (α 1)	G <u>F</u> P <u>G</u> L <u>P</u> GPSG <u>E</u> PGK	1343.63571	1344.64299	672.82514	672.82520	-0.10
415-429 (α 1)	GPSGPQGPS <u>G</u> PP <u>G</u> PK	1347.64186	1348.64914	674.82821	674.82815	0.09
<u>472-486 (α1)</u>	<u>GEP</u> GPAGL <u>P</u> GPPGER	<u>1434.67388</u>	<u>1435.68116</u>	<u>718.34422</u>	<u>718.34419</u>	<u>0.04</u>
865-881 (α 1)	GSAG <u>P</u> PGATGF <u>P</u> GAAGR	1458.68511	1459.69239	730.34984	730.35006	-0.31
775-792 (α 2)	GDGG <u>P</u> PGATGF <u>P</u> GAAGR	1472.66437	1473.67165	737.33947	737.33971	-0.33
<u>977-994 (α2)</u>	<u>GEP</u> GPAGAVGPAGAVGPR ²	<u>1515.77932</u>	<u>1516.78660</u>	<u>758.89694</u>	<u>758.89666</u>	<u>0.37</u>

TABLE 2.1: Tryptic peptides of Collagen I. Confirmed using CAD (Continued)

Length coll-I(chain)	Identified peptides	MW	[M+H] ¹⁺	[M+2H] ²⁺	Exp.	Error ppm
977-994 (α 2)	GE <u>P</u> GPAGAVGPAGAVGPR ²	1531.77424	1532.78152	766.89440	766.89480	-0.52
674-691 (α 2)	GAPGAIGAP <u>P</u> GPAGANGDR	1537.71204	1538.71932	769.86330	769.86295	0.45
1066-1083 (α 1)	GETGPAGPAGPIGPVGAR	1559.80554	1560.81282	780.91005	780.90981	0.31
1152-1167 (α 1)	DGLNGL <u>P</u> GP <u>I</u> GP <u>P</u> PGPR	1560.78958	1561.79686	781.40207	781.40222	-0.19
1152-1167 (α 1)	DGLNGL <u>P</u> GP <u>I</u> GP <u>P</u> PGPR	1561.77359	1562.78087	781.89408	781.89371	0.47
397-414 (α 1)	GANGA <u>P</u> GIAGAP <u>P</u> GF <u>P</u> PGAR	1584.76442	1585.7717	793.38949	793.38943	0.08
397-414 (α 1)	GANGA <u>P</u> GIAGAP <u>P</u> GF <u>P</u> PGAR	1585.74843	1586.75571	793.88150	793.88087	0.79
612-630 (α 1)	DGEAGAQQ <u>P</u> PGPAGPAGER	1705.76553	1706.77281	853.89005	853.88995	0.11
1022-1051 (α 2)	GHNGLQGL <u>P</u> GLAGHHGDQGAP <u>P</u> GAVGPAGPR	2783.35994	2784.36722	696.84726(+4)	696.84696	0.43
1022-1051 (α 2)	GH <u>D</u> GLQGL <u>P</u> GLAGHHGDQGAP <u>P</u> GAVGPAGPR	2784.34395	2785.35123	697.09326(+4)	697.09295	0.44
713-736 (α 2)	GEVGPAGPNGFAG <u>P</u> AGAAGQ <u>P</u> GAK	2066.96567	2067.97295	1034.49012	1034.49083	-0.69

^{1,2}Same peptide (different modification)

P: Hydroxyproline (Hyp)

N: Deamidation of N (D)

Underlined peptides (+1 and +2) were used to calibrate de instrument

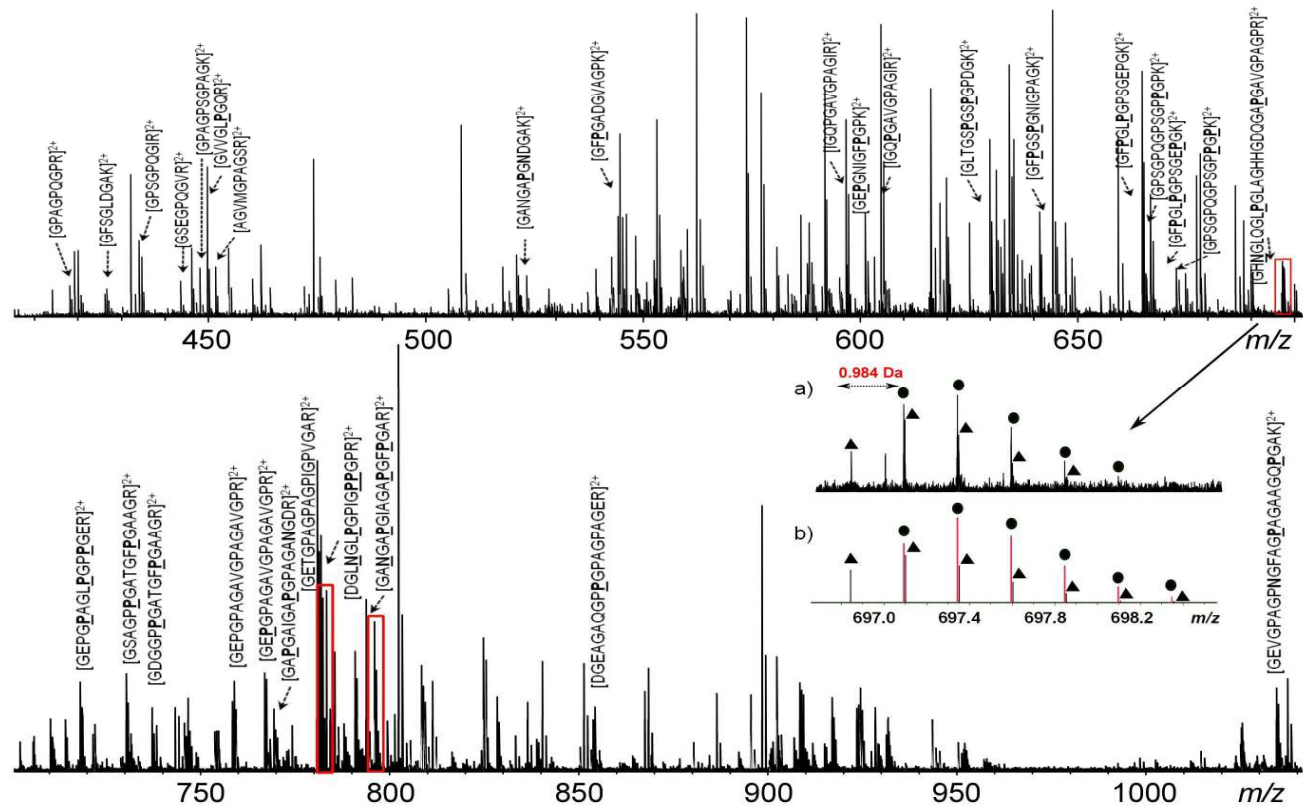


FIGURE 2.3: Mass spectrum of collagen tryptic digest. Peptides marked with N (Asn) were selected for deamidation studies. **P** represents Hyp. Insert a) shows the isotopic pattern for the non-deamidated, m/z [696.84696]⁴⁺, and the deamidated peptide, m/z [697.09326]⁴⁺, where the triangles represent the non-deamidated peptide and the circles represent the deamidated peptide. Insert b) shows the theoretical isotopic distribution for the peptides.

2.3.2. Identification of Deamidation Sites in Collagen by Using FT-ICR-MS

In order to assess deamidation, a shift of +0.984 Da was monitored in all peptides in the mass spectrum of the collagen tryptic digest (see red boxes and insert in Fig 2.3.). Analysis of the spectrum showed three of the six possible peptides that may undergo Asn deamidation due to the presence of Asn in the sequence. These three observed peptides corresponded well with the expected mass (<1 ppm). Therefore, they were chosen for further deamidation studies: [coll_I (α 2) 1022-1051]⁴⁺, [coll_I (α 1) 1152-1167]²⁺, and [coll_I (α 1) 397-414]²⁺.

Fig. 2.4 shows the ECD spectra corresponding to [coll_I (α 2) 1022-1051]⁴⁺. With ECD, 80% cleavage coverage can be achieved for the peptide shown in Fig 2.4 (excluding cleavages next to Pro, the fragmentation of this peptide shows 100% of the expected cleavages). After manual interpretation of the spectrum, the deamidation site was easily identified as position 3 from the N-terminus. Fragmentation with CAD for this peptide showed the location of two Hyp at positions 10 and 22 from the C-terminus. ECD spectrum for [coll_I (α 1) 1152-1167]²⁺ and coll_I (α 1) 397-414 are shown in Figs. 2.5 and 2.6 respectively; they show less fragmentation than the CAD spectra due to the number of prolines present in each peptide. However, the deamidation sites were easily located in each peptide. Figs. 2.7, 2.8 and 2.9 show the CAD spectra showing the positions of Hyp for each peptide. The data obtained from the CAD spectra complemented the data from ECD to complete the sequences. Isotopic distributions and high mass accuracy were used to increase the confidence in assignments of each peptide and its fragments. The complete list of masses is included at the end of the chapter.

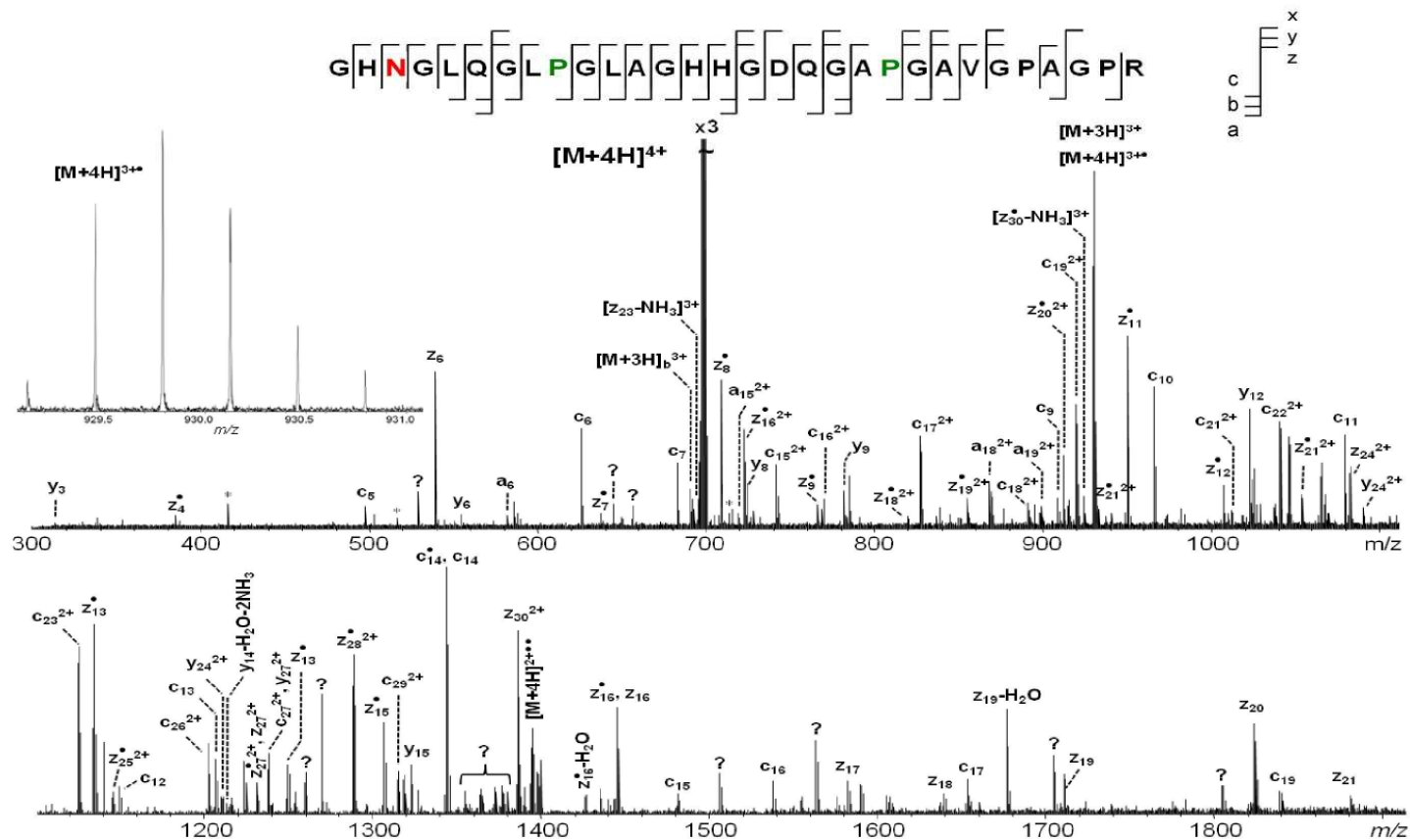


FIGURE 2.4: ECD of tryptic digest peptide [coll_I (α_2) -1022-1051]⁴⁺ ion at m/z 697.09326. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Insert shows the charge reduced species at m/z 929.4584. (*) represents noise peaks and (?) represents unknown peaks, (\blacktriangle) represents a different peptide present in the mixture.

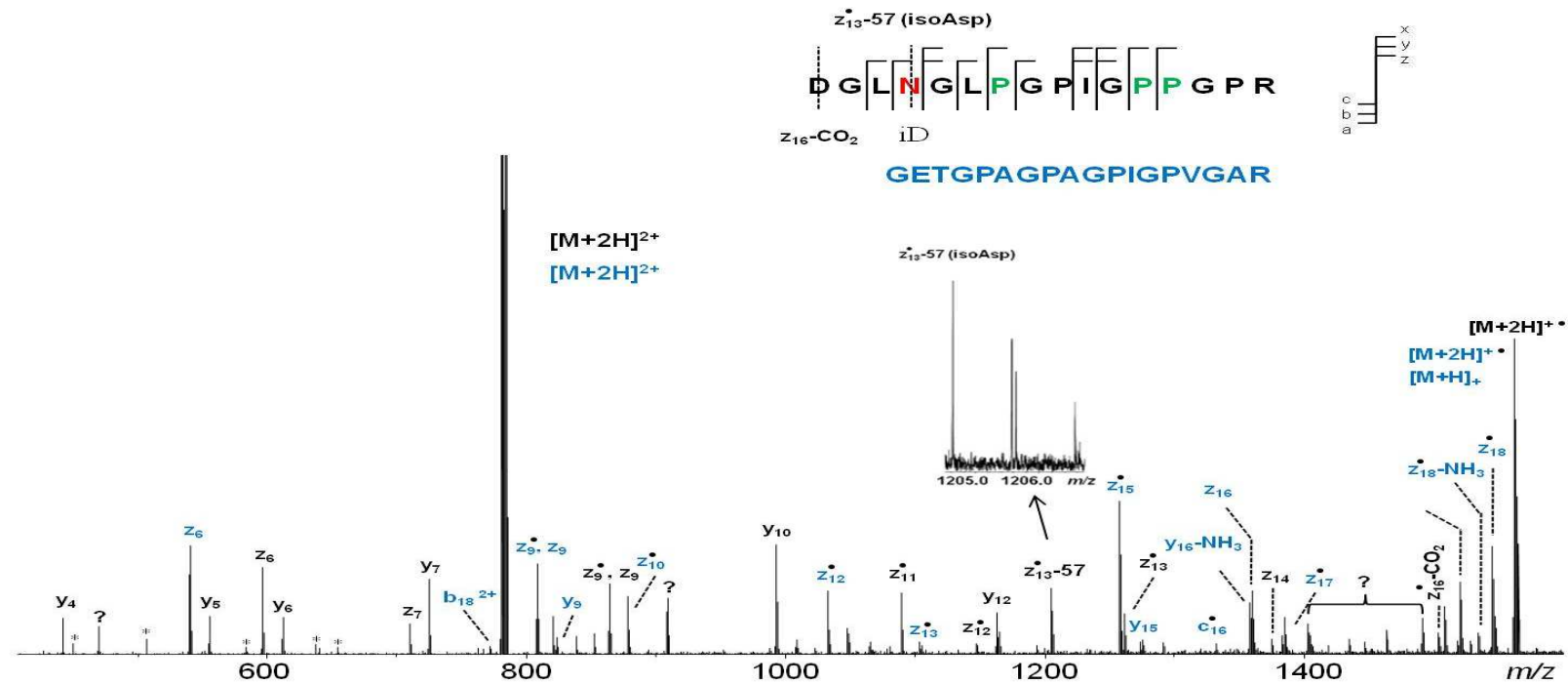


FIGURE 2.5: ECD of tryptic digest peptide [coll_I (α 1) 1152-1167] $^{2+}$ ion at m/z 781.89408. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Insert show the region where IsoAsp diagnostic ion appears ($z_{13}^{\bullet}-57$, m/z 1204.63171) indicating position 13, and the lost CO_2 from z_{16} ($z_{16}-44$, m/z 1502.77117) represents the diagnostic ion for the presence of Asp in position 16. Ions labelled with blue are originated from another peptide present in the digestion mixture at m/z 780.90998.

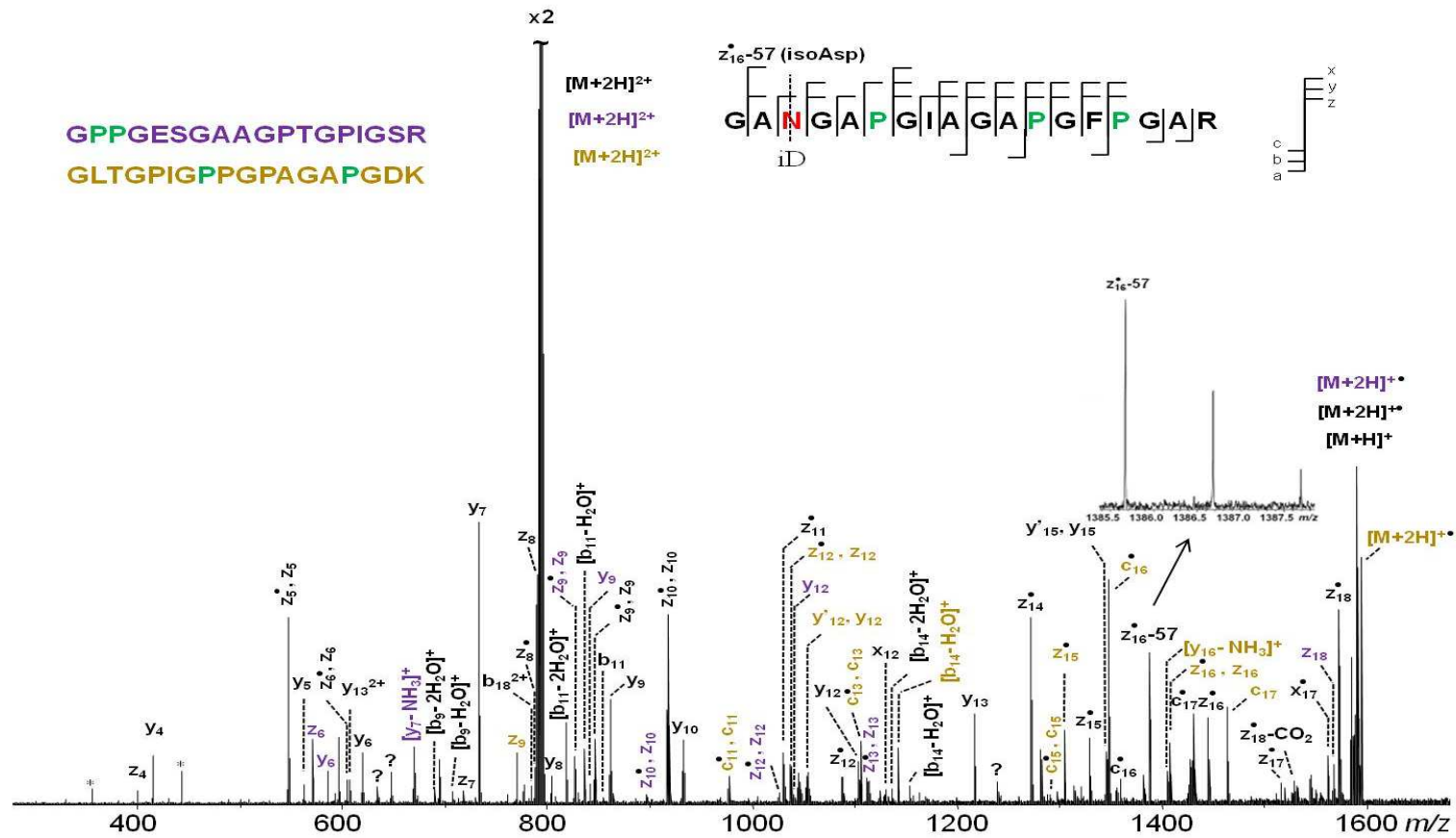


FIGURE 2.6: ECD of tryptic digest peptide [coll_I (α 1) 397-414] $^{2+}$ ion at 793.88150 m/z . Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. The insert show the region where the Asp diagnostic ion (z_{16}^{-57} , m/z 1385.68082) appears, this one correspond to position 3 (N-terminal), 16 (C-terminal) in the peptide. Ions labelled with purple and brown are originated from other two peptides present in the digestion mixture with m/z 790.88673 and m/z 795.90975.

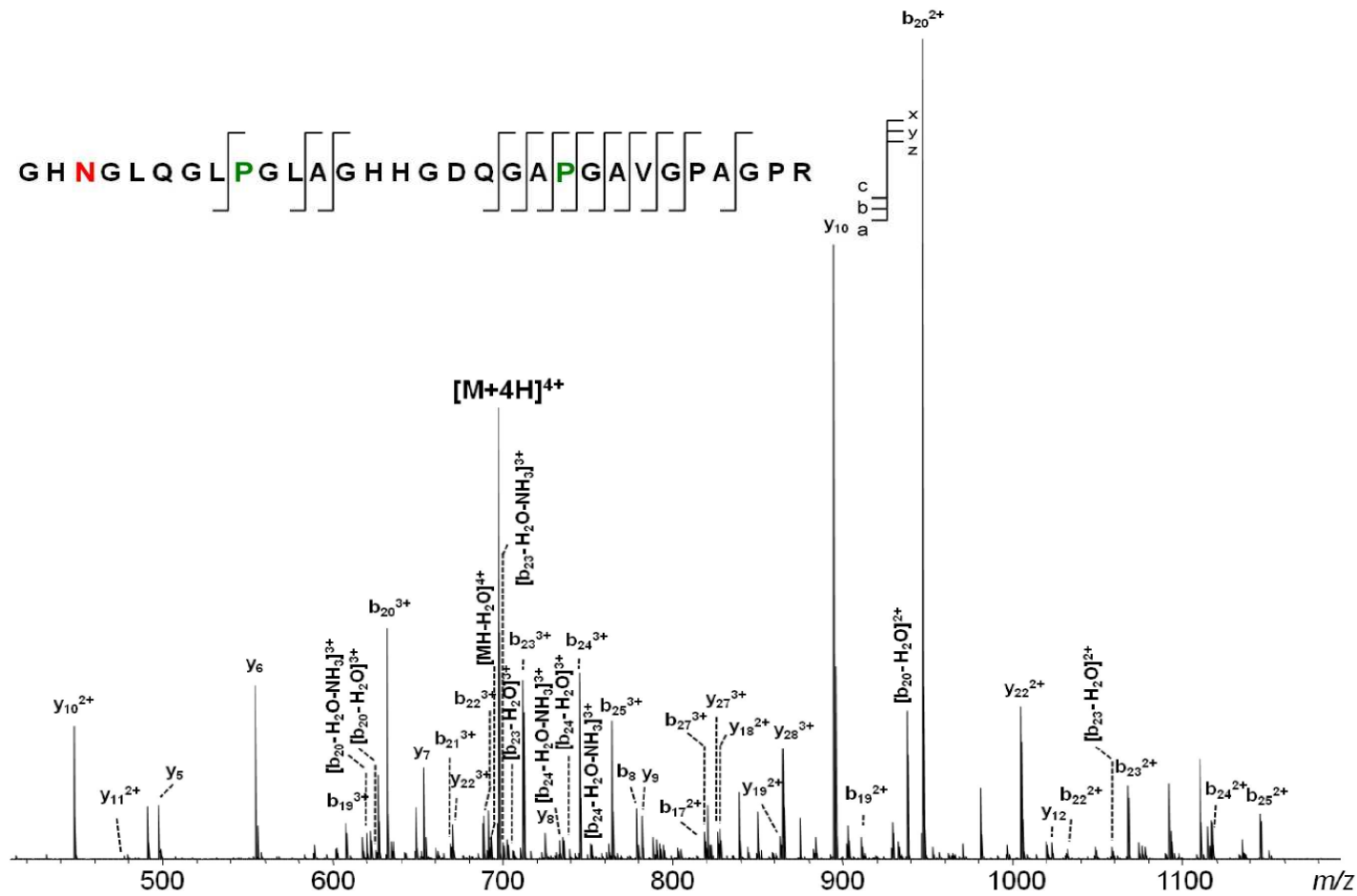


FIGURE 2.7: CAD of tryptic digest peptide [coll_I (α_2) -1022-1051] $^{4+}$ ion at m/z 697.09326. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site.

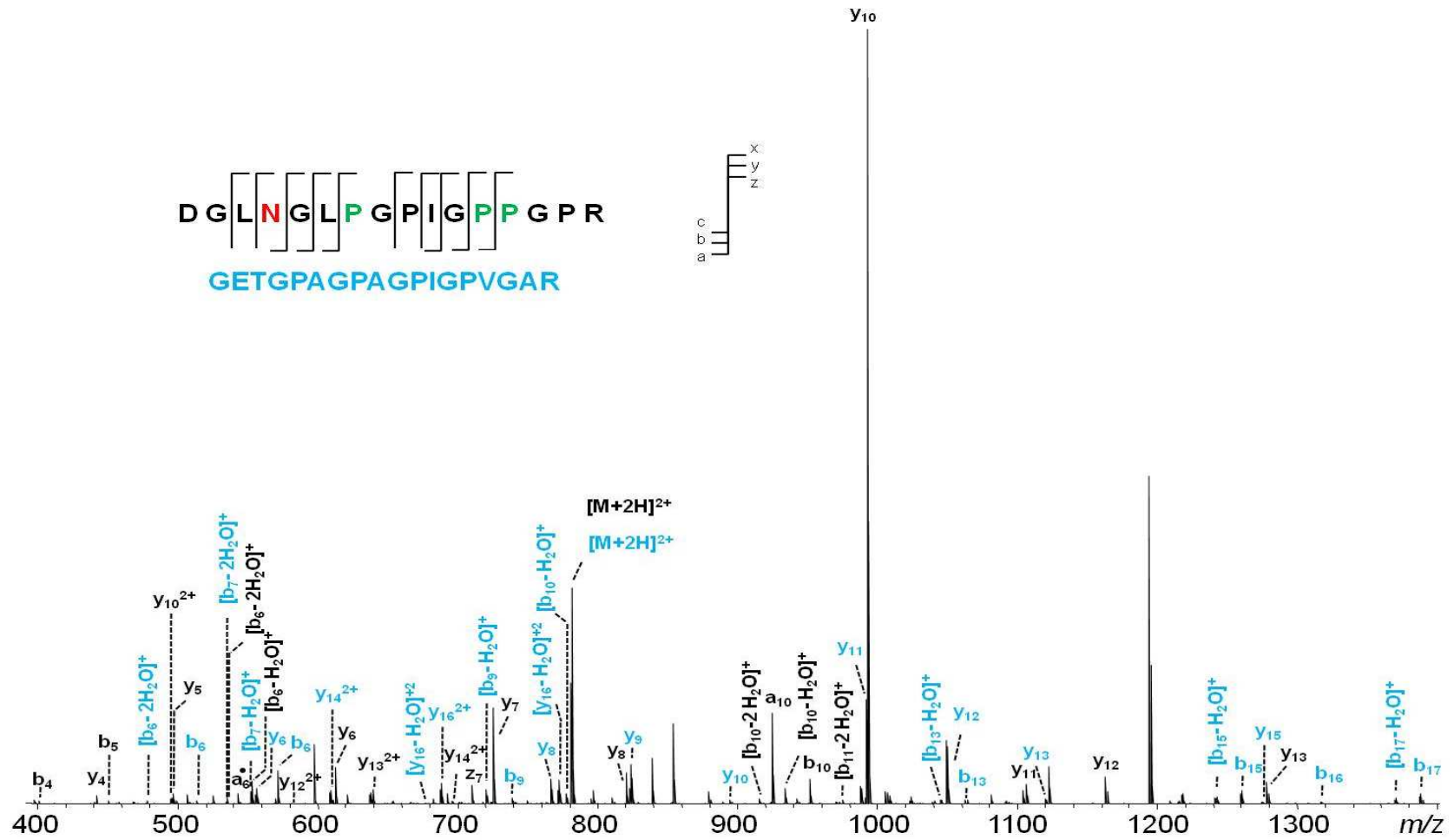


FIGURE 2.8: CAD of tryptic digest peptide [coll_I(α 1)-1152-1167]²⁺ ion at m/z 781.89408. Asn highlighted in red shows the deamidation site. P highlighted in green show hydroxylation site. Ions labelled with blue are originated from another peptide present in the digestion mixture at m/z 780.90998.

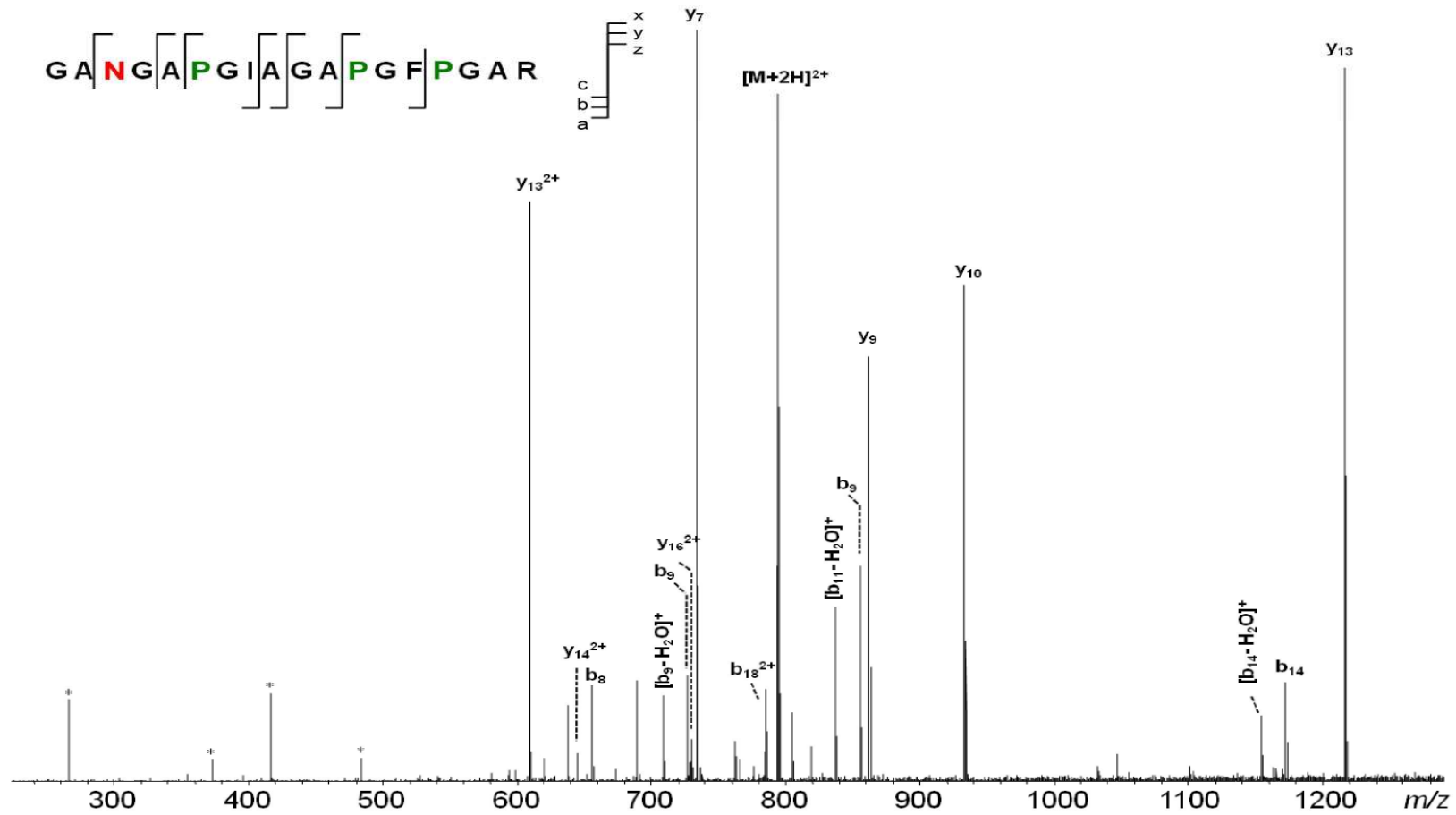


FIGURE 2.9: CAD of tryptic digest peptide [coll_I (α 1) 397-414, coll_I(α 1)]²⁺ ion at m/z 793.88150. Asn highlighted in red shows the deamidation site. Prolines highlighted in green show hydroxylation site.

2.3.3. Deamidation Rates

To study the deamidation of these peptides, collagen was denatured (heated) at different times (40, 60, 80, 100, and 120 minutes) holding the temperature at 62 °C. First order reaction was assumed, which, to date, is the most accepted mechanism for this particular post-translational modification.⁵⁸⁻⁶⁰ As previously described in Chapter 1, this assumption implies that the reaction occurs spontaneously under physiological conditions through a succinimide intermediate. The rate of the deamidation reaction is affected by both the amino acid sequence and the three-dimensional structure. In addition to the latter, other parameters play a significant role in deamidation rates such as pH, temperature, ionic strength, buffer ions, and other solution properties. For temperatures above 70 °C, collagen peptides were not detected in the spectra. This can be explained by the temperature of gelatinization or shrinkage temperature, at which the coils of collagen shrink. One of the most significant studies about the behavior of collagen in solution at different temperatures was reported by Sinex in 1960,⁶¹ in which it is explained that the deamidation of collagen follows a temperature dependent relationship only above the temperature at which the triple helix melts (T_m), which for all collagen molecules is above physiological temperatures. At this melting temperature, or T_m , the three alpha helices break and they start to exist as “statistical random coils” in solution. Once the temperature has reached the gelatinization point, or T_s (shrinkage temperature), which is above ≈ 65 °C for collagen molecules, the coils start to shrink.⁶¹ The shrinkage of collagen will affect its solubility. Therefore, there is a narrow window of temperatures which allow the study of this molecule in solution.

Optimization of the temperature for deamidation of collagen was the most difficult step. After several trials of denaturation of collagen at different times (40, 60, 80, 100, 120 min) at temperatures ranging from 40 °C to 85 °C the optimal temperature for the denaturation experiment was found to be around 62 °C. The extent of deamidation of the peptides was calculated by measuring the intensity ratios between the non-deamidated and the deamidated species at different points in the reaction (See Fig 2.10 and 2.11). Non-deamidated and deamidated peptides behave similarly under ESI conditions,⁶² thus allowing a relationship to be established between the peak intensities of each peptide. Deamidation at time zero was determined by using the first order equation and extrapolating back to time zero as previously explained in Chapter 1.

$$\mathbf{A} = \mathbf{A}_0 e^{-kt} \quad (2.1)$$

Where \mathbf{A} = concentration of species A ($N_{(t)}$, intensity of the non-deamidated monoisotopic peak), \mathbf{A}_0 = initial concentration of species A ($N_{(t)}+D_{(t)}$, where D is the intensity of the monoisotopic peak of the deamidated species), k = rate constant, t = time. The exponential coefficient in equation (2.1) represents the deamidation rate constant for the reaction. By constructing a rate curve following a first order mechanism for each peptide, as shown in Fig 2.11, it is possible to extrapolate back to time zero and obtain the initial extent of deamidation for each peptide before sample preparation, which is given by the value in the “y” axis or ($N_{(t)}/ N_{(t)}+D_{(t)}$), when time or “x” = 0 (Table 2.2).

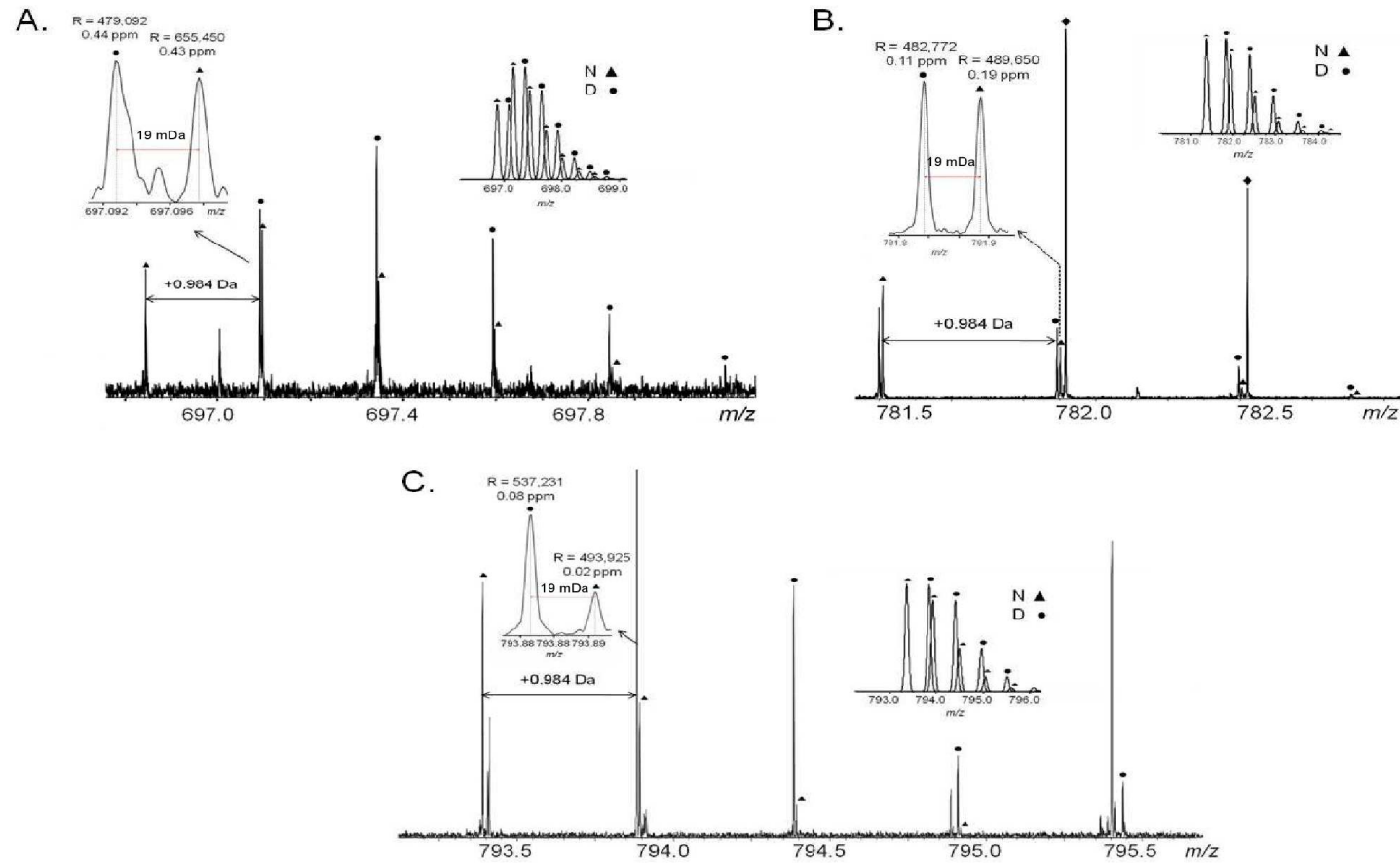


FIGURE 2.10: Snapshot of the deamidation reaction of peptides at 80 min after denatured time at 62 °C. A. $[\text{coll_I}\alpha_2]_{-1022-1051}]^{4+}$ ion at m/z 697.09326, B. $[\text{coll_I}(\alpha_1)_{-1152-1167}]^{2+}$ ion at m/z 781.89408, and C. $[\text{coll_I}(\alpha_1)_{397-414}, \text{coll_I}(\alpha_1)]^{2+}$ ion at m/z 793.88150. Triangles represent the non-deamidated peptides and the circles represent the deamidated peptides. Inserts to the right are the theoretical isotopic distributions. Inserts to the left are expansion of the A+1 isotope of the non-deamidated and the A isotope of the deamidated species.

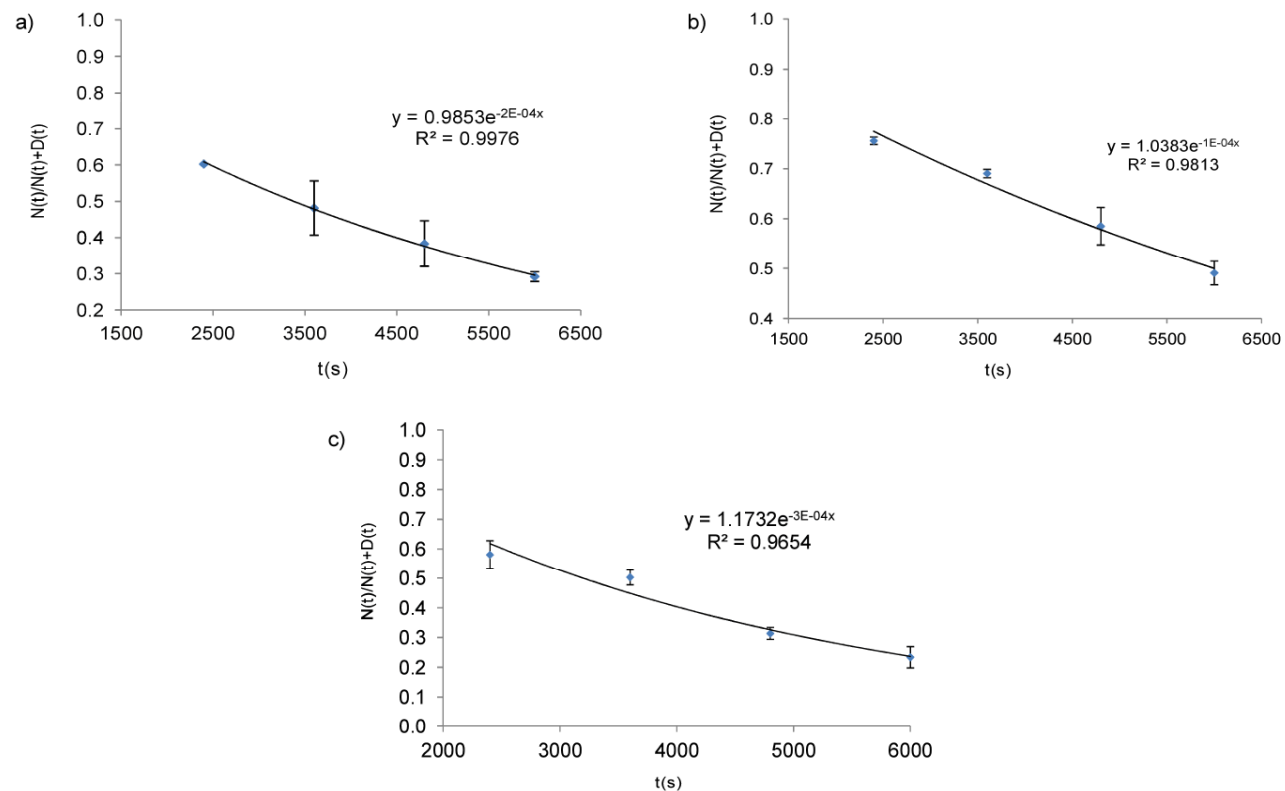


FIGURE 2.11: Exponential decay curves of the non-deamidated peptide at 62 °C. A. [coll_I (α_2) -1022-1051]⁴⁺ ion at m/z 697.09326, B.[coll_I (α_1) -1152-1167]²⁺ ion at m/z 781.89408, and C. [coll_I (α_1) 397-414, coll_I (α_1)]²⁺ ion at m/z 793.88150. The exponential coefficients represent the rate constants of deamidation for each peptide. The intercept at the y axis represents the extent of deamidation before sample preparation in each case.

TABLE 2.2: Deamidation rates and extent of deamidation of collagen peptides at 62 °C

Peptide	coll_I(α 2) -1022-1051	coll_I(α 1) -1152-1167	coll_I(α 1) 397-414
	62°C	62°C	62°C
k sec⁻¹ ± SD	(2.02±0.23)E-4	(1.23±0.10)E-4	(2.69±0.28)E-4
t_{1/2} (sec) ± SD	3428±389	5629±463	2575±269
(N_(t)/N_{(t)+D_(t))}	0.99	1.04	1.17
% D_(t)	1.0	-1.8	-0.6

It is important to point out that by inducing deamidation in the denaturation step and assuming proportionality in the increase of deamidation during protein digestion (or no deamidation at all),⁵¹ it is possible to create a curve and measure the initial extent of deamidation by extrapolating back to zero time. This assumption was proved by monitoring deamidation during protein digestion. The experiment that validates this assumption will be shown in Chapter 3.

Deamidation half lives were (3428±389) s, (5629±463) s, and (2575±269) s for coll_I (α 2) -1022-1051, coll_I (α 1) -1152-1167, and coll_I (α 1) 397-414 respectively at 62 °C. The extent of deamidation of these collagen samples was given by the intercept in the “y” axis of the exponential curves shown in Fig. 2.11 (N_(t)/N_{(t)+D_(t)) when t = 0. The extent of deamidation the peptides: coll_I (α 2) -1022-1051, coll_I (α 1) -1152-1167, and coll_I (α 1) 397-414 was found to be 0.99, 1.05, and 1.16 respectively, which means that the extent of deamidation was not elevated in the sample before preparation. The author recommends using these peptides as potential indicators of deamidation of bovine collagen. Therefore, the presence of these peptides needs further investigation in real samples.}

The deamidation half lives were found to be shorter when compared to calculated models for pentapeptides used by Robinson's group.²⁰ According to Robinson's pentapeptide deamidation calculation model, the half-lives for the Asn located between histidine (His) and glycine (Gly) is predicted to be 98,496 s, between Leu and Gly is 93,312 s, and between alanine (Ala) and Gly is 90,720 s respectively.

These half-lives were measured at 37 °C, and they range between 10 and 40 times the half-lives obtained in the work presented here, but this is mainly due to the use of high temperature which is known to accelerate the reaction remarkably.²² The complexity of collagen molecules provides a real challenge to establish a direct correlation with this model at the protein level, but it provides an idea in terms of velocity of the reaction at different temperatures.

Another published study available in the literature has reported a range of 10,800-17,400 s to release 50% of all the labile amides of whole collagen from kangaroo tail at similar temperatures.⁶¹ This last study was performed in all the amides present in collagen, and the reaction would be expected to take longer than individual releases of amides in each peptide, which is the case in the present study.

2.3.4. Differentiation of IsoAsp from Asp

The ECD of tryptic peptide [coll_I (α 1) 1152-1167]²⁺ ion at m/z 781.89408 (Fig. 2.5) shows the region where the IsoAsp diagnostic ion appears ($z_{13}^{\bullet-57}$, m/z 1204.63171) indicating deamidation at position 4 from the N-terminus or 13 from the C-terminus, and the lost of CO₂ from z_{16} (z_{16-44} , m/z 1502.77117) represents the diagnostic ion

for the presence of Asp in position 16. Ions labelled with blue originated from another peptide present in the digestion mixture with m/z 780.90998. It is important to mention that mass accuracy is critical in these experiments in order to allow the extent of deamidation to be calculated. The ECD of tryptic digest peptide [coll_I (α 1) 397-414]²⁺ ion at 793.88150 m/z (Fig. 2.6), shows the region where the IsoAsp diagnostic ion (z^{\bullet}_{16-57} , m/z 1385.68082) appears. This one corresponds to position 3 (N-terminal), 16 (C-terminal) in the peptide. Ions labelled with purple and brown, originated from other two peptides present in the complex digestion mixture with mass-to-charge m/z 790.88673 and m/z 795.90975 respectively.

In the ECD spectrum of the chosen collagen peptides, only the isoAsp diagnostic peaks ($z^{\bullet}-57$) corresponding to the deamidated Asn were observed in two of the peptides Fig. 2.5 and 2.6. One possible reason to explain the lack of Asp diagnostic ions may be interference with high intensity peaks located in the $(M+nH)^{(n-1)+\bullet}-60$ region where Asp ions should appear. Due to the absence of these peaks, deamidation rates could not be calculated using the peak intensity ratio of these two isomers.

2.4. CONCLUSIONS

A total of 30 tryptic peptides from collagen were successfully identified using FT-ICR-MS, three of which were chosen to monitor deamidation and are recommended to serve as indicators of deamidation in collagen from bovine bone. ECD was shown to be the best fragmentation technique to sequence the peptides in the study and identify the deamidation site. The use of FT-ICR-MS was critical in the differentiation of Hyp (113.04768 Da) from Leu/Ile (113.08406 Da). High resolution was needed to monitor the intensity ratios of the non-deamidated and the deamidated peptide. The first order kinetic mechanism fitted the deamidation process for each peptide and the extent of deamidation at time zero was successfully determined to be $\approx 1\%$. The method can be used to determine the extent of deamidated peptide present in the sample before sample preparation, which can be advantageous when deamidation is artificially induced.

IsoAsp diagnostic peaks ($z^{\bullet}-57$) were only observed in the spectra of two peptides. One possible reason to explain the lack of Asp peaks may be the overlapping with high intensity peaks located in the $(M+nH)^{(n-1)+\bullet}-60$ region where Asp ions should show.

2.5. REFERENCES

1. Schwarcz, H.P. Chronometric Dating in Archaeology: A Review. *Accounts of Chemical Research*. **2002**, *35*, 637-643.
2. R. E. Taylor, M.J.A., *Chronometric Dating in Archaeology*. Plenum Press: New York, **1997**; p 395.
3. Pettitt, P.B., in *Archaeological Methods*, ed. C.C. Herbert D. G. Maschner. AltaMira Press: Oxford, 2005, vol. I, p 669.
4. Willerslev, E. and Cooper, A. Ancient DNA. *Proceedings of the Royal Society B-Biological Sciences*. **2005**, *272*, 3-16.
5. Asara, M. Protein Sequences from Mastodon and Tyrannosaurus Rex Revealed by Mass Spectrometry *Science*. **2007**, *316*, 1698-1698.
6. Schweitzer, M.H. The Future of Molecular Paleontology. *Palaeontologia Electronica*. **2003**, *5*, 1-11 (<http://palaeo-electronica.org>).
7. Bada, J.L. Aspartic Acid Racemization Ages of California Paleindian Skeletons. *American Antiquity*. **1985**, *50*, 645-647.
8. Woodhead, J. Collagen: The Anatomy of a Protein. *Studies in Biology No. 117*. **1980**.
9. Leo, G., Andreotti, A., Marino, G., Pucci, P., Colombini M. P., and Birolo, L. Deamidation at Asparagine and Glutamine as a Major Modification Upon Deterioration/Aging of Proteinaceous Binders in Mural Paintings. *Analytical Chemistry*. **2011**, *83*, 2056-2064.
10. Buckley, M., Larkin, N., and Collins, M. Mammoth and Mastodon Collagen Sequences; Survival and Utility. *Geochimica et Cosmochimica Acta*. **2011**, *75*, 2007-2016.
11. Julg, A., Lafont, R., and Perinet, G. Mechanisms of Collagen Racemization in Fossil Bones. Application to Absolute Dating. *Quaternary Science Reviews*. **1987**, *6*, 25-28.
12. Van Duin, A.C.T. and Collins, M.J. The Effects of Conformational Constraints on Aspartic Acid Racemization. *Organic Geochemistry*. **1998**, *29*, 1227-1232.
13. Collins, M.J., Penkman, K.E.H., Rohland, N., Shapiro, B., Dobberstein, R.C., Ritz-Timme, S., and Hofreiter, M. Is Amino Acid Racemization a Useful Tool for Screening for Ancient DNA in Bone? *Proceedings of the Royal Society B-Biological Sciences*. **2009**, *276*, 2971-2977.
14. Buckley, M., Walker, A., Ho, S.Y.W., Yang, Y., Smith, C., Ashton, P., Thomas-Oates, J., Cappellini, E., Koon, H., Penkman, K., Elsworth, B., Ashford, D., Solazzo, C., Andrews, P., Strahler, J., Shapiro, B., Ostrom, P., Gandhi, H., Miller, W., Raney, B., Zylber, M.I., Gilbert, M.T.P., Prigodich, R.V., Ryan, M., Rijdsdijk, K.F., Janoo, A., and Collins, M.J. Comment on "Protein Sequences from Mastodon and Tyrannosaurus Rex Revealed by Mass Spectrometry". *Science*. **2008**, *319*, 5859.
15. Robinson, A.B. and Rudd, C.J. Deamidation of Glutaminyl and Asparaginyl Residues in Peptides and Proteins. *Current Topics in Cellular Regulation*. **1974**, *8*, 247-95.
16. Robinson, A.B., Mckerrow, J.H., and Cary, P. Controlled Deamidation of Peptides and Proteins - an Experimental Hazard and a Possible Biological Timer. *Proceedings of the National Academy of Sciences of the United States of America*. **1970**, *66*, 753-757.

17. Midelfor, C.F. and Mehler, A.H. Deamidation in-Vivo of an Asparagine Residue of Rabbit Muscle Aldolase. *Proceedings of the National Academy of Sciences of the United States of America*. **1972**, *69*, 1816-1819.
18. Robinson, A.B., Mckerrow, J.H., and Legaz, M. Sequence Dependent Deamidation Rates for Model Peptides of Cytochrome-C. *International Journal of Peptide and Protein Research*. **1974**, *6*, 31-35.
19. Mckerrow, J.H. and Robinson, A.B. Primary Sequence Dependence of Deamidation of Rabbit Muscle Aldolase. *Science*. **1974**, *183*, 85-85.
20. Robinson, N.E. and Robinson, A.B. Molecular Clocks. *Proceedings of the National Academy of Sciences of the United States of America*. **2001**, *98*, 944-949.
21. Robinson, N.E., Robinson, A.B., and Merrifield, R.B. Mass Spectrometric Evaluation of Synthetic Peptides as Primary Structure Models for Peptide and Protein Deamidation. *Journal of Peptide Research*. **2001**, *57*, 483-493.
22. Robinson, N.E. and Robinson, A.B., *Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*. Althouse Press: Cave Junction, OR, **2004**.
23. Miesbauer, L.R., Zhou, X., Yang, Z., Sun, Y., Smith, D.L., and Smith, J.B. Post-Translational Modifications of Water-Soluble Human Lens Crystallins from Young Adults. *Journal of Biological Chemistry*. **1994**, *269*, 12494-502.
24. David, L.L., Lampi, K.J., Lund, A.L., and Smith, J.B. The Sequence of Human B1-Crystallin Cdna Allows Mass Spectrometric Detection of B1 Protein Missing Portions of Its N-Terminal Extension. *Journal of Biological Chemistry*. **1996**, *271*, 4273-4279.
25. Lampi, K.J., Ma, Z., Shih, M., Shearer, T.R., Smith, J.B., Smith, D.L., and David, L.L. Sequence Analysis of Ba3, Bb3, and Ba4 Crystallins Completes the Identification of the Major Proteins in Young Human Lens. *Journal of Biological Chemistry*. **1997**, *272*, 2268-2275.
26. Smith, J.B., Miesbauer, L.R., and Sun, Y.P. Post Translational Modification of Crystallins Isolated from Human Lenses. *Pure and Applied Chemistry*. **1994**, *66*, 95-100.
27. Hanson, S.R.A., Smith, D.L., and Smith, J.B. Deamidation and Disulfide Bonding in Human Lens Γ -Crystallins. *Experimental Eye Research*. **1998**, *67*, 301-312.
28. Bloemendal, H., De Jong, W., Jaenicke, R., Lubsen, N.H., Slingsby, C., and Tardieu, A. Ageing and Vision: Structure, Stability and Function of Lens Crystallins. *Progress in Biophysics & Molecular Biology*. **2004**, *86*, 407-485.
29. Wistow, G.J. and Piatigorsky, J. Lens Crystallins - the Evolution and Expression of Proteins for a Highly Specialized Tissue. *Annual Review of Biochemistry*. **1988**, *57*, 479-504.
30. Wilmarth, P.A., Tanner, S., Dasari, S., Nagalla, S.R., Riviere, M.A., Bafna, V., Pevzner, P.A., and David, L.L. Age-Related Changes in Human Crystallins Determined from Comparative Analysis of Post-Translational Modifications in Young and Aged Lens: Does Deamidation Contribute to Crystallin Insolubility? *Journal of Proteome Research*. **2006**, *5*, 2554-2566.
31. Lund, A.L., Smith, J.B., and Smith, D.L. Modifications of the Water-Insoluble Human Lens A-Crystallins. *Experimental Eye Research*. **1996**, *63*, 661-672.

32. Zhang, Z.L., Smith, D.L., and Smith, J.B. Human Beta-Crystallins Modified by Backbone Cleavage, Deamidation and Oxidation Are Prone to Associate. *Experimental Eye Research*. **2003**, *77*, 259-272.
33. Lampi, K.J., Amyx, K.K., Ahmann, P., and Steel, E.A. Deamidation in Human Lens Beta B2-Crystallin Destabilizes the Dimer. *Biochemistry*. **2006**, *45*, 3146-3153.
34. Hanson, S.R.A., Hasan, A., Smith, D.L., and Smith, J.B. The Major in Vivo Modifications of the Human Water-Insoluble Lens Crystallins Are Disulfide Bonds, Deamidation, Methionine Oxidation and Backbone Cleavage. *Experimental Eye Research*. **2000**, *71*, 195-207.
35. Robinson, N.E., Robinson, Z.W., Robinson, B.R., Robinson, A.L., Robinson, J.A., Robinson, M.L., and Robinson, A.B. Structure-Dependent Nonenzymatic Deamidation of Glutamyl and Asparagyl Pentapeptides. *Journal of Peptide Research*. **2004**, *63*, 426-436.
36. Noah E. Robinson, M.L.R., Stephanie E. S. Schulze, Bert T. Lai, and Harry B. Gray. Deamidation of α -Synuclein. *Protein Science*. **2010**, *18*, 1766-1773.
37. Lapko, V.N., Purkiss, A.G., Smith, D.L., and Smith, J.B. Deamidation in Human Gamma S-Crystallin from Cataractous Lenses Is Influenced by Surface Exposure. *Biochemistry*. **2002**, *41*, 8638-8648.
38. Harms, M.J., Wilmarth, P.A., Kapfer, D.M., Steel, E.A., David, L.L., Bachinger, H.P., and Lampi, K.J. Laser Light-Scattering Evidence for an Altered Association of Beta B1-Crystallin Deamidated in the Connecting Peptide. *Protein Science*. **2004**, *13*, 678-686.
39. Zomber, G., Reuveny, S., Garti, N., Shafferman, A., and Elhanany, E. Effects of Spontaneous Deamidation on the Cytotoxic Activity of the Bacillus Anthracis Protective Antigen. *Journal of Biological Chemistry*. **2005**, *280*, 39897-39906.
40. Vanbelle, C., Halgand, F., Cedervall, T., Thulin, E., Akerfeldt, K.S., Laprevote, O., and Linse, S. Deamidation and Disulfide Bridge Formation in Human Calbindin D-28k with Effects on Calcium Binding. *Protein Science*. **2005**, *14*, 968-979.
41. Solstad, T., Carvalho, R.N., Andersen, O.A., Waidelich, D., and Flatmark, T. Deamidation of Labile Asparagine Residues in the Autoregulatory Sequence of Human Phenylalanine Hydroxylase - Structural and Functional Implications. *European Journal of Biochemistry*. **2003**, *270*, 929-938.
42. Carvalho, R.N., Solstad, T., Bjorgo, E., Barroso, J.F., and Flatmark, T. Deamidations in Recombinant Human Phenylalanine Hydroxylase - Identification of Labile Asparagine Residues and Functional Characterization of Asn -> Asp Mutant Forms. *Journal of Biological Chemistry*. **2003**, *278*, 15142-15152.
43. Cournoyer, J.J., Pittman, J.L., Ivleva, V.B., Fallows, E., Waskell, L., Costello, C.E., and O'connor, P.B. Deamidation: Differentiation of Aspartyl from Isoaspartyl Products in Peptides by Electron Capture Dissociation. *Protein Science*. **2005**, *14*, 452-463.
44. Cournoyer, J.J., Lin, C., and O'connor, P.B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. *Analytical Chemistry*. **2006**, *78*, 1264-1271.
45. Sargaeva, N.P., Lin, C., and O'connor, P.B. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid Beta Peptides, Including a Beta 1-42, Using Electron-Ion Reactions. *Analytical Chemistry*. **2009**, *81*, 9778-9786.

46. Gonzalez, L.J., Shimizu, T., Satomi, Y., Betancourt, L., Besada, V., Padron, G., Orlando, R., Shirasawa, T., Shimonishi, Y., and Takao, T. Differentiating Alpha- and Beta-Aspartic Acids by Electrospray Ionization and Low-Energy Tandem Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2000**, *14*, 2092-2102.
47. Lloyd, J.R., Cotter, M.L., Ohori, D., and Doyle, D.L. Distinction of Alpha-Aspartyl and Beta-Aspartyl and Alpha-Glutamyl and Gamma-Glutamyl-Transferase Peptides by Fast Atom Bombardment Tandem Mass Spectrometry. *Biomedical and Environmental Mass Spectrometry*. **1988**, *15*, 399-402.
48. Cournoyer, J.J., Lin, C., Bowman, M.J., and O'connor, P.B. Quantitating the Relative Abundance of Isoaspartyl Residues in Deamidated Proteins by Electron Capture Dissociation. *Journal of the American Society for Mass Spectrometry*. **2007**, *18*, 48-56.
49. Cournoyer, J.J., Pittman, J.L., Ivleva, V.B., Fallows, E., Waskell, L., Costello, C.E., and O'connor, P.B. Deamidation: Differentiation of Aspartyl from Isoaspartyl Products in Peptides by Electron Capture Dissociation. *Protein Science*. **2005**, *14*, 452-463.
50. Sargaeva, N.P., Lin, C., and O'connor, P.B. Differentiating N-Terminal Aspartic and Isoaspartic Acid Residues in Peptides. *Analytical Chemistry*. **2011**, *83*, 6675-6682.
51. Li, X., Cournoyer, J.J., Lin, C., and O'connor, P.B. Use of O-18 Labels to Monitor Deamidation During Protein and Peptide Sample Processing. *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 855-864.
52. Jennings, K.R. Collision-Induced Decompositions of Aromatic Molecular Ions. *International Journal of Mass Spectrometry and Ion Physics*. **1968**, *1*, 227-235.
53. Tsybin, Y., Quinn, J., Tsybin, O., Hendrickson, C., and Marshall, A. Electron Capture Dissociation Implementation Progress in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 762-771.
54. Whitford, D., *Proteins Structure and Function*. John Wiley & Sons: Chichester, **2005**; Vol. I.
55. Berg, R.A. and Prockop, D.J. Thermal Transition of a Non-Hydroxylated Form of Collagen. Evidence for a Role for Hydroxyproline in Stabilizing Triple Helix of Collagen. *Biochemical and Biophysical Research Communications*. **1973**, *52*, 115-120.
56. Kassel, D.B. and Biemann, K. Differentiation of Hydroxyproline Isomers and Isobars in Peptides by Tandem Mass Spectrometry. *Analytical Chemistry*. **1990**, *62*, 1691-1695.
57. Leymarie, N., Berg, E.A., McComb, M.E., O'connor, P.B., Grogan, J., Oppenheim, F.G., and Costello, C.E. Tandem Mass Spectrometry for Structural Characterization of Proline-Rich Proteins: Application to Salivary Prp-3. *Analytical Chemistry*. **2002**, *74*, 4124-4132.
58. Geiger, T. and Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides. *J Biol Chem*. **1987**, *262*, 785-794.
59. Aswad, D.W., Paranandi, M.V., and Schurter, B.T. Isoaspartate in Peptides and Proteins: Formation, Significance, and Analysis. *Journal of Pharmaceutical and Biomedical Analysis*. **2000**, *21*, 1129-1136.

60. Radkiewicz, J.L., Zipse, H., Clarke, S., and Houk, K.N. Accelerated Racemization of Aspartic Acid and Asparagine Residues Via Succinimide Intermediates: An Ab Initio Theoretical Exploration of Mechanism. *Journal of the American Chemical Society*. **1996**, *118*, 9148-9155.
61. Sinex, F.M. Aging and the Lability of Irreplaceable Molecules. Ii. The Amide Groups of Collagen. *Journals of Gerontology*. **1960**, *15*, 15-18.
62. Stroop, S.D. A Modified Peptide Mapping Strategy for Quantifying Site-Specific Deamidation by Electrospray Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2007**, *21*, 830-836.

2.6. MASS ACCURACY AND PEPTIDE ASSIGNMENTS

TABLE 2.3: Peak list generated from the externally calibrated CAD spectrum of [GHNGLQGLPGLAGHHGDQGAPGAVGPAGPR+4H]⁴⁺
m/z 697.09326

Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)	Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)
y ₁₀ ²⁺	447.74322	447.74341	-0.42	b ₂₄ ³⁺	744.68952	744.69007	-0.74
y ₅	497.28304	497.28328	-0.48	[b ₂₅ -H ₂ O-NH ₃] ³⁺	752.01804	752.01822	-0.24
y ₁₁ ²⁺	483.261775	483.26201	-0.49	b ₂₅ ³⁺	763.69667	763.69727	-0.78
y ₆	554.30450	554.3048	-0.54	b ₈	778.38420	778.38479	-0.76
b ₁₈ ³⁺	588.61178	588.61214	-0.62	y ₉	781.43148	781.43207	-0.76
[b ₁₉ -H ₂ O-NH ₃] ³⁺	595.94029	595.9403	-0.01	b ₁₇ ²⁺	818.38474	818.38542	-0.84
b ₁₉ ³⁺	607.61893	607.6193	-0.61	b ₂₇ ³⁺	819.72663	819.72722	-0.72
[b ₂₀ -H ₂ O-NH ₃] ³⁺	619.61933	619.61933	0.00	y ₂₇ ³⁺	826.08616	826.0868	-0.77
[b ₂₀ -H ₂ O] ³⁺	625.29482	625.29482	0.00	y ₁₈ ²⁺	827.39564	827.3963	-0.80
b ₂₀ ³⁺	631.29797	631.29835	-0.61	y ₁₉ ²⁺	862.91419	862.91495	-0.88
y ₇	653.37291	653.37333	-0.64	y ₂₈ ³⁺	864.42847	864.42921	-0.85
b ₂₁ ³⁺	668.98053	668.98098	-0.68	y ₁₀	894.47916	894.4799	-0.83
y ₂₂ ³⁺	669.99629	669.99674	-0.68	b ₁₉ ²⁺	910.92476	910.9256	-0.93
b ₂₂ ³⁺	687.98768	687.98813	-0.65	[b ₂₀ -H ₂ O-NH ₃] ₂ ⁺	928.92509	928.92558	-0.53
[MH-H ₂ O] ⁴⁺	692.59104	692.59109	-0.07	[b ₂₀ -H ₂ O] ²⁺	937.43832	937.43891	-0.63
[M+4H] ⁴⁺	697.09327	697.09374	-0.68	b ₂₀ ²⁺	946.44331	946.44417	-0.91
[b ₂₃ -H ₂ O-NH ₃] ³⁺	699.98808	699.98819	-0.16	y ₂₂ ²⁺	1004.49079	1004.49182	-1.03
[b ₂₃ -H ₂ O] ³⁺	705.66357	705.66361	-0.05	y ₁₂	1022.53773	1022.53877	-1.02
b ₂₃ ³⁺	711.66672	711.66720	-0.68	b ₂₂ ²⁺	1031.47788	1031.47926	-1.34

TABLE 2.3: Peak list generated from the externally calibrated CAD spectrum of [GHNGLQGLPGLAGHHGDQGAPGAVGPAGPR+4H]⁴⁺
m/z 697.09326 (Continued)

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
y ₈	724.41002	724.41053	-0.70	[b ₂₃ -H ₂ O] ²⁺	1057.99145	1057.99228	-0.79
y ₈	724.41002	724.41053	-0.70	[b ₂₃ -H ₂ O] ²⁺	1057.99145	1057.99228	-0.79
y ₂₄ ³⁺	726.69813	726.69857	-0.61	b ₂₃ ²⁺	1066.99644	1066.99751	-1.01
y ₁₆ ²⁺	730.35545	730.35592	-0.64	b ₂₄ ²⁺	1116.53064	1116.53185	-1.08
[b ₂₄ -H ₂ O-NH ₃] ³⁺	733.01088	733.011	-0.16	b ₂₅ ²⁺	1145.04137	1145.0427	-1.16
[b ₂₄ -H ₂ O] ³⁺	738.68638	738.68657	-0.26	y ₁₃	1150.59631	1150.59768	-1.19

TABLE 2.4: Peak list generated from the ECD spectrum of [GHNGLQGLPGLAGHHGDQGAPGAVGPAGPR+4H]⁴⁺ m/z 697.09326

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z ₃	313.17445	313.17445	0.00	c ₁₁	1078.56395	1078.56399	-0.04
y ₃	329.19317	329.19326	-0.27	z ₂₄ ²⁺ •	1081.53419	1081.53434	-0.14
z ₄ [•]	384.21156	384.21163	-0.18	z ₂₄ ²⁺	1082.03581	1082.03588	-0.06
c ₅	497.24665	497.24667	-0.04	y ₂₄ ²⁺	1089.54355	1089.54399	-0.40
z ₆ [•]	538.28578	538.28579	-0.02	c ₂₄ ²⁺	1125.043915	1125.04394	-0.02
a ₆	581.29159	581.29126	0.57	y ₁₃ -NH ₃	1133.569836	1133.56947	0.32
c ₆	625.30523	625.30517	0.10	z ₁₃ [•]	1134.57759	1134.57768	-0.08
z ₇ [•]	637.35419	637.35414	0.08	z ₂₅ ²⁺ •	1145.56348	1145.56277	0.62
c ₇	682.32669	682.3267	-0.01	c ₁₂	1149.60106	1149.60096	0.09
[M+3H]3+	689.99584	689.99498	1.24	c ₂₆ ²⁺	1202.081025	1202.08104	-0.01
[z ₂₃ -NH ₃] ³⁺	691.33672	691.33759	-1.26	c ₁₃	1206.62252	1206.62251	0.01
[M+4H]4+	697.09327	697.09272	0.79	y ₂₆ ²⁺	1210.11487	1210.1146	0.22
z ₈ [•]	708.3913	708.39126	0.06	y ₁₄ -H ₂ O-2NH ₃	1213.559787	1213.55928	0.42
a ₁₅ ²⁺	718.86699	718.86622	1.07	z ₂₇ ²⁺	1230.61624	1230.6161	0.11
z ₁₆ ²⁺ •	722.34609	722.34611	-0.03	z ₂₇ ²⁺	1231.12015	1231.11936	0.64
y ₈	724.41002	724.40995	0.10	c ₂₇ ²⁺	1237.59958	1237.59952	0.05
c ₁₅ ²⁺	740.87381	740.87384	-0.04	y ₂₇ ²⁺	1238.6256	1238.6256	0.00
z ₉ [•]	765.41276	765.41291	-0.20	y ₁₄ -NH ₃	1248.596776	1248.59636	0.33
c ₁₆ ²⁺	769.38454	769.3845	0.05	z ₁₄ [•]	1249.60453	1249.60468	-0.12

TABLE 2.4: Peak list generated from the ECD spectrum of [GHNGLQGLPGLAGHHGDQGAPGAVGPAGPR+4H]⁴⁺ *m/z* 697.09326 (Continued)

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
y ₉	781.43148	781.43118	0.38	z ₂₈ ^{2+•}	1288.12971	1288.12981	-0.08
z ₁₈ ^{2+•}	819.386275	819.38635	-0.09	z ₁₅ [•]	1306.62599	1306.62596	0.02
c ₁₇ ²⁺	826.89801	826.89782	0.23	c ₂₉ ²⁺	1314.63669	1314.63653	0.12
z ₁₉ ^{2+•}	854.90483	854.90465	0.21	y ₁₅	1322.64471	1322.64465	0.05
a ₁₈ ²⁺	868.92048	868.92031	0.20	c ₁₄ [•]	1342.67361	1342.67348	0.10
c ₁₈ ²⁺	890.9273	890.92736	-0.07	c ₁₄	1343.68143	1343.68138	0.04
a ₁₉ ²⁺	897.43121	897.43052	0.77	z ₃₀ ²⁺	1385.673805	1385.67385	-0.03
c ₉	908.45843	908.4585	-0.08	[M+4H] ^{2+••}	1394.187615	1394.18696	0.47
z ₂₀ ^{2+•}	911.44686	911.44699	-0.14	z ₁₆ [•] -H ₂ O	1425.674385	1425.67534	-0.67
c ₁₉ ²⁺	919.43803	919.43813	-0.11	z ₁₆ [•]	1443.6849	1443.68545	-0.38
[z ₃₀ [•] -NH ₃] ³⁺	923.7827167	923.78237	0.38	z ₁₆	1444.69272	1444.69268	0.03
[M+3H] ³⁺	929.12193	929.12269	-0.82	c ₁₅	1480.74034	1480.74034	0.00
[M+4H] ^{3+•}	929.45841	929.4579	0.55	c ₁₆	1537.7618	1537.76187	-0.05
z ₂₁ ^{2+•}	939.95759	939.95731	0.30	z ₁₇	1581.75163	1581.75193	-0.19
y ₁₁ -NH ₃	948.489796	948.48938	0.44	z ₁₈	1638.77309	1638.77306	0.02
z ₁₁ [•]	949.49755	949.4975	0.05	c ₁₇	1652.78874	1652.78875	-0.01
c ₁₀	965.47989	965.47989	0.00	z ₁₉ -H ₂ O	1675.804765	1675.8032	0.93

TABLE 2.4: Peak list generated from the ECD spectrum of
 GHNGLQGLPGLAGHHGDQGAPGAVGPAGPR+4H]⁴⁺ *m/z* 697.09326 (Continued)

Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)	Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)
z_{12}^{\bullet}	1006.51901	1006.5192	-0.19	z_{19}	1709.8102	1709.81047	-0.16
c_{21}^{2+}	1011.480425	1011.48064	-0.21	z_{20}	1822.89426	1822.89436	-0.05
y_{12}	1022.53773	1022.53768	0.05	c_{19}	1837.86878	1837.86939	-0.33
c_{22}^{2+}	1039.991155	1039.99121	-0.05	z_{21}	1879.91572	1879.91525	0.25
$z_{23}^{2+\bullet}$	1053.02346	1053.02357	-0.10				

TABLE 2.5: CAD peak table [DGLNGLPGPIGPPGPR+2H]²⁺ *m/z* 781.89408

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
b ₄	401.16667	401.16687	-0.50	(b ₁₀ -H ₂ O)	777.35261	777.35343	-1.06
y ₄	442.24085	442.24110	-0.57	[M+2H] ²⁺	780.91005	780.91088	-1.06
b ₅	458.18813	458.18843	-0.65	[M+2H] ²⁺	781.89408	781.89497	-1.14
(b ₆ -2H ₂ O)	477.20930	477.20951	-0.44	y ₈	822.44681	822.44770	-1.08
y ₁₀ ²⁺	496.76162	496.76195	-0.67	y ₉	823.47843	823.47934	-1.11
y ₅	499.29869	499.29903	-0.68	y ₉	879.46827	879.46929	-1.16
b ₆	513.23033	513.23070	-0.72	y ₁₀	894.51554	894.51662	-1.21
(b ₇ -2H ₂ O)	534.23076	534.23104	-0.52	b ₁₀ -2H ₂ O	915.45712	915.45811	-1.08
b ₆ -2H ₂ O	535.25116	535.25146	-0.56	a ₁₀	924.49106	924.49082	0.26
a _{.6}	543.27734	543.27769	-0.64	b ₁₀ -H ₂ O	933.46764	933.46878	-1.23
(b ₇ -H ₂ O)	552.24128	552.24166	-0.70	b ₁₀	951.47815	951.47936	-1.27
b ₆ -H ₂ O	553.26168	553.26206	-0.70	b ₁₁ -2H ₂ O	972.47858	972.47973	-1.18
y ₆	556.32015	556.32058	-0.77	(b ₁₂ -H ₂ O)	987.48943	987.49075	-1.34
b ₇	570.25179	570.25224	-0.79	y ₁₁	991.56830	991.56969	-1.40
b ₆	571.27219	571.27264	-0.79	y ₁₀	992.51595	992.51699	-1.05
y ₁₂ ²⁺	581.81438	581.81482	-0.76	(b ₁₃ -H ₂ O)	1044.51089	1044.51221	-1.27
y ₁₄ ²⁺	608.84346	608.84399	-0.88	y ₁₂	1048.58976	1048.59116	-1.34
y ₆	612.30999	612.31001	-0.03	b ₁₃	1062.52140	1062.52299	-1.50
y ₁₃ ²⁺	639.32785	639.32841	-0.88	y ₁₁	1105.60001	1105.60154	-1.38

TABLE 2.5: CAD peak table [DGLNGLPGPIGPPPGPR+2H]²⁺ m/z 781.89408 (Continued)

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
[y ₁₆ -H ₂ O] ²⁺	678.87304	678.87342	-0.56	y ₁₃	1119.62687	1119.62839	-1.36
y ₁₆ ²⁺	687.87803	687.87868	-0.95	y ₁₂	1162.62147	1162.62318	-1.47
y ₁₄ ²⁺	695.86988	695.87059	-1.03	(b ₁₅ -H ₂ O)	1240.63206	1240.63403	-1.59
(b ₉ -2H ₂ O)	702.32063	702.32124	-0.87	b ₁₅	1258.64257	1258.64448	-1.52
z ₇	710.38315	710.38335	-0.28	y ₁₅	1273.70109	1273.70316	-1.63
(b ₉ -H ₂ O)	720.33115	720.33184	-0.96	y ₁₃	1277.64841	1277.65028	-1.46
y ₇	725.39405	725.39473	-0.94	(b ₁₆ -H ₂ O)	1297.65352	1297.65532	-1.39
b ₉	738.34166	738.34241	-1.02	b ₁₆	1315.66403	1315.66629	-1.72
(b ₁₀ -2H ₂ O)	759.34209	759.34281	-0.95	(b ₁₇ -H ₂ O)	1368.69063	1368.69292	-1.68
y ₈	766.45697	766.45780	-1.08	b ₁₇	1386.70114	1386.70340	-1.63
[y ₁₈ -H ₂ O] ²⁺	771.90506	771.90562	-0.72				

Ions labelled with colour blue are generated from the peptide **GETGPAGPAGPIGPVGAR** with m/z 780.90998

TABLE 2.6: Peak-list generated from the ECD spectrum of [DGLNGLPGPIGPPGPR+2H]²⁺ *m/z* 781.89408

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z ₃	314.18227	314.18224	0.10	z ₁₂ [•]	1146.60275	1146.60289	-0.12
y ₄	442.24085	442.24085	0.00	y ₁₂	1162.62147	1162.62144	0.03
z₆	541.30925	541.30924	0.02	z₁₃[•]-57	1204.63209	1204.63171	0.32
y ₅	555.28853	555.28871	-0.32	z₁₅[•]	1257.68237	1257.68187	0.40
z ₆	597.29909	597.29902	0.12	z ₁₃ [•]	1261.62969	1261.62935	0.27
y ₆	612.30999	612.31001	-0.03	y ₁₅	1273.70109	1273.70100	0.07
z ₇	710.38315	710.38335	-0.28	c₁₆[•]	1331.68276	1331.68298	-0.17
y ₇	725.39405	725.39422	-0.23	(y₁₆-NH₃)	1357.72230	1357.72180	0.37
b₁₈²⁺	771.90477	771.90467	0.12	z ₁₆	1359.73787	1359.73759	0.21
[M+2H]²⁺	780.91005	780.90998	0.09	z ₁₄ [•]	1374.71375	1374.71388	-0.09
[M+2H] ²⁺	781.89408	781.89399	0.11	c₁₇[•]	1402.71987	1402.71961	0.19
z₉[•]	807.45971	807.45940	0.38	c ₁₅ [•]	1404.68792	1404.68640	1.08
z₉	808.46753	808.46752	0.01	z₁₆[•]-CO₂	1502.77235	1502.77117	0.79
y₉	823.47843	823.47865	-0.27	a ₁₆	1517.78321	1517.78225	0.63
z ₉ [•]	863.44955	863.44947	0.09	(z₁₈[•]-NH₃)	1527.76763	1527.76514	1.63
z ₉	864.45737	864.45730	0.08	z₁₈[•]	1544.79410	1544.79270	0.91
z₁₀[•]	878.49682	878.49666	0.18	z ₁₆ [•]	1546.76215	1546.76170	-0.29
y ₁₀	992.51595	992.51586	0.09	[M+H]⁺	1560.81282	1560.81188	0.60

TABLE 2.6: Peak list generated from the ECD spectrum of [DGLNGLPGPIGPPGPR+2H]²⁺ *m/z* 781.89408
(Continued)

Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)	Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)
<i>z</i> ₁₁ [•]	1089.58129	1089.58071	0.53	[M+2H] ⁺ [•]	1563.78923	1563.78801	0.78
<i>z</i> ₁₃ [•]	1103.60815	1103.60669	1.32				

Ions labelled with colour blue are originated from the peptide **GETGPAGPAGPIGPVGAR** with *m/z* 780.90998

TABLE 2.7: Peak list generated from the externally calibrated CAD spectrum of [GANGAPPGIAGAPPGFPGAR]²⁺ *m/z* 793.88150

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
y ₁₃ ²⁺	608.30946	608.30957	-0.19	[M+2H]2+	793.88150	793.88202	-0.66
b ₈ -H ₂ O	637.29404	637.29378	0.40	b ₁₁ -H ₂ O	836.389715	836.38951	0.25
y ₁₄ ²⁺	643.82801	643.82811	-0.16	b ₁₁	854.40023	854.40100	-0.90
b ₈	655.30455	655.30464	-0.14	y ₉	861.42132	861.42173	-0.48
b ₉ -H ₂ O	708.33115	708.33106	0.12	y ₁₀	932.45843	932.45776	0.72
b ₉	726.34166	726.34182	-0.22	b ₁₄ -H ₂ O	1153.527265	1153.52593	1.16
y ₁₆ ²⁺	729.85221	729.85244	-0.32	b ₁₄	1171.53778	1171.53679	0.85
y ₇	733.36275	733.36295	-0.27	y ₁₃	1215.61163	1215.61007	1.28
b ₁₈ ²⁺	784.87621	784.87719	-1.25				

TABLE 2.8: Peak list generated from the ECD spectrum of [GANGAPGIAGAPGFPGAR]²⁺ *m/z* 793.88150

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z ₄	401.21430	401.21450	-0.50	y ₁₂	1040.5483	1040.54833	-0.03
y₄	416.22520	416.22520	0.00	y' ₁₂	1051.49286	1051.49291	-0.05
z ₅	547.27489	547.27480	0.16	y ₁₂	1052.50068	1052.5008	-0.11
z ₅	548.28271	548.28270	0.02	z ₁₂ [•]	1086.54523	1086.54563	-0.37
y ₅	563.29361	563.29348	0.23	y ₁₂	1102.56395	1102.56413	-0.16
z₆	571.31982	571.31979	0.05	c ₁₃ [•]	1104.59217	1104.59229	-0.11
y₆	586.33072	586.33079	-0.12	c ₁₃	1105.59999	1105.59996	0.03
z ₆ [•]	604.29635	604.29605	0.50	z ₁₃ [•]	1111.56161	1111.56136	0.22
z ₆	605.30417	605.30423	-0.10	z ₁₃	1112.56943	1112.57062	-1.07
y ₁₃ ²⁺	608.30946	608.30924	0.35	x ₁₂	1129.55104	1129.55067	0.33
b ₈ -2H ₂ O	619.28352	619.28326	0.42	b ₁₄ -2H ₂ O	1135.51675	1135.51701	-0.23
y ₆	620.31507	620.31500	0.11	(b₁₄-2H₂O)[◊]	1141.60004	1141.6002	-0.14
(y₇-NH₃)	670.35193	670.35187	0.08	b ₁₄ -H ₂ O	1153.527265	1153.52668	0.51
b ₉ -2H ₂ O	690.32063	690.32074	-0.16	y ₁₃	1215.61163	1215.61167	-0.03
b ₉ -H ₂ O	708.331145	708.33075	0.56	z ₁₄ [•]	1270.63002	1270.63002	0.00
z ₇	718.35185	718.35193	-0.11	c ₁₅ [•]	1288.67696	1288.67543	1.19
y₇	733.36275	733.36275	0.00	c ₁₅	1289.68478	1289.6849	-0.09
z₉	770.36788	770.36799	-0.14	z ₁₅ [•]	1303.64024	1303.64032	-0.06
b ₁₈ ²⁺	784.87621	784.87612	0.11	z ₁₅ [•]	1327.65148	1327.65151	-0.02

TABLE 2.8: Peak list generated from the ECD spectrum of [GANGAPGIAGAPGFPGAR]²⁺ m/z 793.88150
(Continued)

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z_8^\bullet	788.38114	788.38099	0.19	y'_{15}	1342.66238	1342.66306	-0.51
z_8	789.38896	789.3889	0.08	y_{15}	1343.6702	1343.67047	-0.20
y_8	804.39986	804.3998	0.07	c_{16}	1346.70624	1346.70633	-0.07
$b_{11}-2H_2O$	818.37920	818.37906	0.17	c_{16}^\bullet	1357.62565	1357.62566	-0.01
z_9^\bullet	825.43390	825.43411	-0.25	$z_{16}^\bullet -57$	1385.68082	1385.68082	0.00
z_9	826.44172	826.44169	0.04	$(y_{16}-NH_3)$	1403.680166	1403.67985	0.23
$b_{11}-H_2O$	836.389715	836.38973	-0.02	z_{16}^\bullet	1404.68792	1404.68965	-1.23
y_9	841.45262	841.45256	0.07	z_{16}	1405.69574	1405.69616	-0.30
z_9^\bullet	845.40260	845.40239	0.25	$(c_{17}-2H_2O)$	1425.71215	1425.70959	1.80
z_9	846.41042	846.41046	-0.05	c_{17}^\bullet	1428.66276	1428.66298	-0.15
b_{11}	854.40023	854.40048	-0.29	z_{16}^\bullet	1442.67842	1442.67857	-0.10
y_9	861.42132	861.42123	0.10	c_{17}	1461.73318	1461.73348	-0.21
z_{10}^\bullet	896.47101	896.47081	0.22	z_{17}^\bullet	1513.71553	1513.71553	0.00
z_{10}	897.47883	897.47926	-0.48	z_{17}^\bullet	1517.77198	1517.77245	-0.31
z_{10}^\bullet	916.43971	916.43986	-0.16	$z_{18}^\bullet -CO_2$	1526.74719	1526.74724	-0.03
z_{10}	917.44753	917.44761	-0.09	x_{17}^\bullet	1559.76997	1559.76861	0.87
y_{10}	932.45843	932.45847	-0.04	z_{18}^\bullet	1564.74756	1564.74658	0.63

TABLE 2.8: Peak list generated from the ECD spectrum of [GANGAPGIAGAPPGFPGAR]²⁺ *m/z* 793.88150
(Continued)

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)	Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
c₁₁[•]	976.5336	976.53346	0.14	z₁₈	1565.75538	1565.75122	2.66
c₁₁	977.54142	977.54175	-0.34	z ₁₈ [•]	1570.73699	1570.73707	-0.05
z₁₂[•]	1024.52958	1024.52976	-0.18	[M+2H]^{+•}	1581.77464	1581.77467	-0.02
z₁₂	1025.5374	1025.53756	-0.16	[M+H] ⁺	1586.75571	1586.75567	0.03
z ₁₁ [•]	1029.52377	1029.52394	-0.17	[M+2H] ^{+•}	1587.76407	1587.76373	0.21
z ₁₂ [•]	1036.48196	1036.4819	0.06	[M+2H]^{+•}	1591.82052	1591.82019	0.21
z ₁₂	1037.48978	1037.48973	0.05				

Ions labelled with colour purple are originated from the peptide **GPPGESGAAGPTGPIGSR** with *m/z* 790.88673

Ions labelled with colour brown are generated from the peptide **GLTGPIGPPGPAGAPGDK** with *m/z* 795.90975

**Chapter 3: FT-ICR-MS Approach to
monitor asparagine deamidation in
collagen from archaeological material**

3.1. INTRODUCTION

For the last 40 years, the preservation of ancient biomolecules in fossil materials has been a topic of extensive debate and controversy. Several studies have reported the detection of biomolecules isolated from dinosaur tissues and bones. The first report of this type was published in 1972.¹ Since then, extensive research has focused on the recovery of molecules such as DNA from fossils.²⁻⁵ However, the conclusions of many fossil-derived DNA reports have been challenged due to the presence of contaminants.⁶⁻⁸ Conversely, the conclusions of fossil-derived protein analyses are considered to be more reliable.⁹ For example, collagen has been isolated from different fossil types such as bird, mammal, and reptile bones, including dinosaurs;¹⁰⁻¹² in the latter case the presence of intact collagen has been heavily criticized, partly based on the presence of intact Asn residues that suggest the material might have originated from intra-lab contamination.

Due to the triple helical structure and intimate association with bone mineral, collagen can persist with extraordinary longevity in preserved animal remains.^{13, 14} Asp racemisation in collagen has been explored as a possible method of archeological dating.¹⁵ However the higher order structure of collagen minimises succinimide formation and concomitant Asx (Asp & Asn) racemisation,¹⁶⁻¹⁸ thus limiting its utility for amino acid dating analyses.

Recent reports have shown that certain protein modifications such as deamidation are elevated in collagen from older mammal bones,^{13, 19, 20} and these observations could potentially provide the basis for developing the measure of this post-translational

modification as a means of assessing the age of archaeological materials as an alternative to amino acid racemisation.

The mass defect method has been previously used to determine the extent of deamidation of collagen peptides and other proteins. In this chapter, a methodology is reported to measure the extent of deamidation, using the mass defect method, in tryptic digested peptides of collagen I extracted from archaeological bones ($\approx 1,000$ years old), using FT-ICR-MS. Further, deamidation rates and half-lives have been measured, to assess the feasibility of using Asn deamidation as a relative age dating technique in collagen and whether this is adversely affected by local environmental differences (e.g. temperature, pH, soil chemistry). A preliminary study reported in Chapter 2²¹ provided the characterization of three peptides that were recommended as potential indicators of deamidation in bovine collagen. However, the presence of these indicators was not tested in collagen extracted from bones; it was only tested in standard collagen.

Therefore, the aim of this study was to test that the previously identified indicators reported in Chapter 2 could be used successfully to identify the presence of bovine collagen from archeological samples. A further aim is to provide information on the potential value of archeological dating that could be provided from analysis of the extent of deamidation in ancient collagen by applying the methodology reported in Chapter 2 to real samples.

3.2. EXPERIMENTAL METHODOLOGY

3.2.1. Materials

A cattle bone sample, excavated from the Anglo-Viking levels at the Coppergate archaeological site (York) was used for this study. The sample was provided by professors Terry O'Connor and Matthew Collins from the University of York. These bones were recovered in 1973 and have been subsequently stored above ground at ambient temperature. Sequencing-grade trypsin, ammonium acetate ($\text{CH}_3\text{COONH}_4$), and ammonium bicarbonate (NH_4HCO_3) were purchased from Sigma Aldrich, Gillingham, UK. HPLC-grade methanol and formic acid (HCOOH) were obtained from Fisher Scientific, Leicestershire, UK.

3.2.2. Collagen extraction from bovine bones obtained from a Coventry (UK) butcher.

Bones were frozen to $-80\text{ }^\circ\text{C}$ and smashed with a hammer to small pieces (see Figs. 3.1 and 3.2). Since these modern bones contained high levels of fat, a cleaning procedure was necessary. This was achieved by soaking the samples in concentrated HCl in the freezer for an hour, which removes surface proteins and mineralization. After rinsing with purified water to reach neutral pH, a mixture of chloroform:methanol ($\text{CH}_2\text{Cl}_2:\text{MeOH}$) 87:13 v/v was added, while constantly stirring, for another hour. Bones were rinsed with MeOH and kept in the $-80\text{ }^\circ\text{C}$ freezer overnight. After the cleaning procedure was completed a coffee grinder was used to grind the bones to powder (see Fig. 3.3). Triplicates of 14 mg of this powder were weighed and demineralised with 1 mL of HCl (0.6 M) at $4\text{ }^\circ\text{C}$ for 4 hours. From this extraction, the non-soluble, collagen-containing fraction was rinsed 3 times with pure water until neutral pH was reached.



FIGURE 3.1: Pieces of bovine bone obtained from a Coventry UK butcher.



FIGURE 3.2: Size of the bones after hammering.



FIGURE 3.3: Resulting bone powder obtained from the grinding process.

3.2.3. Digestion of collagen

Prior to digestion, the non-soluble fraction was denatured in 1 mL of 50 mM ammonium bicarbonate following a previously described methodology in Chapter 2.²¹ Adjustments to the enzyme:protein ratio had to be optimized. The supernatant (700 μ L) was collected and 24 μ L of trypsin solution (2 μ g/ μ L) was added using an approximate ratio of 1:200 (w/w) enzyme:protein. To minimise deamidation during tryptic digestion the mixture was allowed to react for 4 hours at 37°C,²² the reaction was halted by adding 5 % (v/v) of formic acid, and the samples were dried in a SpeedVac system (Savant SPD121P Fisher Scientific, Leicestershire, UK) to evaporate volatiles at 35 °C, following storage at -80 °C until time of analysis (\approx 1 day). The samples were reconstituted into a mixture of 1:1 MeOH:H₂O, and diluted into a mixture of 49.5:49.5:1 MeOH:H₂O:formic acid to a final approximate concentration of 2.5 μ M.

3.2.4. Collagen extraction from archeological bones

In order to extract the collagen from archaeological bones the procedure explained above was followed, with the difference that no fat was present in the samples. Thus no previous cleaning procedure was needed.

3.2.5. FT-ICR-MS measurements

The experiments were carried out on a Bruker 12T solariX FT-ICR-MS (Bruker Daltonics). In order to identify deamidation sites, samples of digested collagen were electrosprayed at approximately 2.5 μ M concentration in 49.5:49.5:1 MeOH:H₂O:formic acid. An average of 200 scans was taken.

For CAD experiments,²³ the parent ions were isolated using a quadrupole and were fragmented in the collision cell (q, typical collision energies ~10-30 eV), and then transmitted into the ICR cell for detection. For ECD experiments, the parent ions were isolated using a quadrupole and accumulated in a hexapole collision cell (at injection energies <5 eV) for 1-10 s. After being transferred and trapped in the ICR cell, ions were irradiated with 1.5 eV electrons from a heated hollow cathode (1.7 A) for 20 to 90 ms.²⁴ The mass spectrometer was initially externally calibrated using collagen I from Sigma Aldrich. After confident assignments using mass accuracy and isotopic distributions, the spectra were internally calibrated using peptides present in the samples (See section 3.6 of this chapter for more information about mass accuracy).

3.3. RESULTS AND DISCUSSION

3.3.1. Characterization of marker peptides from bovine collagen

The spectrum of modern bone shown in Fig. 3.4a, shows the presence of the three peptides indicators of deamidation corresponding to bovine collagen assigned in the previous study.²¹

Curiously collagen from contemporary bones is more difficult to purify than from ancient bones. This is mainly due to the high fat content. Thus, collagen was extracted and analyzed from ancient bovine bones. These bones originated from the excavation performed in Coppergate, York, in 1973 by the York Archaeological Trust.

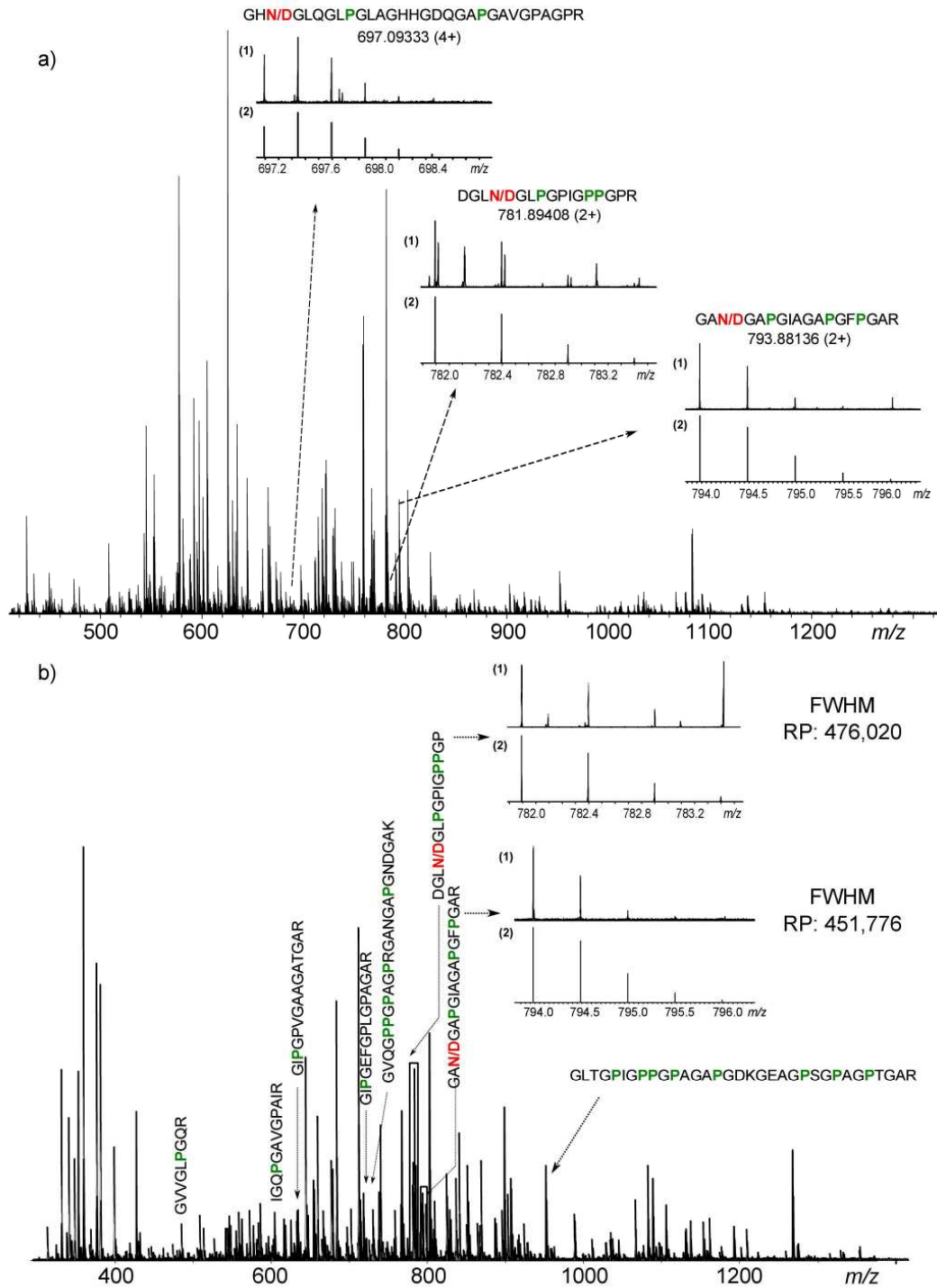


FIGURE 3.4: a) Mass spectrum of tryptic digest of collagen extracted from contemporary cow bones. **N/D** represents deamidation site. **P** represents Hyp. The inserts show three possible deamidated peptides corresponding to m/z [697.09326]⁴⁺, m/z [781.89408]²⁺, and m/z [793.88133]²⁺. b) Tryptic digest collagen extracted from 1,050 year-old bone. Inserts (1) and (2) show the experimental and theoretical isotopic distributions respectively.

Dating back to 960 AD (Viking period) these bones were well preserved due to the moist peaty layers in the region, where flooding hinders influx of oxygen from the atmosphere, thus reducing rates of degradation.²⁵ Such anaerobic preservation allowed us to extract the collagen from the bone and analyse the tryptic peptides using FT-ICR-MS. Figure 3.4b shows the spectrum of the digested collagen extracted from the archaeological samples.

Twelve peptides, including peptides with N/D, were successfully assigned to tryptic peptides predicted from the published bovine collagen sequence (P02453 and P02465 Uniprot database).²⁶ This was achieved using mass accuracy, isotopic distributions, and in some cases fragmentation using CAD and ECD (see Table 3.1 and tables in section 3.6 of this chapter). These sequences were compared against known sequences for human collagen, in order to test for any sample handling contamination issues. The peptide corresponding to 970-983 (α 1) was found to be a marker for bovine collagen, rather than that from a human source.

TABLE 3.1: Assignments of peptides in archeological bone material

Length coll-I (chain)	Identified peptides	MW	[M+2H] ²⁺	[M+3H] ³⁺	[M+4H] ⁴⁺	Exp.	Error ppm
958-966 (α 1)	GVVGL P GQR	897.50319	449.75888	300.17501		449.75884	0.08
970-983 (α 1)	IGQ P GAVGPAGIR	1343.63571	604.84092	403.56304		604.84091	0.02
326-340 (α 2)	G P GPVGAAGATGAR	1266.66800	634.34128	423.22995		634.34122	0.09
572-586 (α 2)	G P GEFGLPGPAGAR (2oxi)	1426.72043	714.36750	476.58076		714.36734	0.22
685-708 (α 1)	GVQGP P PAGPRGANGAPGNDGAK (4oxi)	2162.01002	1082.01229	721.67729		721.67757	-0.39
1152-1167 (α 1)	DGL N GL P GP I G P P G R	1560.78958	781.40207	521.27047		781.40189	0.23
1152-1167 (α 1)	DGL D GL P GP I G P P G R	1561.77359	781.89408	521.59848		781.89408	0.00
397-414 (α 1)	G A N G A P G I A G A P G F P G A R	1584.76442	793.38949	529.26209		793.38933	0.20
397-414 (α 1)	G A D G A P G I A G A P G F P G A R	1585.74843	793.88150	529.59009		793.88133	0.21
762-795(α 1)	GLT G P I G P P G P A G A P G D K GE A G P S G P A G P T G A R (2oxi)	2852.40520	1427.20988	951.80901		951.80833	0.72
1022-1051 (α 2)	G H N GL Q GL P GL A GH H GD Q G A P G AV G P A G P R	2783.35994	1392.68725	928.79393	696.84726	696.84722	0.06
1022-1051 (α 2)	G H D GL Q GL P GL A GH H GD Q G A P G AV G P A G P R	2784.34395	1393.17926	929.12193	697.09326	697.09377	-0.73

P: Hydroxyproline

N: Deamidated

3.3.2. Identification of deamidation sites in collagen extracted from archeological bones using FT-ICR-MS

In order to assess deamidation, a shift of +0.984 Da was monitored in the spectrum of the collagen tryptic digest (Fig. 3.4.b). Analysis of the spectra showed two potential deamidating peptides, and these two observed peptides corresponded well with the expected masses (<1 ppm). Therefore, these 2 peptides were chosen for further deamidation studies, [coll_I (α 1) 1152-1167]²⁺, and [coll_I (α 1) 397-414]²⁺. Fig. 3.5 and 3.6 show the ECD spectrum for each of these peptides.

The use of infrared multiphoton dissociation (IRMPD) along with ECD assisted sequencing of each peptide. The advantages of performing IRMPD along with ECD rest on the ability of photon irradiation to produce extensive peptide backbone fragmentation, including sites next to Pro, while ECD provides complementary sequence information. IR-photon irradiation also increases the orbit of the ion packet improving interaction with the electron beam.^{27, 28} Consequently, more efficient fragmentation is achieved. In this particular case, the use of both techniques facilitated the localisation of the Hyp, which are a marker for the presence of collagen.

Deamidation sites were also easily located in each peptide. CAD spectra for each of the selected peptides are shown in Fig. 3.7 and 3.8. Isotopic distributions and high mass accuracy were used to increase the confidence in assignments of each peptide and its fragments. The complete list of masses and assignments are included in the section 3.6 of this chapter.

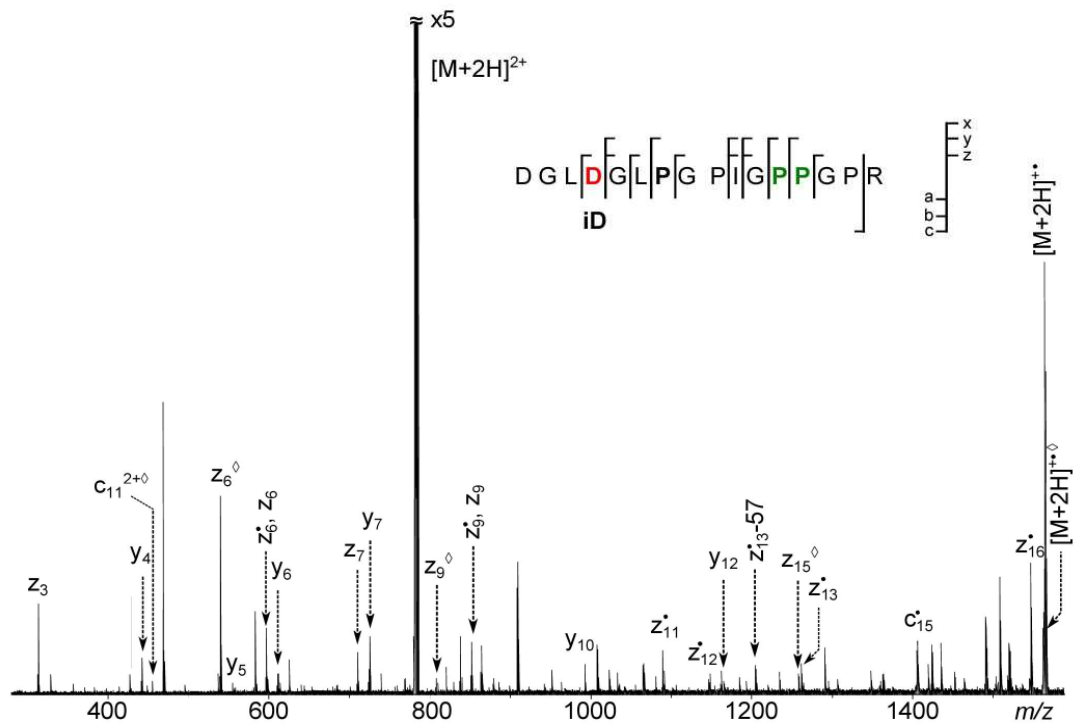


FIGURE 3.5: ECD-IRMPD of tryptic digest peptide [coll_I (α 1) 1152-1167] $^{2+}$ ion at m/z 781.89408. The deamidation is located in position 4 from the N-terminus (**D**). P highlighted in green represent the modification of hydroxylation. Ions labelled with superscript (\diamond) originate from another peptide present in the digestion mixture at m/z 780.90998.

The ECD of the tryptic digest peptide [coll_I (α 1) 1152-1167] $^{2+}$ ion at m/z 781.89408 (Fig. 3.5) shows the region where the IsoAsp diagnostic ion appears (z_{13-57}^{\bullet} , m/z 1204.63209) indicating deamidation at position 13. Due to the density of peaks in the collagen spectrum (\sim 400 peaks in a window of 400-1300 m/z), ions labelled with superscript (\diamond) originate from another peptide present in the digestion mixture with m/z 780.90998, which could not be isolated using the quadrupole, but which could be assigned from the resulting MS/MS spectrum.

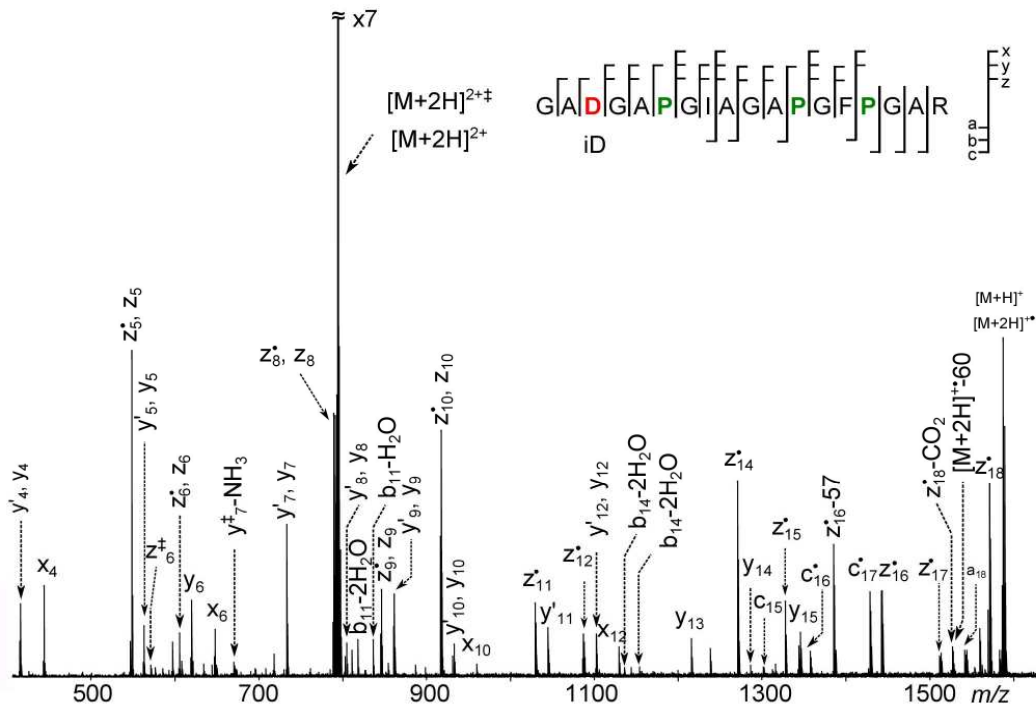


FIGURE 3.6: ECD-IRMPD of tryptic digest peptide [coll_I (α 1) 397-414]²⁺ ion at m/z 793.88150. The deamidation site is located in position 3 from the N-terminus (**D**), in the peptide. P highlighted in green represent the modification of hydroxylation. Ions labelled with superscript (\ddagger) are originated from other peptide present in the digestion mixture with m/z 790.88678.

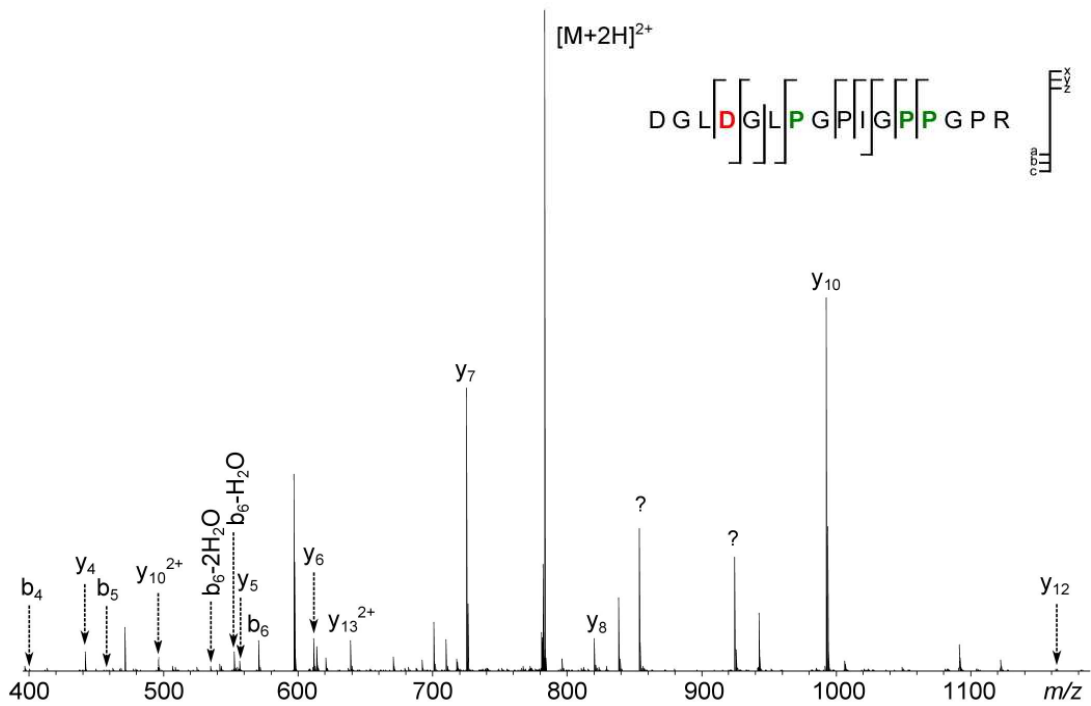


FIGURE 3.7: CAD of tryptic digest peptide [coll_I(α 1) -1152-1167]²⁺ ion at m/z 781.89408. Red colour represent the deamidation site, green indicates the presence of Hyp.

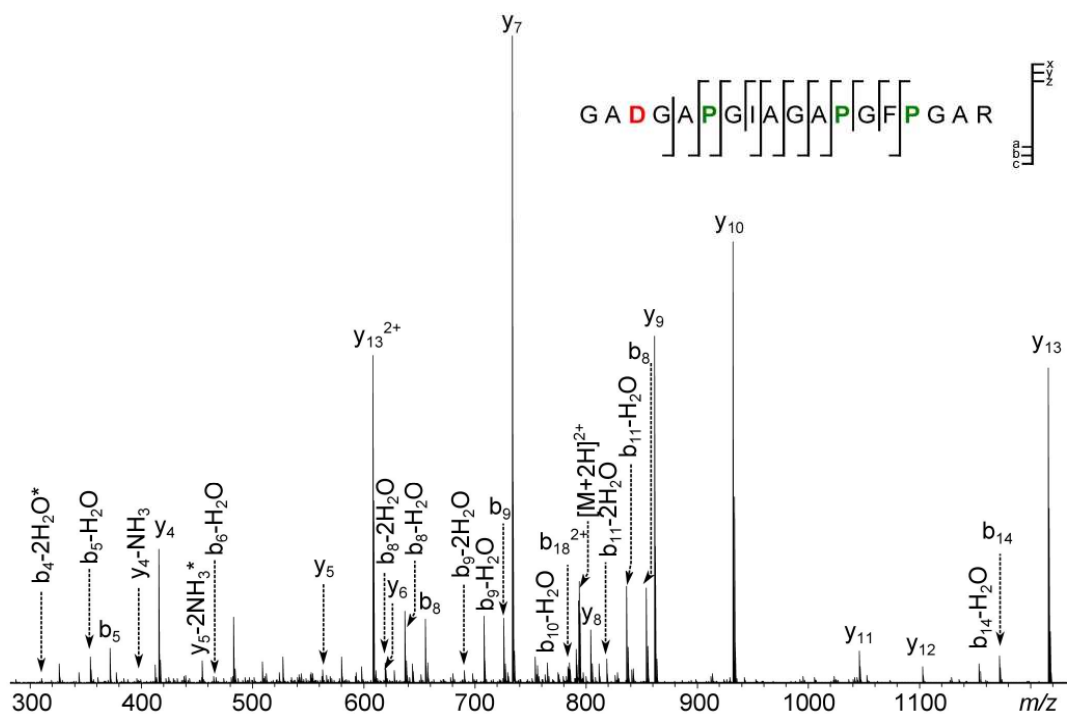


FIGURE 3.8: CAD of tryptic digest peptide [coll_I ($\alpha 1$) 397-414, coll_I($\alpha 1$)] $^{2+}$ ion at m/z 793.88150. Red colour represent the deamidation site, green indicates the presence of Hyp.

The ECD of the other tryptic digest peptide which showed evidence of deamidation [coll_I ($\alpha 1$) 397-414] $^{2+}$ ion at m/z 793.88150 (Fig. 3.6), shows the region where the IsoAsp diagnostic ion (z^{\bullet}_{16-57} , m/z 1385.68082) appears, corresponding to position 3 (from the N-terminus) in the peptide. (Ions labelled with superscript (\ddagger) originated from yet another peptide present in the complex digestion mixture (790.88678 m/z)). This peptide was assigned due to the presence of additional fragments in the ECD spectrum (Fig. 3.6).

In the ECD spectrum of the chosen collagen peptides, the IsoAsp diagnostic peaks ($z^{\bullet}-57$) corresponding to the deamidated Asn were observed in the spectrum, Fig. 3.5 and 3.6. The Asp diagnostic fragment ion also appeared in the spectrum shown in Fig. 3.6. A possible explanation for the lack of the Asp diagnostic ion in the spectrum shown in Fig. 3.5 may be interference with high intensity peaks located in

the $(M+nH)^{(n-1)+\bullet}$ -60 region where Asp ions should appear; this effect has been previously observed by Cournoyer et al.²⁹

3.3.3. Determination of the reaction rates for deamidation reaction

In order to study the deamidation of the peptides present in the protein, deamidation was deliberately induced by heating the sample as previously explained in the experimental section. Consequently, a deamidation decay curve was produced and the extent of deamidation is back calculated by extrapolating the curve to time zero. A first order reaction was assumed, as explained in previous chapters.

It is important to note that, by inducing deamidation in the denaturation step and assuming no deamidation during acidic treatment and protein digestion,²² it is possible to create a curve and measure the initial extent of deamidation by extrapolating back to zero. Fig. 3.9 shows the exponential decay curves for the peptides in study. This assumption was later tested by monitoring the deamidation during acidic treatment and protein digestion.

Deamidation half lives were calculated to be (5639 ± 823) and (4127 ± 981) s for coll_I ($\alpha 1$) 1152-1167 and coll_I ($\alpha 1$) 397-414 respectively at 62 °C, as shown in table 3.2. The extent of deamidation (expressed as the remaining non-deamidated fraction) of these collagen samples was given by the “y” axis intercept of the exponential curves $N_t/(N_t+D_t)$, and was found to be (0.93 ± 0.09) , and (0.66 ± 0.03) for coll_I ($\alpha 1$) 1152-1167, and coll_I ($\alpha 1$) 397-414 respectively. This means that the coll_I ($\alpha 1$) 1152-1167 peptide in the sample, prior to the deliberate deamidation in this experiment,

was 7 % deamidated and the coll_I (α 1) 397-414 peptide in the sample was 34% deamidated.

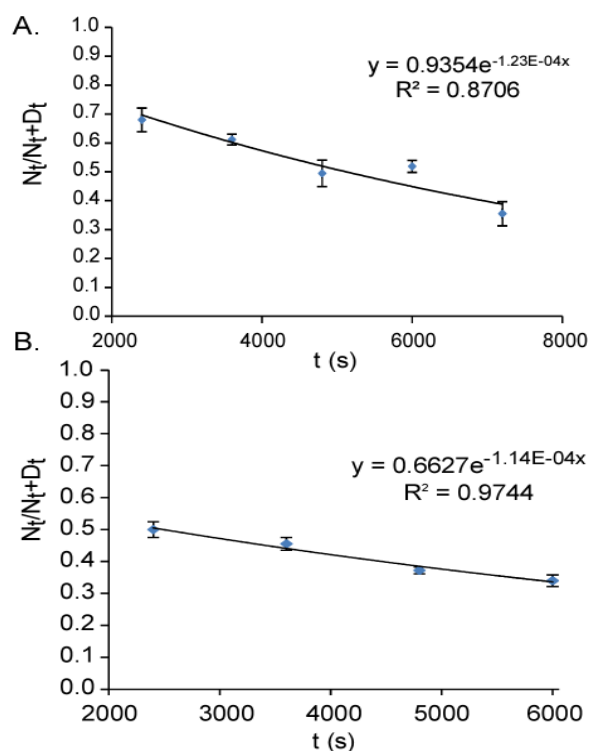


FIGURE 3.9: Exponential decay curves of the non-deamidated peptide at 62 °C. A. [coll_I (α 1) - 1152-1167]²⁺ ion at m/z 781.89408, and B. [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150. The exponential coefficients represent the rate constants of deamidation reaction for each peptide. The intercept represents the extent of deamidation in each case.

TABLE 3.2: Deamidation rates and grade of deamidation of collagen peptides from young collagen and ancient collagen at 62°C

Peptide	coll_I(α 1) -1152-1167 (62°C)		coll_I(α 1) 397-414 (62°C)	
	Young	Ancient	Young	Ancient
$k (\times 10^4 \text{s}^{-1}) \pm \text{SD}$	1.23 ± 0.10	1.23 ± 0.17	2.69 ± 0.28	1.13 ± 0.13
$t^{1/2} (\text{s}) \pm \text{SD}$	5629 ± 463	5639 ± 823	2575 ± 269	4127 ± 981
$(N_t/(N_t+D_t))$	1.05 ± 0.08	0.93 ± 0.09	1.18 ± 0.12	0.66 ± 0.03
% D (t=0)	≈ 3	≈ 7	≈ 6	≈ 34

An elevated extent of deamidation in the ancient sample would be expected for the peptide [coll_I(α 1)-1152-1167] corresponding to a sequence of DGLNGLPGPIGPPGPR. As shown in table 3.2, the extent of deamidation for this peptide in the ancient sample is $\approx 7\%$. The difference in the extent of deamidation for this peptide when compared to the peptide [coll_I(α 1) 397-414], with a sequence of GANGLPGIAGAPGFPGAR, could be explained by structure impediment of deamidation due to the presence of a Leu next to Asn in the former peptide. The presence of a small amino acid such as alanine (Ala) would facilitate the deamidation in the latter peptide showing an extent of deamidation of $\approx 34\%$, which correspond well with previous studies.³⁰

3.3.4. Estimation of the Activation Energy

The rates of Asn deamidation are influenced by several factors including protein primary sequence, pH, and temperature.³¹ It has been previously reported that the deamidation rate increases proportionally with temperature. For example in the case of the peptide (VYPNGA) the rate increases 2.4 fold per 10 °C in basic buffer.³² It was shown that some of deamidation inhibition in protein is essentially stable at normal temperature. Unless a specific structural impediment to deamidation is greatly interrupted as a function of temperature, the temperature dependence of that inhibition would be expected to be diminished.

Since the future desire is to use the extent of deamidation as a measure of the relative age of ancient materials, it is necessary to account for the temperature dependence of the reaction. Thus the kinetic molecular theory can be used, which states that the higher the temperature the higher the kinetic energy of the molecules. An increase in

temperature means that molecules have more chance to collide each other and react. In order to find the necessary activation energy for deamidation, the Arrhenius approximation was used. The rate constant dependence of temperature was monitored by using the following equation:

$$k = Ae^{-E_a/RT} \quad (3.1)$$

Where A is a factor which accounts for the frequency of collision and the probability that the molecules are in correct orientation, E_a is activation energy, R is the gas constant (8.314 J/molK) and T is temperature in Kelvin. By applying natural log to equation 3.1, equation 3.2 is obtained and when plotted gives a straight line.

$$\ln k = \ln A - E_a/RT \quad (3.2)$$

The activation energy for deamidation was determined through the use of experimental data obtained at the temperatures of 50 °C and 62 °C for the peptide coll_I(α 1) 397-414 present in collagen, which so far is the most recognizable marker throughout a range of different bovine bone samples. The activation energy was calculated by using equation 3.3.

$$\ln(k_1/k_2) = E_a/RT(1/T_1 - 1/T_2) \quad (3.3)$$

Where k_1 and k_2 are deamidation rate constants at 50 °C and 62 °C respectively. These values are directly extracted from the exponential decays of the non-deamidated species (Fig. 3.10). By using equation 3.3 a value for activation energy was calculated to be 49 kJ/mol,³³ which is significantly less than previous reports for deamidation of collagen (173 kJ/mol).¹⁹ A possible explanation for the difference could be that in the present study the E_a was calculated for one particular peptide, and not for the entire population of deamidated peptides.

This value of activation energy can be used to calculate the rate constants at any desired temperature, which will enable a more reliable calculation of the sample's age by adjusting the deamidation rates at a more realistic temperature (e.g. 10 °C).

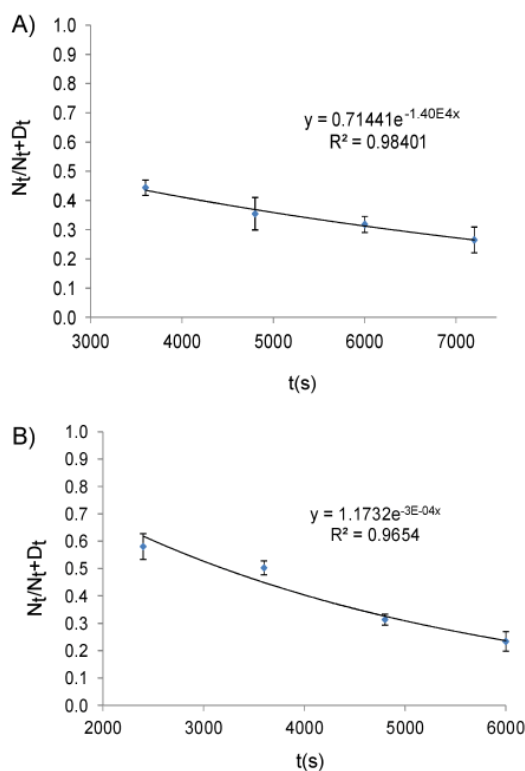


FIGURE 3.10: Exponential decay curves of the non-deamidated peptide [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 at A) 50 °C and B) 62 °C. The exponential coefficients represent the rate constants of deamidation reaction. The intercept represents the extent of deamidation in each case.

3.3.5. Effect of digestion treatment on deamidation

It was hypothesised that the protein digestion step in the experimental method could have artificially induced extra deamidation which would affect the result. Therefore, an experiment was performed to test the effect of protein digestion on deamidation, by isolating the digestion treatment. This was achieved by adjusting the length of time given for digestion from 2 hours up to 24 hours, at 37 °C, and keeping the rest of the parameters in the process constant. Fig. 3.11. A-B shows that the extent of measured deamidation can indeed be affected by the conditions used in the digestion

step. These results show that the time window for this part of the procedure should be kept to less than 4 hours, where the effect is kept to a minimum. It has been previously reported that 4 hours is enough time for tryptic digestion.^{13,22}

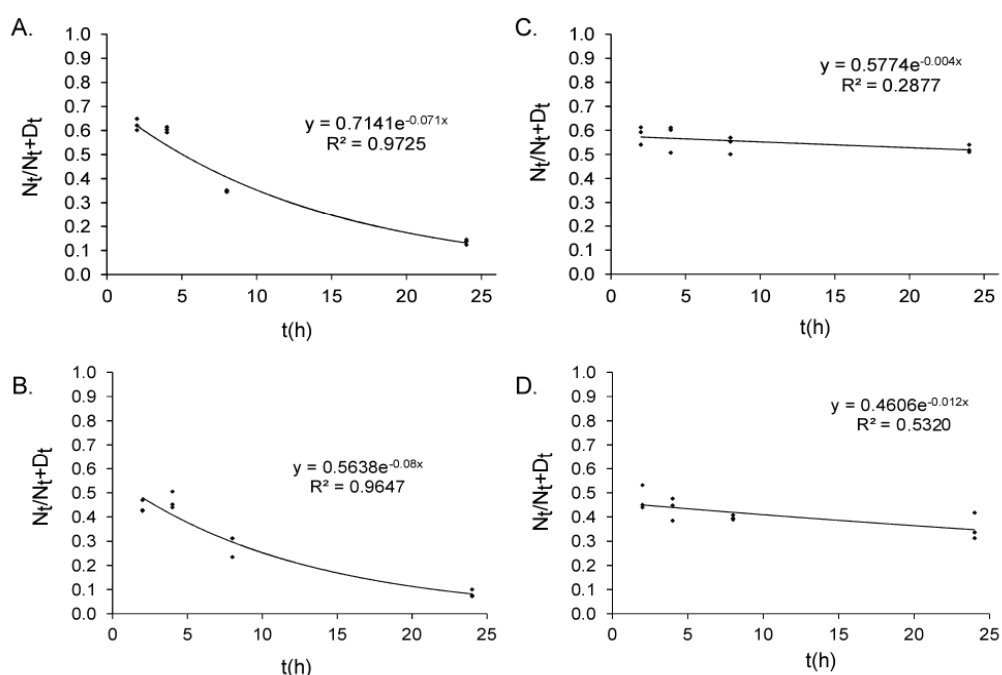


FIGURE 3.11: Exponential decay curves of the non-deamidated peptide at 62 °C. A, B. Effect of digestion conditions on [coll_I (α 1) -1152-1167]²⁺ ion at m/z 781.89408, and [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 respectively. C, and D. Effect of acidic treatment on [coll_I (α 1) -1152-1167]²⁺ ion at 781.89408 m/z , and [coll_I (α 1) 397-414, coll_I (α 1)]²⁺ ion at m/z 793.88150 respectively.

3.3.6. Effect of acidic treatment on deamidation

A similar experiment was performed to test the effect of acidic treatment on the measured extent of deamidation. The isolation of the acidic treatment step was carried out at times ranging from 2 hours to 24 hours. Fig. 3.10. C-D shows that acidic treatment had little effect on deamidation over the course of 24 hours. However, the time window selected for the denaturation experiment was 4 hours,

which has been previously reported to be sufficient time to extract the inorganic material from the bone matrix.

3.3.7. Effect of pH and buffers

Bada and Shou (1980)³⁴ reported a study of racemisation rates of Asp in modern bone fragments. In this experiment they proved that the rates were independent of pH in the range 3-9 at a temperature of 96 °C. They suggest that the bone matrix itself acts as a buffer that stabilises the pH of the system. Thus they believe that environmental acid pH is not important with respect to archaeological bones.

3.4. CONCLUSIONS

A total of eight tryptic peptides from bovine collagen were successfully assigned in a Viking period bone sample (~1,000 years old) using FT-ICR-MS. Two of the peptides were chosen as probes to monitor the extent of deamidation.

ECD combined with IRMPD was shown to be a good fragmentation technique to sequence the peptides in the study and identify the deamidation sites. The use of FT-ICR-MS facilitated the differentiation of Hyp (113.04768 Da) from Leu/Ile (113.08406 Da). Ultrahigh resolution was needed to monitor the intensity ratios of the non-deamidated and the deamidated peptides.

The first order kinetic mechanism fitted the deamidation process for each peptide and the extent of deamidation at time zero was successfully determined accordingly. The method can be used to determine the extent of deamidation of the peptides present in samples of archaeological interest (in this case bovine bone) before sample preparation, which can be advantageous when deamidation is artificially induced. The study proved that deamidation was elevated in one of the peptides in the ancient sample. However this deamidation was subject to structural composition of the peptide.

It is proposed that the deamidated peptides identified in this study could be potential age markers in these types of samples. However their presence needs to be further investigated in older samples (> 1,000 years old).

3.5. REFERENCES

1. Nowicki, Z., Sarnecka-Keller, M., Pawlicki, R., Kubiak, H., and Korbel, A. Morphological and Chemical Analysis of the Structure of the Fossil Dinosaur Bone. *Anatomischer Anzeiger*. **1972**, 132, 10-23.
2. Paabo, S., Higuchi, R.G., and Wilson, A.C. Ancient DNA and the Polymerase Chain-Reaction - the Emerging Field of Molecular Archaeology. *Journal of Biological Chemistry*. **1989**, 264, 9709-9712.
3. Vickers-Rich, P., Trusler, P., Rowley, M.J., Cooper, A., Chambers, G.K., Bock, W.J., Millener, P.R., Worthy, T.H., and Yaldwyn, J.C. Morphology, Myology, Collagen and DNA of a Mummified Upland Moa, *Megalapteryx Didinus* (Aves: Dinornithiformes) from New Zealand. *Tuhinga*. **1995**, 0, 1-26.
4. Woodward, S.R., Weyand, N.J., and Bunnell, M. DNA-Sequence from Cretaceous Period Bone Fragments. *Science*. **1994**, 266, 1229-1232.
5. Pawlicki, R. Histochemical-Demonstration of DNA in Osteocytes from Dinosaur Bones. *Folia Histochemica Et Cytobiologica*. **1995**, 33, 183-186.
6. Asara, M. Protein Sequences from Mastodon and Tyrannosaurus Rex Revealed by Mass Spectrometry *Science*. **2007**, 316, 1698-1698.
7. Hedges, S.B. and Schweitzer, M.H. Detecting Dinosaur DNA. *Science*. **1995**, 268, 1191-1192.
8. Zischler, H., Hoss, M., Handt, O., Vonhaeseler, A., Vanderkuyl, A.C., Goudsmit, J., and Paabo, S. Detecting Dinosaur DNA. *Science*. **1995**, 268, 1192-1193.
9. Schweitzer, M.H. The Future of Molecular Paleontology. *Palaeontologia Electronica*. **2003**, 5, 1-11 (<http://palaeo-electronica.org>).
10. Buckley, M., Collins, M., and Thomas-Oates, J. A Method of Isolating the Collagen (I) Alpha 2 Chain Carboxyterminal Peptide for Species Identification in Bone Fragments. *Analytical Biochemistry*. **2008**, 374, 325-334.
11. Buckley, M., Collins, M., Thomas-Oates, J., and Wilson, J.C. Species Identification by Analysis of Bone Collagen Using Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2009**, 23, 3843-3854.
12. Gurley, L.R., Valdez, J.G., Spall, W.D., Smith, B.F., and Gillette, D.D. Proteins in the Fossil Bone of the Dinosaur, *Seismosaurus*. *Journal of Protein Chemistry*. **1991**, 10, 75-90.
13. Leo, G., Andreotti, A., Marino, G., Pucci, P., Colombini M. P., and Birolo, L. Deamidation at Asparagine and Glutamine as a Major Modification Upon Deterioration/Aging of Proteinaceous Binders in Mural Paintings. *Analytical Chemistry*. **2011**, 83, 2056-2064.
14. Buckley, M., Larkin, N., and Collins, M. Mammoth and Mastodon Collagen Sequences; Survival and Utility. *Geochimica et Cosmochimica Acta*. **2011**, 75, 2007-2016.
15. Eglinton, G. and Collins, M.J. Predicting Protein Decomposition: The Case of Aspartic-Acid Racemization Kinetics - Discussion. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*. **1999**, 354, 64-64.
16. Julg, A., Lafont, R., and Perinet, G. Mechanisms of Collagen Racemization in Fossil Bones. Application to Absolute Dating. *Quaternary Science Reviews*. **1987**, 6, 25-28.

17. Van Duin, A.C.T. and Collins, M.J. The Effects of Conformational Constraints on Aspartic Acid Racemization. *Organic Geochemistry*. **1998**, *29*, 1227-1232.
18. Collins, M.J., Penkman, K.E.H., Rohland, N., Shapiro, B., Dobberstein, R.C., Ritz-Timme, S., and Hofreiter, M. Is Amino Acid Racemization a Useful Tool for Screening for Ancient DNA in Bone? *Proceedings of the Royal Society B-Biological Sciences*. **2009**, *276*, 2971-2977.
19. Buckley, M., Walker, A., Ho, S.Y.W., Yang, Y., Smith, C., Ashton, P., Thomas-Oates, J., Cappellini, E., Koon, H., Penkman, K., Elsworth, B., Ashford, D., Solazzo, C., Andrews, P., Strahler, J., Shapiro, B., Ostrom, P., Gandhi, H., Miller, W., Raney, B., Zylber, M.I., Gilbert, M.T.P., Prigodich, R.V., Ryan, M., Rijdsdijk, K.F., Janoo, A., and Collins, M.J. Comment on "Protein Sequences from Mastodon and Tyrannosaurus Rex Revealed by Mass Spectrometry". *Science*. **2008**, *319*, 5859.
20. Wilson, J., Van Doorn, N.L., and Collins, M.J. Assessing the Extent of Bone Degradation Using Glutamine Deamidation in Collagen. *Analytical Chemistry*. **2012**, *84*, 9041-9048.
21. Perez Hurtado, P. and O'connor, P.B. Deamidation of Collagen. *Analytical Chemistry*. **2012**, *84*, 3017-3025.
22. Li, X., Cournoyer, J.J., Lin, C., and O'connor, P.B. Use of O-18 Labels to Monitor Deamidation During Protein and Peptide Sample Processing. *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 855-864.
23. Jennings, K.R. Collision-Induced Decompositions of Aromatic Molecular Ions. *International Journal of Mass Spectrometry and Ion Physics*. **1968**, *1*, 227-235.
24. Tsybin, Y., Quinn, J., Tsybin, O., Hendrickson, C., and Marshall, A. Electron Capture Dissociation Implementation Progress in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 762-771.
25. Keddy, P.A., in *Wetland Ecology: Principles and Conservation*, ed. P.A. Keddy. Cambridge University Press: UK, 2010, p 497.
26. Uniprot, C. Reorganizing the Protein Space at the Universal Protein Resource (Uniprot). *Nucleic Acids Research*. **2012**, *40*, D71-D75.
27. Hofstadler, S.A., Sannes-Lowery, K.A., and Griffey, R.H. M/Z-Selective Infrared Multiphoton Dissociation in a Penning Trap Using Sidekick Trapping and an Rf-Tickle Pulse. *Rapid Communications in Mass Spectrometry*. **2001**, *15*, 945-951.
28. Leymarie, N., Berg, E.A., McComb, M.E., O'connor, P.B., Grogan, J., Oppenheim, F.G., and Costello, C.E. Tandem Mass Spectrometry for Structural Characterization of Proline-Rich Proteins: Application to Salivary Prp-3. *Analytical Chemistry*. **2002**, *74*, 4124-4132.
29. Cournoyer, J.J., Lin, C., and O'connor, P.B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. *Analytical Chemistry*. **2006**, *78*, 1264-1271.
30. Robinson, N.E. and Robinson, A.B. Molecular Clocks. *Proceedings of the National Academy of Sciences of the United States of America*. **2001**, *98*, 944-949.
31. Robinson, N.E. and Robinson, A.B., *Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*. Althouse Press: Cave Junction, OR, **2004**.

32. Geiger, T. and Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides - Succinimide-Linked Reactions That Contribute to Protein-Degradation. *Journal of Biological Chemistry*. **1987**, 262, 785-794.
33. Mbugi, A., *Ft-Icr –Ms Approach to Determine the Temperature Dependence in Collagen Deamidation*, in *Chemistry2012*, University of Warwick: UK. p. 46.
34. Bada, J.L., Shou, M. Y., in *Biogeochemistry of Amino Acids*, ed. P.E. Hare, Hoering, T. C., King, K. J. Wiley: New York, 1980, pp 235-255.

3.6. Mass accuracy and peptide ions assignments

TABLE 3.3: CAD peak table [DGLNGLPGPIGPPGPR+2H]²⁺ *m/z* 781.89408

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
b ₄	401.16667	401.16667	0.00
y ₄	442.24085	442.24067	0.41
b ₅	458.18813	458.18807	0.13
y ₁₀ ²⁺	496.76162	496.76181	-0.39
b ₆ -2H ₂ O	535.25116	535.25107	0.17
b ₆ -H ₂ O	553.26168	553.26146	0.39
y ₅	555.28853	555.28857	-0.07
b ₆	571.27219	571.27213	0.11
y ₆	612.30999	612.31005	-0.10
y ₁₃ ²⁺	639.32785	639.32836	-0.81
y ₇	725.39405	725.39453	-0.66
y ₈	822.44681	822.44565	1.41
y ₁₀	992.51595	992.51693	-0.99
y ₁₂	1162.62147	1162.62145	0.02

TABLE 3.4: Peak list generated from the ECD spectrum of [DGLNGLPGPIGPPGPR+2H]²⁺ *m/z* 781.89408

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z ₃	314.18227	314.18227	0.00
z ₄	427.22995	427.22965	0.70
y ₄	442.24085	442.24046	0.88
c ₁₁ ^{2+◊}	455.224855	455.22465	0.45
y ₅ -H ₂ O	537.278015	537.27802	-0.01
y ₅	555.28853	555.28832	0.38
z ₆	597.29909	597.29908	0.02
y ₇	725.39405	725.39371	0.47
[M+2H] ^{2+◊}	780.90998	780.90943	0.70
[M+2H] ²⁺	781.89408	781.89443	-0.45
z ₉ [◊]	808.46753	808.46771	-0.22
z ₉ [•]	863.44955	863.44960	-0.06
z ₉	864.45737	864.45729	0.09
y ₁₀	992.51595	992.51693	-0.99
z ₁₁ [•]	1089.58129	1089.58134	-0.05
z ₁₂ [•]	1146.60275	1146.60252	0.20
z ₁₃₋₅₇ [•]	1204.63209	1204.63302	-0.77
z ₁₅ [◊]	1257.68237	1257.68125	0.89
z ₁₃ [•]	1261.62969	1261.63019	-0.40
c ₁₅ [•]	1404.68792	1404.68755	0.26
a ₁₆	1517.78321	1517.78263	0.38
z ₁₈ [◊]	1544.79410	1544.79335	0.49
z ₁₆ [•]	1546.76215	1546.76263	0.31
[M+2H] ^{+◊}	1561.82118	1561.82075	0.28
[M+2H] ^{+•}	1563.78923	1563.79001	-0.50

TABLE 3.5: Peak list generated from CAD spectrum of[GANGAPGIAGAPGFPGAR]²⁺ *m/z* 793.88150

Assignments	Calc. Mass (<i>m/z</i>)	Exp. Mass(<i>m/z</i>)	Error (ppm)
b ₄ -2H ₂ O*	309.11943	309.11920	0.74
b ₅ -H ₂ O	354.140835	354.14077	0.18
b ₅	372.15135	372.15135	0.00
y ₄ -NH ₃	399.19873	399.19870	0.07
y ₄	416.22520	416.22524	-0.10
y ₅ -2NH ₃ *	465.24574	465.24567	0.15
b ₆ -H ₂ O	467.18852	467.18862	-0.22
y ₅	563.29361	563.29370	-0.16
y ₁₃ ²⁺	608.30946	608.30957	-0.19
b ₈ -2H ₂ O	619.28352	619.28360	-0.13
y ₆	620.31507	620.31518	-0.18
b ₈ -H ₂ O	637.29404	637.29408	-0.07
y ₁₄ ²⁺	643.82801	643.82815	-0.22
b ₈	655.30455	655.30469	-0.21
b ₉ -2H ₂ O	690.32063	690.32069	-0.09
b ₉ -H ₂ O	708.33115	708.33124	-0.13
b ₉	726.34166	726.34180	-0.19
y ₁₆ ²⁺	729.85221	729.85232	-0.15
y ₇	733.36275	733.36276	-0.01
b ₁₀ -H ₂ O	765.352605	765.35264	-0.05
b ₁₈ ²⁺	784.87621	784.87642	-0.27
[M+2H]2+	793.88150	793.88158	-0.11
y ₈	804.39986	804.40001	-0.19
b ₁₁ -2H ₂ O	818.37920	818.37929	-0.11
b ₁₁ -H ₂ O	836.389715	836.38980	-0.10
b ₁₁	854.40023	854.40041	-0.21
y ₉	861.42132	861.42142	-0.12
y ₁₀	932.45843	932.45841	0.02
y ₁₁	1045.54249	1045.54233	0.15
y ₁₂	1102.56395	1102.56383	0.11
b ₁₄ -H ₂ O	1153.527265	1153.52629	0.85
b ₁₄	1171.53778	1171.53751	0.23
y ₁₃	1215.61163	1215.61162	0.01

TABLE 3.6: Peak list generated from the ECD spectrum of
 $[\text{GANGAPGIAGAPGFPGAR}]^{2+}$ m/z 793.88150

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
<u>z₄</u>	<u>401.21430</u>	<u>401.21430</u>	<u>0.00</u>
y' ₄	415.21738	415.21741	-0.07
y ₄	416.22520	416.22523	-0.07
x ₄	443.21229	443.21229	0.00
z [•] ₅	547.27489	547.27481	0.15
z ₅	548.28271	548.28269	0.04
y' ₅	562.28579	562.28587	-0.14
y ₅	563.29361	563.29363	-0.04
z [†] ₆	571.31982	571.32000	-0.32
z [•] ₆	604.29635	604.29635	0.00
z ₆	605.30417	605.30414	0.05
y ₁₃ ²⁺	608.30946	608.30944	0.02
b ₈ -2H ₂ O	619.28352	619.28341	0.18
y ₆	620.31507	620.31507	0.00
x ₆	647.30216	647.30209	0.11
y [†] ₇ -NH ₃	670.35193	670.35176	0.25
b ₉ -H ₂ O	708.331145	708.33112	0.04
z [•] ₇	717.34403	717.34419	-0.22
z ₇	718.35185	718.35192	-0.10
y ₁₆ ²⁺	729.85221	729.85249	-0.38
y' ₇	732.35493	732.35488	0.07
y ₇	733.36275	733.36286	-0.15
z [•] ₈	788.38114	788.38110	0.05
z ₈	789.38896	789.38895	0.01
[M+2H] ^{2+†}	790.88678	790.88692	-0.18
[M+2H] ²⁺	793.881495	793.88139	0.13
y' ₈	803.39204	803.39218	-0.17
y ₈	804.39986	804.39986	0.00
b ₁₁ -2H ₂ O	818.37920	818.37920	0.00
b ₁₁ -H ₂ O	836.389715	836.38976	-0.05
z [•] ₉	845.40260	845.40256	0.05
z ₉	846.41042	846.41043	-0.01
b ₁₁	854.40023	854.40046	-0.27
y' ₉	860.4135	860.41354	-0.05
<u>y₉</u>	<u>861.42132</u>	<u>861.42132</u>	<u>0.00</u>

Continued

Assignments	Calc. Mass (m/z)	Exp. Mass(m/z)	Error (ppm)
z^{\bullet}_{10}	916.43971	916.43967	0.04
z_{10}	917.44753	917.44754	-0.01
y_{10}	932.45843	932.45842	0.01
x_{10}	959.44552	959.44529	0.24
z^{\bullet}_{11}	<u>1029.52377</u>	<u>1029.52376</u>	<u>0.01</u>
y'_{11}	1044.53467	1044.53470	-0.03
z^{\bullet}_{12}	1086.54523	1086.54527	-0.04
y'_{12}	1101.55613	1101.55615	-0.02
y_{12}	1102.56395	1102.56413	-0.16
x_{12}	1129.55104	1129.55113	-0.08
$b_{14}-2H_2O$	1135.51675	1135.51566	0.96
$b_{14}-H_2O$	1153.527265	1153.52723	0.03
y_{13}	1215.61163	1215.61167	-0.03
z^{\bullet}_{14}	1270.63002	1270.62990	0.09
y_{14}	1286.64874	1286.64832	0.33
c_{15}	1301.61201	1301.61260	-0.45
z^{\bullet}_{15}	1327.65148	1327.65144	0.03
y_{15}	1343.6702	1343.67078	-0.43
c^{\bullet}_{16}	1357.62565	1357.62596	-0.23
$z^{\bullet}_{16}-57$	1385.68082	1385.68068	0.10
c^{\bullet}_{17}	1428.66276	1428.66237	0.27
z^{\bullet}_{16}	1442.67842	1442.67828	0.10
z^{\bullet}_{17}	<u>1513.71553</u>	<u>1513.71553</u>	<u>0.00</u>
$z^{\bullet}_{18}-CO_2$	1526.74719	1526.74746	-0.18
$[M+2H]^{\bullet+}-60$	1527.74297	1527.74169	0.84
a_{18}	1541.75805	1541.75742	0.41
z^{\bullet}_{18}	1570.73699	1570.73669	0.19
$[M+H]^+$	1586.75571	1586.75553	0.11
$[M+2H]^{\bullet+}$	1587.76407	1587.76360	0.30

TABLE 3.7: Internal calibrants used in the MS mode

Molecular Formulae	Cal. Mass m/z	Exp. Mass m/z	Error (ppm)
$C_{35}H_{57}N_{13}O_{11}$	418.72229	418.72229	0.001
$C_{38}H_{67}N_{13}O_{12}$	449.75888	449.75888	0.005
$C_{48}H_{73}N_{13}O_{16}$	544.77217	544.77216	-0.018
$C_{52}H_{89}N_{17}O_{16}$	604.84092	604.8409	-0.004
$C_{65}H_{105}N_{21}O_{22}$	766.89440	766.8944	0.052
$C_{38}H_{67}N_{13}O_{12}$	898.51047	898.5104	-0.036

Chapter 4: Monitoring deamidation and differentiation of the reaction products in ancient collagen using ECD

4.1. INTRODUCCION

Deamidation of Asn and isomerisation of Asp have been closely linked to long-lived proteins and have been related with the concept of molecular timers of biological events.^{1, 2} Deamidation and Asn isomerisation have also been associated with ageing,^{3, 4} type 2 diabetes mellitus,⁵ eye lens anomalies,⁶ and Alzheimer disease.^{7, 8} The study of these types of modifications brings potentially new alternatives to different treatments of these diseases and also a new approach to relative dating.

Deamidation is very straightforward to determine in mass spectrometry by following a shift of +0.984 Da in the spectrum. Isomerisation is not as simple. Numerous tactics have been developed to successfully differentiate the two isomers using non-MS methods as well as MS-based methods. Most of the MS tactics are based around the separation of the two isomers by liquid-chromatography and the detection of characteristic immonium ions at low m/z .⁹⁻¹¹

4.1.1. Non-MS Methods to Determine IsoAsp residues

IsoAsp Antibody

Some antibodies have been successfully used in the localisation of racemisation and deamidation products.¹²⁻¹⁴ The use of antibodies against IsoAsp residues has focused principally on the study of Alzheimer's disease by the multiple antigen peptide system procedure (MAP).¹⁵ In general, the antibodies produced are specific to the IsoAsp residue.

Proteolytic Digestion

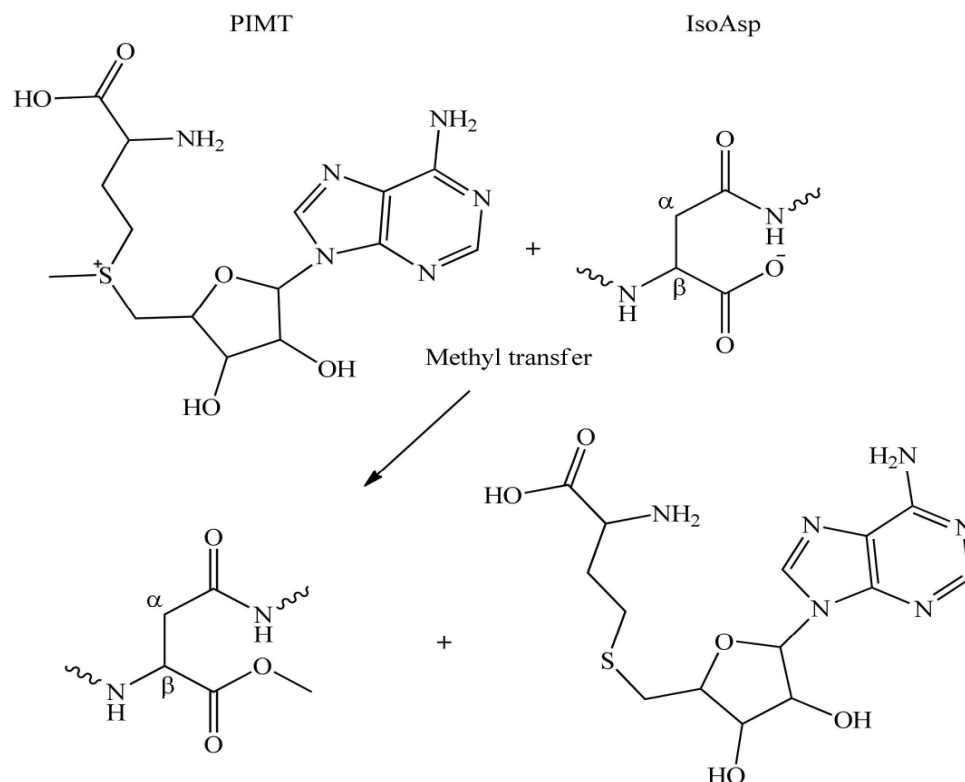
The most common enzyme capable of cleaving at the N-terminus side of L-Asp and not at D-Asp or D/L-IsoAsp residues is endoproteinase Asp-N. It was first used by Kameoka and co-workers., to detect deamidation of Asn and isomerisation of Asp residues in a protein.¹⁶ Although the use of mass spectrometry to detect deamidation is very straightforward (following the shift of +0.948 Da), proteolytic digestion is a simple and quick method to detect Asp isomerisation, when it works successfully.

PIMT enzyme

Protein L-IsoAsp-O-methyltransferase (PIMT) is an enzyme that selectively methylates IsoAsp residues. Consequently, the enzyme can be used to detect deamidation of Asn and isomerisation of Asp. The reaction involves the transfer of methyl groups from S-adenosyl-L-methionine (Adomet) to the alpha side chain carboxyl groups of the IsoAsp residue as seen in Scheme 4.1. IsoAsp detection using PIMT is commonly monitored by HPLC. Recent developments using PIMT enzyme and HPLC are used to detect and measure the abundance of IsoAsp residues. The process is monitored by following the change in UV absorbance of Adomet before and after methylation.¹⁷

Edman Degradation

The Edman reaction involves the nucleophilic attack from the nitrogen of the N-terminus of the protein under investigation to the carbon of the cyanate group of phenylisothiocyanate (Scheme 4.2). This interaction leads to a cyclic intermediate that facilitates hydrolysis of the adjacent amide bond.

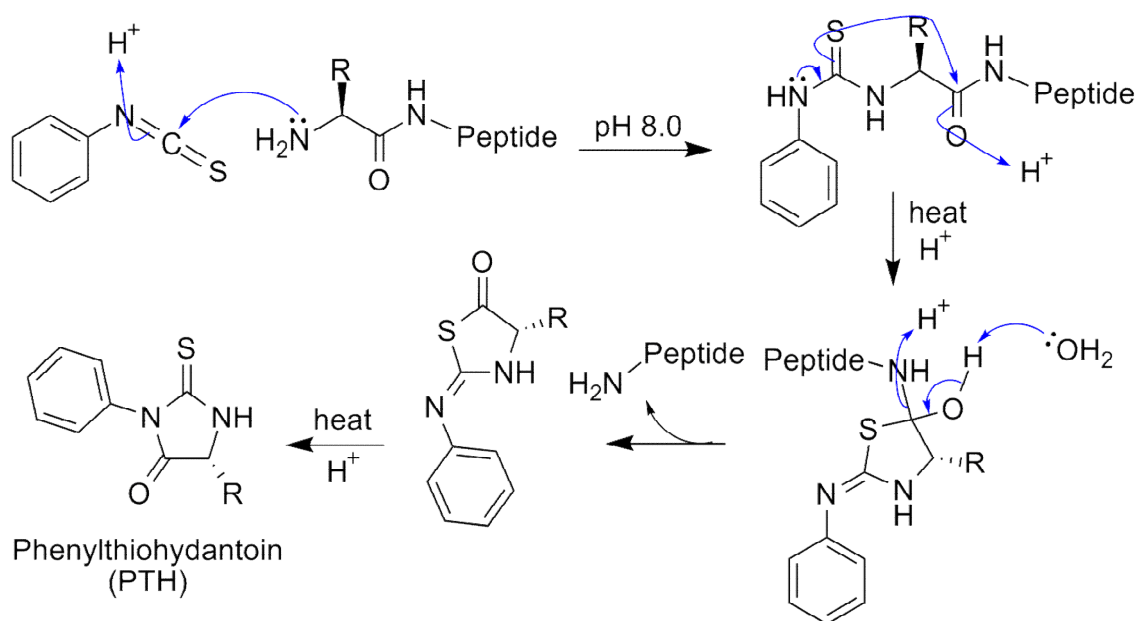


SCHEME 4.1: Methyl transfer into the protein IsoAsp residue.

The effect of this reaction is to specifically remove the N-terminal amino acid from the protein and thereby generate a cleaved derivative of that amino acid and a protein that is shortened by one amino acid and has a new N-terminal. The term degradation suitably describes the effect of repeated applications of the Edman reaction to repeatedly degrade the protein amino acid by amino acid from the N-terminal. This process was found to be particularly useful to locate the presence of an IsoAsp residue in a sequence in 1962.¹⁸

Asp isomerisation can be challenging to detect using mass spectrometry, as there is no mass difference.¹⁹⁻²¹ Early methods were developed to differentiate the two isomeric forms^{9, 10} based on relative fragment ion abundance, but the introduction of new fragmentation techniques in mass spectrometry such as electron capture (ECD)

and electron transfer dissociation (ETD) has improved the ability of researchers to differentiate between Asp and IsoAsp residues in proteins.¹⁹⁻²²



SCHEME 4.2: Edman degradation reaction.

4.1.2. MS methods to determine IsoAsp from Asp residues

An extensive discussion about MS methodologies to differentiate IsoAsp from Asp has been presented in Chapter 1 of this thesis. This section shows a summary to remind the reader about the different methodologies.

Reversed phase High-Performance Liquid Chromatography (RP-HPLC)

Separation of the isomer products of the deamidation reaction can be achieved by HPLC, and the resultant eluent can be analyzed by mass spectrometry. The usual order of elution has been previously reported to be as follows; IsoAsp, Asp, and then the aminosuccinyl form.^{20, 23-26}

A typical HPLC experiment to separate the isomers consists of a linear gradient, a C18 column, and a mobile phase constituted by an acidic solution such as 0.1%

trifluoroacetic acid (TFA), and an organic phase such as acetonitrile (ACN). The acidic solution plays an important role in the separation. Changing the pH of the mobile phase can assist the overall separation by varying the position of interfering species in the chromatogram. For instance, pH 6 was found to be useful to separate four forms of a peptide that was deamidated and isomerized.²⁷ Because the IsoAsp/Asp forms are ionised at pH 6, the non-deamidated form is displaced to higher retention times, therefore separating them for an easier analysis of the chromatogram.

Diagnostic $b+H_2O$ for IsoAsp residues in CAD MS spectra

Several reports of the use of MS to differentiate between IsoAsp from Asp appeared around 1990's.^{28,29} The differentiation was achieved by the presence of the fragment corresponding to the IsoAsp $[b_{n-1} + H_2O]^+$ in the spectrum, where n is the position of the Asp/IsoAsp; this fragment was not found in the spectrum corresponding to the Asp form of the same peptide. The same fragment was also used to differentiate the IsoAsp from Asp forms of a deamidated proteolytic peptide, from hirudin, an anticoagulant peptide.³⁰ The formation of the fragment $[b_{n-1} + H_2O]^+$ could be explained using the mechanism proposed by Papayannopoulos,²⁸ where nucleophilic attack of the carboxyl OH group of β -Asp at carbonyl-protonated serine can be assumed to form an oxazolidone intermediate. This intermediate should easily decompose to the observed $[b_{n-1} + H_2O]^+$ ion, along with CO and a C-terminal aldimine-type moiety that represent a pair of highly stable neutral species. Since the IsoAsp side chain resembles the C-terminus, generation of $[b_{n-1} + H_2O]^+$ for IsoAsp residues is suggested to resemble the fragmentation channel shown to occur in the

low-energy MS/MS spectra of peptides where a C-terminal hydroxyl rearrangement generates $[b_{n-1} + H_2O]^+$ fragment ions.

Immonium ions of IsoAsp residues

Immonium ions are small internal fragments with a single side chain formed by a combination of *a* type and *y* type cleavage, as previously shown in Scheme 1.3. Immonium ions can be used to differentiate the IsoAsp from Asp. Reports have shown that the intensity of the Asp immonium ion in the MS/MS spectrum of an IsoAsp peptide is decreased compared to the intensity of the immonium ion from the Asp form.

In 2000, a study by Gonzalez and co-workers⁹ made use of immonium ions to differentiate IsoAsp from Asp. In their spectrum a signal at 70 Da was shown, which was not detected for the Asp isoform. A mechanism for the formation of this immonium ion was proposed via a rearrangement of a primary carbocation (unstable structure), since it was not found in the IsoAsp spectrum. The loss of water and a further rearrangement yields a more stable acylium ion ($C_3H_4NO^+$), detected at m/z 70.02928. Regardless, these immonium ions could not be used to predict the presence of an IsoAsp residue because Pro also yields an intense immonium ion ($C_4H_8N^+$) at m/z 70.06567.

Use of ECD to differentiate IsoAsp from Asp species

O'Connor and co-workers proposed an alternative way of identifying Asp from IsoAsp.^{19, 20, 22} This method uses ECD to differentiate the isomers products of the deamidation reaction and it has been discussed in depth in Chapter 1. Thus the scope

of the work presented in this chapter is to show the application of this method to differentiate the isomer products of the deamidation reaction in collagen peptides extracted from ancient bones (\approx 3,000 years old). The application can be further investigated as an alternative method of calculating the extent of deamidation of Asn. To date, no reports of this type are recorded in the literature.

4.2. EXPERIMENTAL METHODOLOGY

4.2.1. Materials

Bovine bone samples used for this study were provided by Thomas Higham from the University of Oxford (see Figs. 4.1-4.5). Sequencing-grade trypsin, ammonium acetate ($\text{CH}_3\text{COONH}_4$), and ammonium bicarbonate (NH_4HCO_3) were purchased from Sigma Aldrich, Gillingham, UK. HPLC-grade methanol and formic acid (HCOOH) were obtained from Fisher Scientific, Leicestershire, UK.

4.2.2. Collagen extraction from archeological bones

Bones were ground into a powder using a coffee grinder (see Fig. 4.6). Triplicates of 14 mg of this powder were weighed and demineralised with 1 mL of HCl (0.6M) at 4 °C for 4 hours. From this extraction, the non-soluble, collagen containing fraction was rinsed three times with pure water until neutral pH was reached. Prior to digestion, the non-soluble fraction was denatured in 1 mL of 50 mM ammonium bicarbonate following a previously described methodology.²⁶

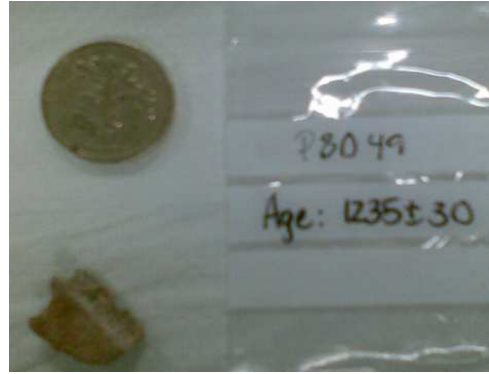
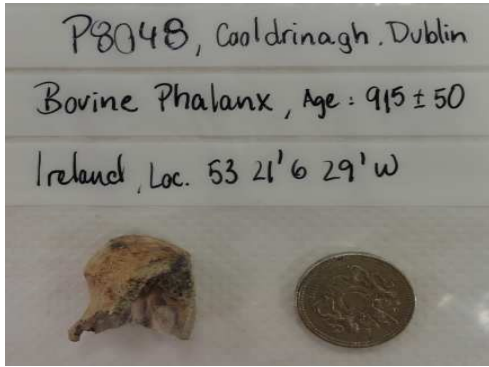


FIGURE 4.1: Samples labeled with number 8 (left) and number 9 (right). Both are bovine phalanx from Cooldrinagh, Dublin. Age according to radio carbon date is (915±50) and (1235±30) years respectively.



FIGURE 4.2: Samples labeled with number 10 (left) and number 11 (right). Both samples are bone from bovine, Ness of Brodgar, Scotland. Age according to radio carbon date is (3829±27) and (3878±26) years old respectively.

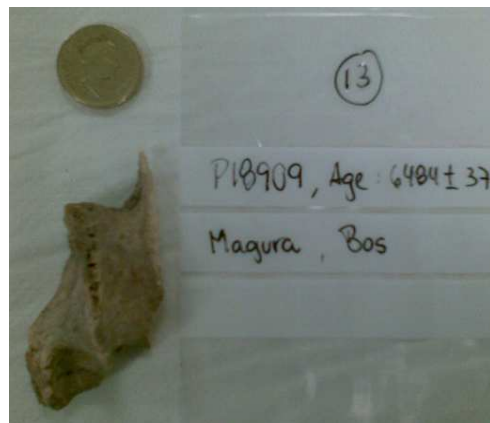


FIGURE 4.3: Samples labeled with number 12 (left) and number 13 (right). Both are bone from Bos Taurus, Magura. Age according to radio carbon date is (6354±37) and (6484±37) years old respectively.



FIGURE 4.4: Samples labeled with number 14 (left) and number 15 (right). Both are bone from Bovini, La Ferrassie, France. Age according to radio carbon date is (31400 ± 400) and (32100 ± 450) years old respectively.



FIGURE 4.5: Samples labeled with number 16 (left) and number 17 (right). Both are bone from Bovini, La Ferrassie, and La Chauverie, France. Age according to radio carbon date is (32550 ± 450) and (50000 ± 3900) years old respectively.



FIGURE 4.6: Powder bone obtained from the grinding process.

Adjustments to the ratio enzyme:protein had to be considered. The supernatant (700 μL) was collected and 24 μL of trypsin solution (2 $\mu\text{g}/\mu\text{L}$) was added using an approximate ratio of 1:200 (w/w) enzyme:protein. To minimise deamidation during tryptic digestion the mixture was allowed to react for 4 hours at 37°C,³¹ the reaction was halted by adding 5 % (v/v) of formic acid, and the samples were dried in a SpeedVac system (Savant SPD121P Fisher Scientific, Leicestershire, UK) to evaporate volatiles at 35 °C, following storage at -80 °C until analysis (\approx 1 day). The samples were reconstituted into a mixture of 1:1 MeOH:H₂O, and diluted into a mixture of 49.5:49.5:1 MeOH:H₂O:formic acid to a final approximate concentration of 2.5 μM .

4.2.3. FT-ICR-MS measurements

The experiments were carried out on a Bruker 12T solariX FT-ICR-MS (Bruker Daltonics). In order to identify deamidation sites, samples of digested collagen were electrosprayed at approximately 2.5 μM concentration in 49.5:49.5:1 MeOH:H₂O:formic acid. An average of 200 scans was taken.

CAD and ECD experiments were carried out following procedures explained in Chapter 2 and 3 respectively.. See tables in section 4.6 of this chapter for information in mass assignments.

4.3. RESULTS AND DISCUSSION

4.3.1. Identification of collagen peptides originated from ancient cow bones

A range of different ages of bone samples were provided by Professor Thomas Higham from the University of Oxford. The purpose of the analysis was to identify potential deamidation sites in peptides extracted from the samples, that could serve as markers to measure the extent of deamidation, using the previous methodology developed in our laboratory.³² Tables 4.1 and 4.2 in section 4.6 of this chapter show the assigned peptides throughout the range of samples selected.

Figure 4.7 shows the mass spectrum of a potential deamidated peptide of collagen extracted from a range of cow bone samples aged from 900 years to 33,000 years old. This peptide corresponds to m/z [793.88153]²⁺, which is consistent with previous studies.²⁶ Figs. 4.8 and 4.9 show other two potential deamidated peptides corresponding to m/z [927.78113]³⁺, and m/z [829.74176]³⁺. From this range of samples, the spectrum with the best signal-to-noise ratio was selected to show the presence of the three potential marker peptides (Fig. 4.10). Consequently, this particular sample was chosen to carry out the experiments for the differentiation of IsoAsp from Asp using ECD. This spectrum relates to a bone sample aged around (3,829±27) years old according to radio carbon dating.

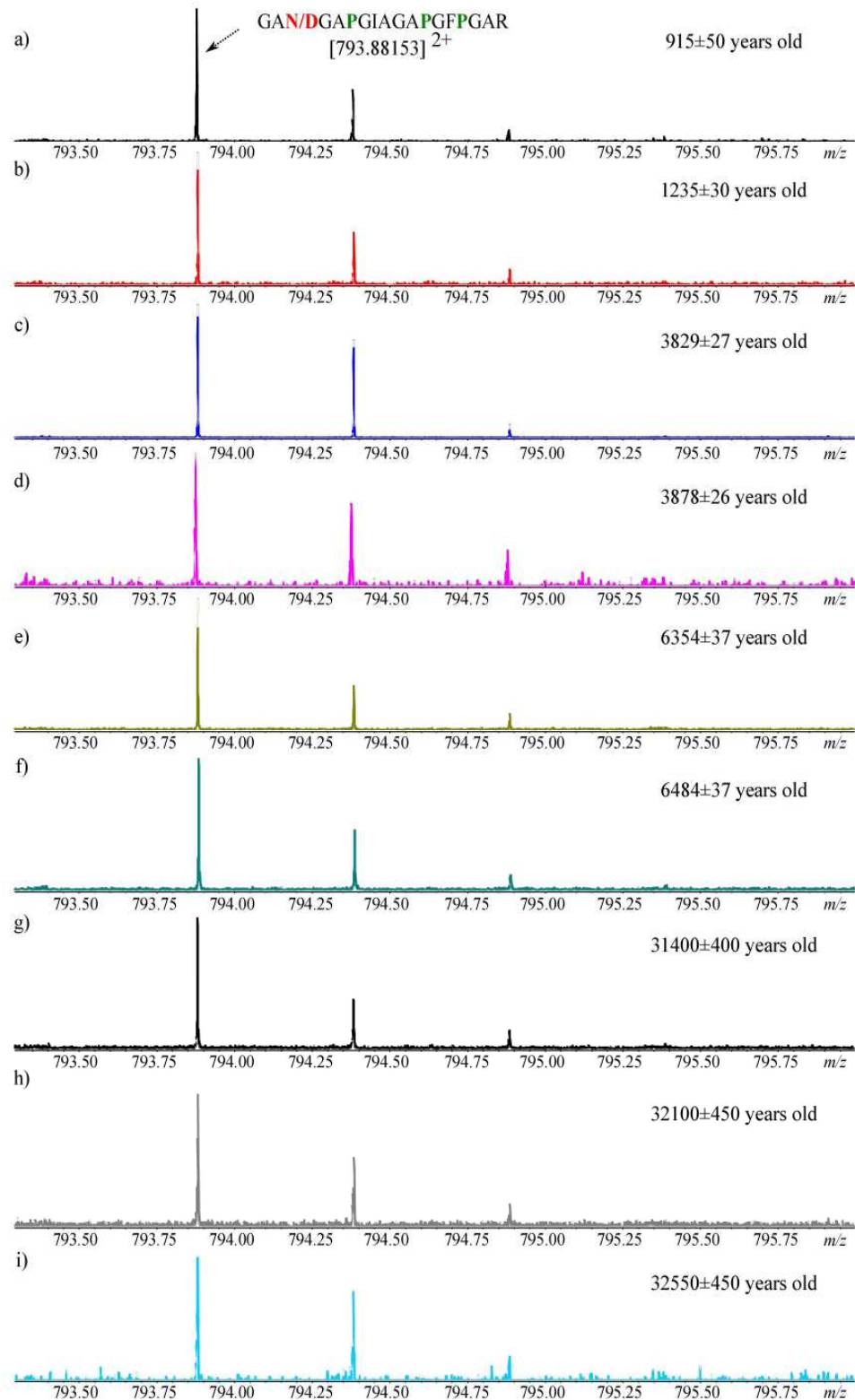


FIGURE 4.7: a) to i) Enlargement of the m/z $[793.88153]^{2+}$ ion present in the spectrum, which corresponds to a fully deamidated peptide of collagen extracted from archaeological bones. The age of the bones used in the experiment are shown in the figure.

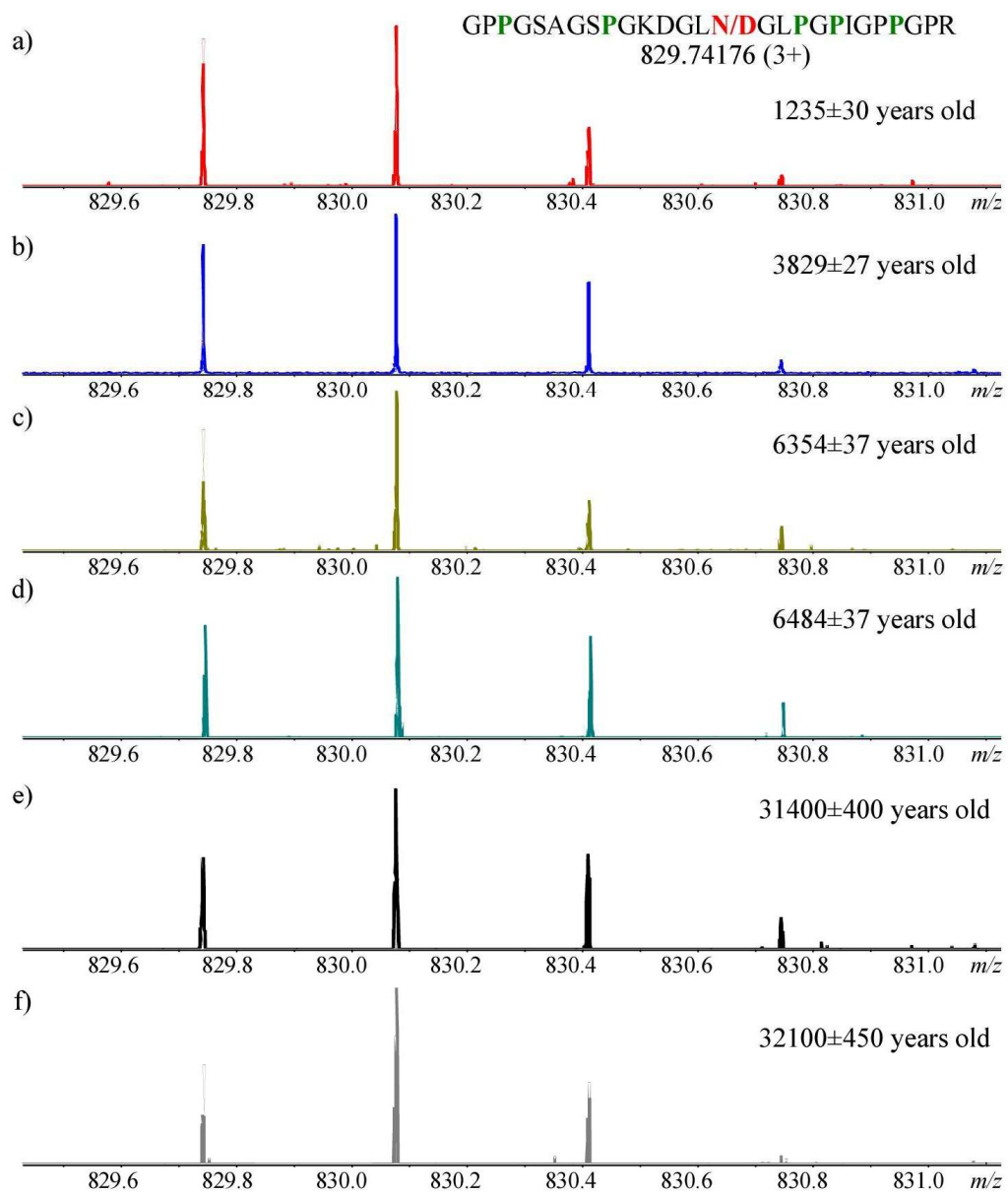


FIGURE 4.8: a) to f): Enlargement of the m/z $[829.74176]^{3+}$ ion present in the spectrum, which corresponds to a deamidated peptide of collagen extracted from cattle bone. The figure shows the age of the bone used.

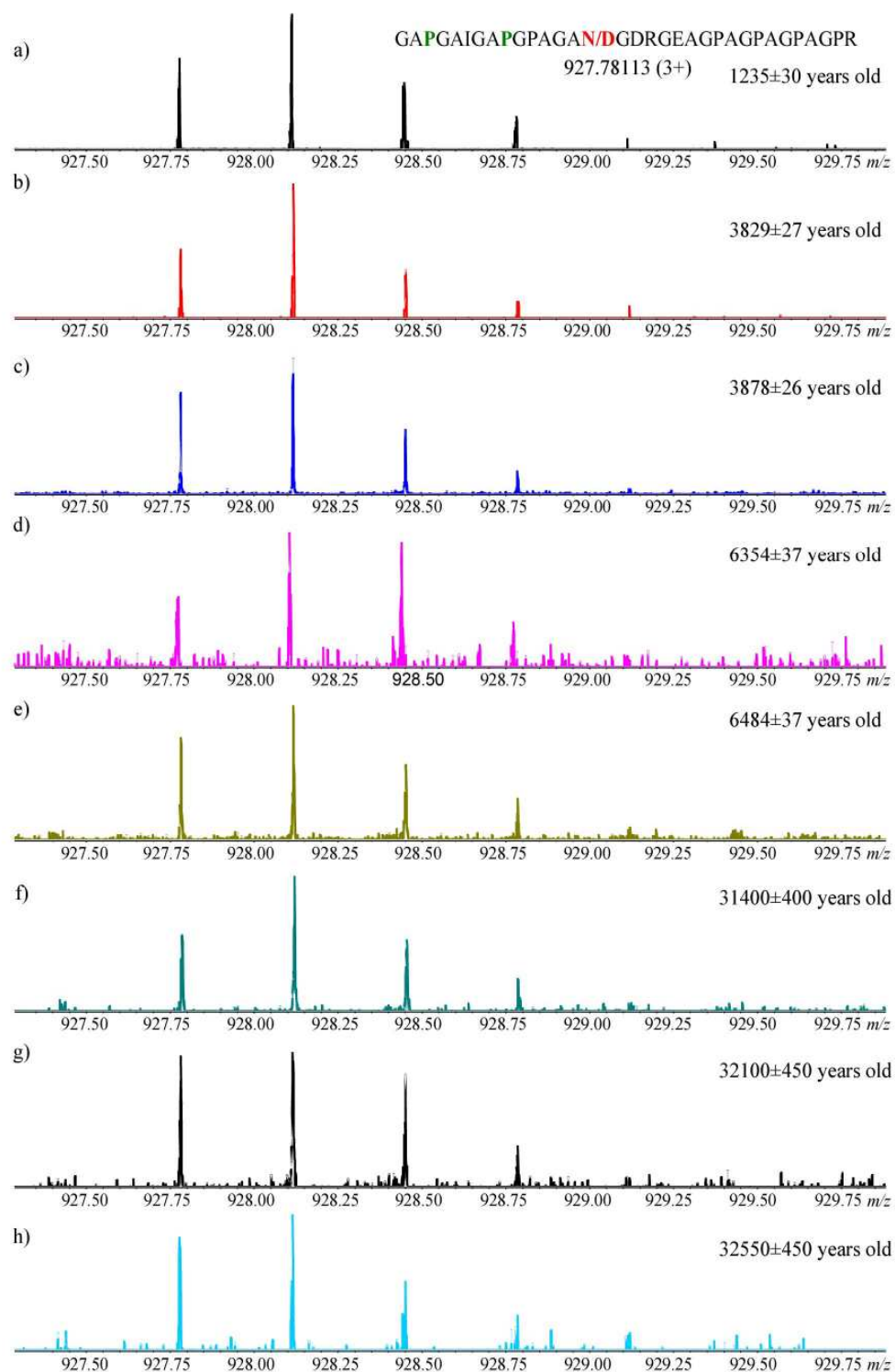


FIGURE 4.9: a) to h): Enlargement of the m/z $[927.78113]^{3+}$ ion present in the spectrum, which correspond to a deamidated peptide of collagen extracted from bone, which ages are shown in the figure.

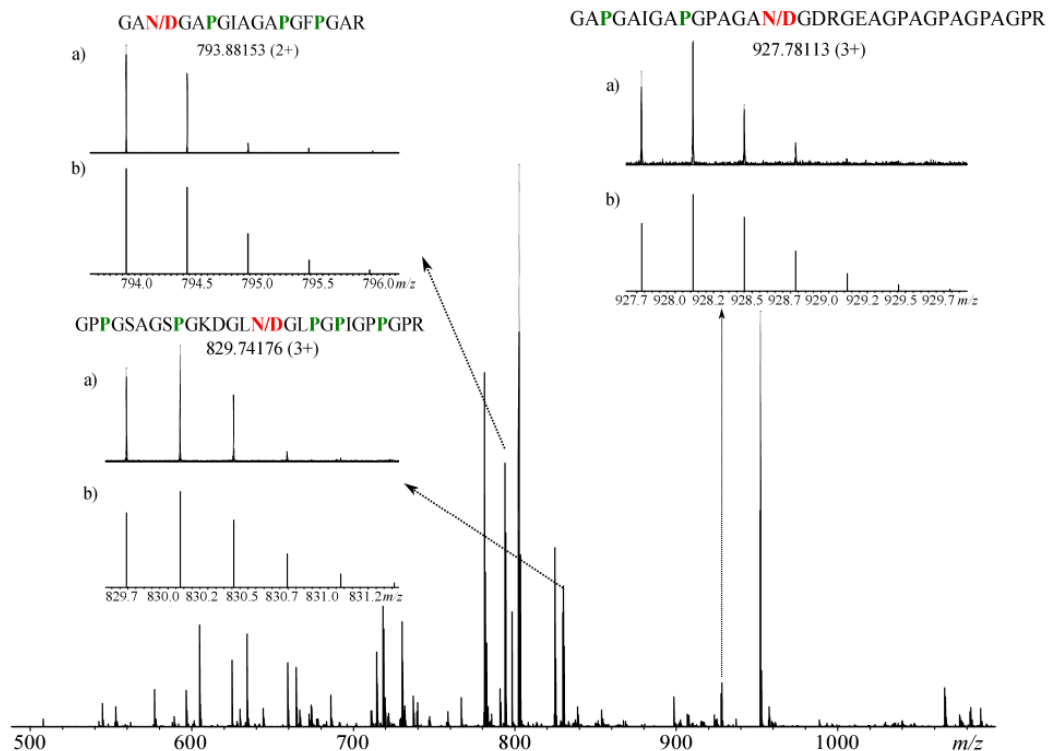


FIGURE 4.10: Mass spectrum of tryptic digest of collagen extracted from 3,000 year-old cow bones. **N/D** represents deamidation site. **P** represents Hyp. The inserts show three possible deamidated peptides corresponding to m/z [793.88153]²⁺, m/z [927.78113]³⁺, and m/z [829.74176]³⁺. a) and b) are real and simulated isotopic distributions respectively.

4.3.2. Differentiation of Asp from IsoAsp in collagen using ECD

Since Asp and IsoAsp residues have exactly the same mass, their differentiation using mass spectrometry can be challenging, but the accessibility to ECD assisted the analysis.²⁰ Under ECD conditions the peptides with Asp residues are likely to suffer neutral loss of the Asp side chain, which are not normally observed in the IsoAsp equivalents.¹⁹ Unique fragmentation patterns have been previously observed in the ECD spectra of peptides with IsoAsp residues that do not appear in the spectra of the peptides with Asp residues only.¹⁹ However, these observations were obtained in model peptides with aspartyl residues and their IsoAsp analogues. In the peptides originate from the deamidation reaction, a mixture of the isomers is expected. This mixture is generally at a ratio of 3:1 IsoAsp:Asp at pH 7.4.³³

The differentiation of IsoAsp from Asp in collagen from ancient samples is highlighted here, particularly as an alternative way of measuring the extent of deamidation, which was reported in Chapter 2 and 3. Although the mass defect method is a faster and a simpler method, its utility is limited to ultrahigh resolution mass spectrometers. The extent of deamidation could be calculated by assuming that the peak intensities of the diagnostic ions for Asp and IsoAsp are additive and equivalent to the peak intensity of D. This can be achieved by plotting $N_t/(N_t+D_t)$ vs time, and using $D_t = \text{IsoAsp}+\text{Asp}$. This approach is particularly useful when the separation of the isotopes involved in the deamidation reaction cannot be achieved, which was the case in this study due to a failure of the instrument as is shown in Fig. 4.11.

Figure 4.12 shows the ECD and CAD spectra of the peptide with the sequence GAN**GA**PGIAGAP**GF**PGAR [coll_I (α 1)-397-414], which corresponds to the ion m/z [793.88153]²⁺. Deamidation of the Asn located at position 3 from the N-terminus is highlighted in red in Fig. 4.12. The ion z_{16-57}^{\bullet} with m/z 1385.68101 is the diagnostic ion that indicates the presence of IsoAsp in position three from the N-terminus in the peptide. Asp diagnostic ion $[M+2H]^{\bullet}-60$ was also found in the spectrum, which m/z corresponds to 1527.74192. The clear distinction between the two ions in this particular peptide could be further used to determine the extent of deamidation before sample preparation.

The use of CAD complemented the fragmentation obtained from ECD, which contributed to sequence the Hyp present in the peptide with higher level of

confidence. For more information related to mass accuracy and assignments of the ions see tables 4.2 and 4.3 in section 4.6 of this chapter.

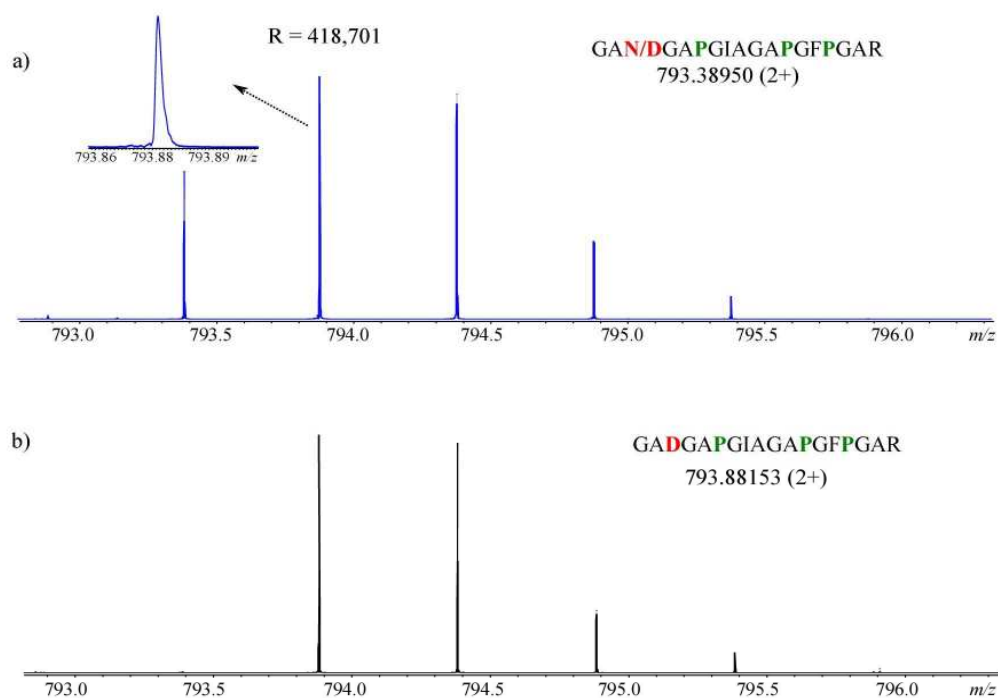


FIGURE 4.11: a) Narrowband spectrum corresponding to m/z [793.88153]²⁺ ion in a mixture of the non-deamidated and the deamidated form. The insert represents the zoom of the peak where the two isotopes should appear. b) Narrowband spectrum of the deamidated form of the peptide m/z [793.88153]²⁺.

Figure 4.13 shows the ECD of another possible deamidated peptide with a sequence **GPPGSAGSPGKDGLNGLPPIGPPGPR** [coll_I (α 1)-1141-1167], which relates to the ion m/z [829.74184]³⁺. The presence of the IsoAsp diagnostic ion m/z 1284.59434 $c_{14} \bullet +58$ indicates the presence of IsoAsp. On the other hand, the loss of Asp side chain from the reduced charged species $[M+3H]^{2+ \bullet} -60$ m/z 1184.59142 could not be used to determine the presence of Asp at position 15 from the N-

terminus. This is due to the presence of another Asp at position 12 from the N-terminus.

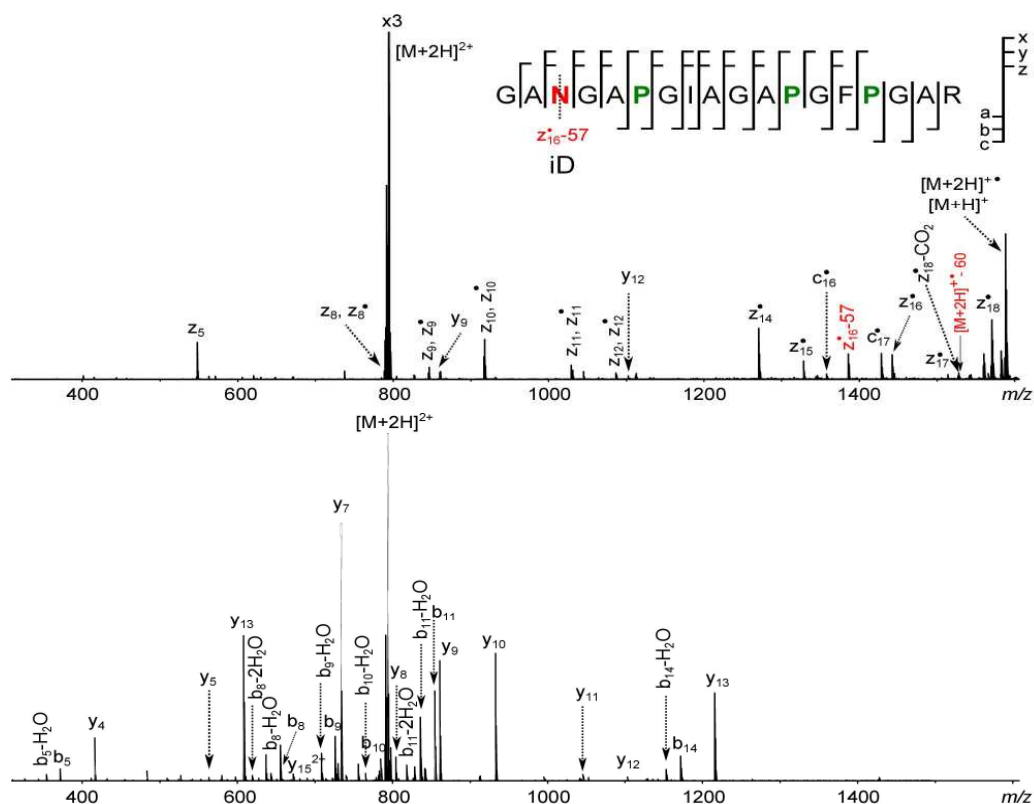


FIGURE 4.12: Up: ECD and Down: CAD of tryptic digest peptide [coll_I (α 1) 397-414]²⁺ ion at m/z 793.88150. The deamidation site is located in position 3 (N-terminal), in the peptide (N). Diagnostic ion corresponding to the deamidation at position 3 is labelled as z_{16}^{*57} . P in green represent the modification of hydroxylation.

Looking across figure 4.14, which corresponds to the spectrum of the peptide with sequence “GAPGAIGAPGPAGANGDRGEAGPAGPAGPR” at [927.78145]³⁺ m/z no diagnostic ions of the isomers were found in the ECD spectrum of this peptide. Due to the absence of these ions, it will not be possible to calculate the extent of deamidation for this particular peptide using the IsoAsp/Asp ratios.

The preliminary analysis reported here shows the presence of potential deamidated peptides in remarkable aged samples, which brings the possibility of mapping the

extent of deamidation against radio carbon date of the sample. Thus, information about the relationship between deamidation and ageing can be further investigated.

In general, the presence of diagnostic ions ($z^{\bullet}-57$ and $[M+nH]^{n+}-60$) along with the mass defect method can be used to determine the extent of deamidation in the sample. In some cases, the absence of the diagnostic ions does not allow the use of the peak intensities, thus limiting its use.

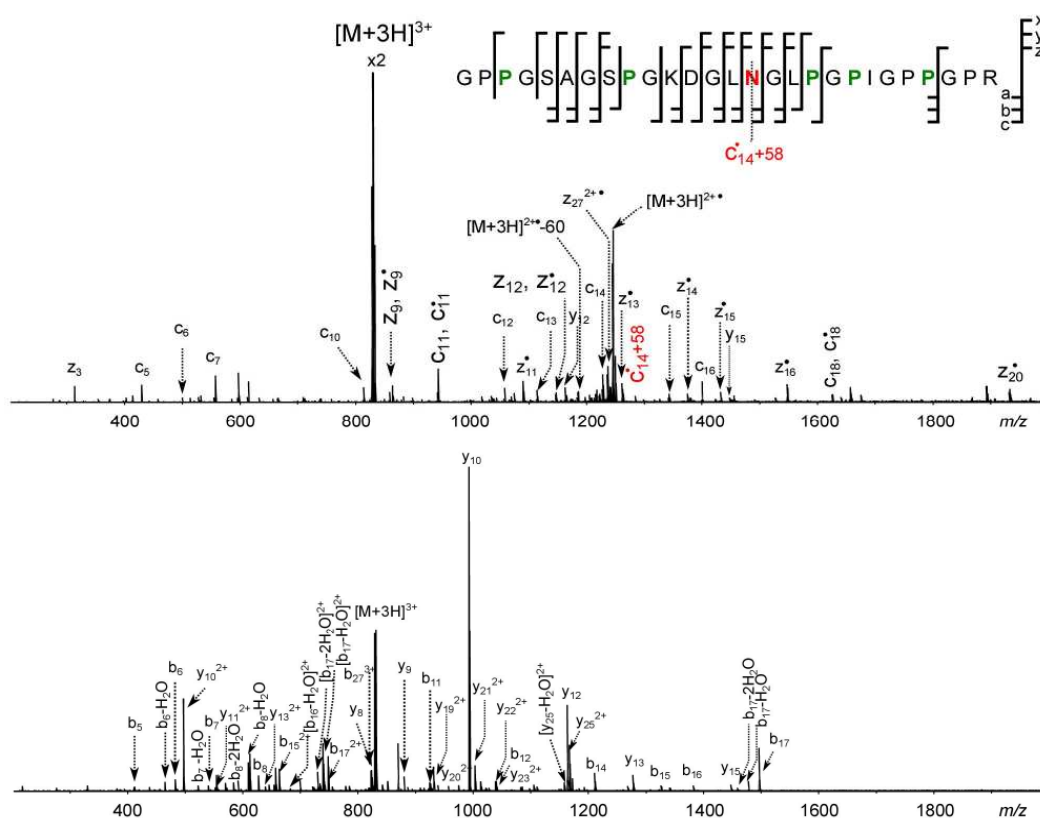


FIGURE 4.13: Up: ECD and Down: CAD of tryptic digest peptide [coll_I ($\alpha 1$) 1141-1167]³⁺ ion at m/z 829.74184. The deamidation site is located in position 15 (N-terminal) labelled as (N), in the peptide. Diagnostic ion corresponding to the deamidation site is labelled as $c_{14}^{\bullet}+58$ in the ECD spectrum. P in green represent the modification of hydroxylation.

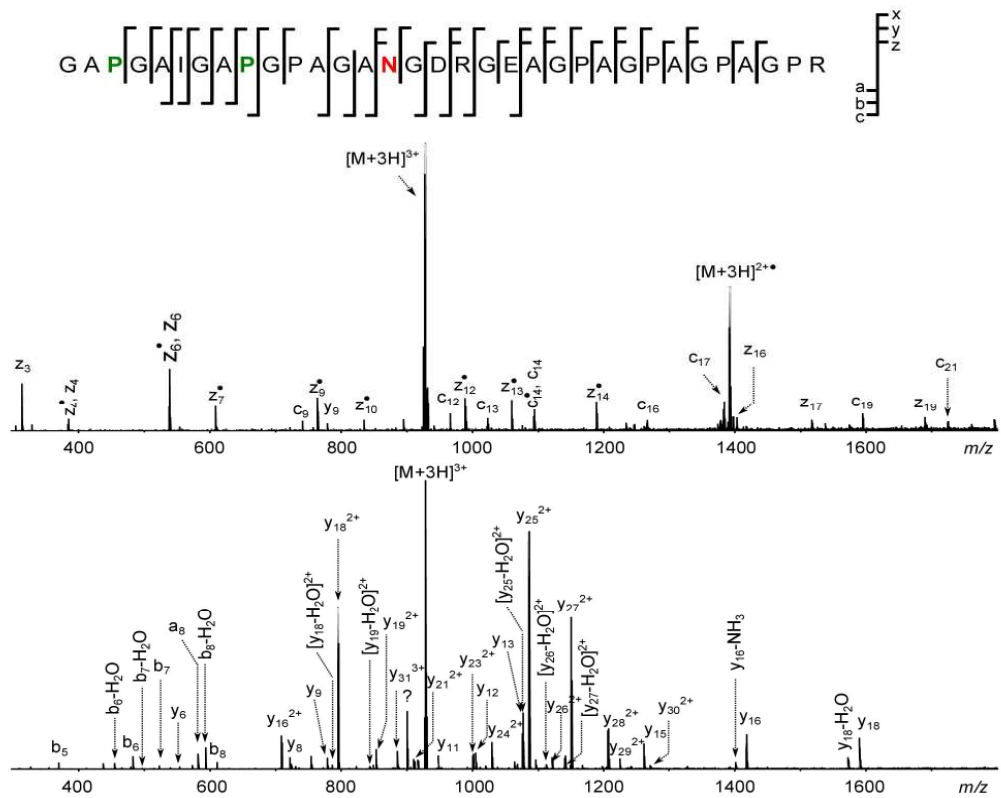


FIGURE 4.14: Up: ECD and Down: CAD of tryptic digest peptide [coll_I ($\alpha 2$) 674-706] $^{3+}$ ion at m/z 927.78145. The deamidation site is located in position 15 (N-terminal), labelled as (N) in the peptide. P highlighted in green represent the modification of hydroxylation.

4.4. CONCLUSIONS

The use of FT-ICR-MS allowed the localisation of several deamidation sites in the peptides found in the bone samples studied. Specifically, the use of ECD assisted in the differentiation of IsoAsp from Asp in some cases. The peptides that showed deamidation and the differentiation of the isomers were possible can be utilized in the analysis of IsoAsp+Asp peak intensities to calculate the extent of deamidation in the samples.

This study represents a preliminary investigation of the peptides, which deamidation extent needs to be further investigated. The proposal is to determine the extent of deamidation of the peptides reported here using a combination of the methodology presented in Chapter 2 and the use of IsoAsp+Asp peak intensities method. The study needs to be carried out throughout the range of different aged samples in order to establish a relationship between age and deamidation.

4.5. REFERENCES

1. Chelius, D., Rehder, D.S., and Bondarenko, P.V. Identification and Characterization of Deamidation Sites in the Conserved Regions of Human Immunoglobulin Gamma Antibodies. *Analytical Chemistry*. **2005**, 77, 6004-6011.
2. Robinson, N.E. and Robinson, A.B. Molecular Clocks. *Proceedings of the National Academy of Sciences of the United States of America*. **2001**, 98, 944-949.
3. Ritz-Timme, S. and Collins, M.J. Racemization of Aspartic Acid in Human Proteins. *Ageing Research Reviews*. **2002**, 1, 43-59.
4. Carvalho, R.N., Solstad, T., Bjorgo, E., Barroso, J.F., and Flatmark, T. Deamidations in Recombinant Human Phenylalanine Hydroxylase - Identification of Labile Asparagine Residues and Functional Characterization of Asn -> Asp Mutant Forms. *Journal of Biological Chemistry*. **2003**, 278, 15142-15152.
5. Nilsson, M.R., Driscoll, M., and Raleigh, D.P. Low Levels of Asparagine Deamidation Can Have a Dramatic Effect on Aggregation of Amyloidogenic Peptides: Implications for the Study of Amyloid Formation. *Protein Science*. **2002**, 11, 342-349.
6. Fujii, N., Ishibashi, Y., Satoh, K., Fujino, M., and Harada, K. Simultaneous Racemization and Isomerization at Specific Aspartic Acid Residues in Alpha B-Crystallin from the Aged Human Lens. *Biochimica Et Biophysica Acta*. **1994**, 1204, 157-163.
7. Roher, A.E., Lowenson, J.D., Clarke, S., Wolkow, C., Wang, R., Cotter, R.J., Reardon, I.M., Zurcherneely, H.A., Heinrikson, R.L., Ball, M.J., and Greenberg, B.D. Structural Alterations in the Peptide Backbone of Beta-Amyloid Core Protein May Account for Its Deposition and Stability in Alzheimers-Disease. *Journal of Biological Chemistry*. **1993**, 268, 3072-3083.
8. Shimizu, T., Matsuoka, Y., and Shirasawa, T. Biological Significance of Isoaspartate and Its Repair System. *Biological & pharmaceutical bulletin*. **2005**, 28, 1590-1596.
9. Gonzalez, L.J., Shimizu, T., Satomi, Y., Betancourt, L., Besada, V., Padron, G., Orlando, R., Shirasawa, T., Shimonishi, Y., and Takao, T. Differentiating Alpha- and Beta-Aspartic Acids by Electrospray Ionization and Low-Energy Tandem Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2000**, 14, 2092-2102.
10. Lloyd, J.R., Cotter, M.L., Otori, D., and Doyle, D.L. Distinction of Alpha-Aspartyl and Beta-Aspartyl and Alpha-Glutamyl and Gamma-Glutamyl-Transferase Peptides by Fast Atom Bombardment Tandem Mass Spectrometry. *Biomedical and Environmental Mass Spectrometry*. **1988**, 15, 399-402.
11. Lehmann, W.D., Schlosser, A., Erben, G., Pipkorn, R., Bossemeyer, D., and Kinzel, V. Analysis of Isoaspartate in Peptides by Electrospray Tandem Mass Spectrometry. *Protein Science*. **2000**, 9, 2260-2268.
12. Bada, J.L., Shou, M. Y., in *Biogeochemistry of Amino Acids*, ed. P.E. Hare, Hoering, T. C., King, K. J. Wiley: New York, 1980, pp 235-255.
13. Lapko, V.N., Purkiss, A.G., Smith, D.L., and Smith, J.B. Deamidation in Human Gamma S-Crystallin from Cataractous Lenses Is Influenced by Surface Exposure. *Biochemistry*. **2002**, 41, 8638-8648.

14. Fernandez, E., Ortiz, J.E., Perez-Perez, A., Prats, E., Turbon, D., Torres, T., and Arroyo-Pardo, E. Aspartic Acid Racemization Variability in Ancient Human Remains: Implications in the Prediction of Ancient DNA Recovery. *Journal of Archaeological Science*. **2009**, *36*, 965-972.
15. Harms, M.J., Wilmarth, P.A., Kapfer, D.M., Steel, E.A., David, L.L., Bachinger, H.P., and Lampi, K.J. Laser Light-Scattering Evidence for an Altered Association of Beta B1-Crystallin Deamidated in the Connecting Peptide. *Protein Science*. **2004**, *13*, 678-686.
16. Zomber, G., Reuveny, S., Garti, N., Shafferman, A., and Elhanany, E. Effects of Spontaneous Deamidation on the Cytotoxic Activity of the Bacillus Anthracis Protective Antigen. *Journal of Biological Chemistry*. **2005**, *280*, 39897-39906.
17. Solstad, T., Carvalho, R.N., Andersen, O.A., Waidelich, D., and Flatmark, T. Deamidation of Labile Asparagine Residues in the Autoregulatory Sequence of Human Phenylalanine Hydroxylase - Structural and Functional Implications. *European Journal of Biochemistry*. **2003**, *270*, 929-938.
18. Vanbelle, C., Halgand, F., Cedervall, T., Thulin, E., Akerfeldt, K.S., Laprevote, O., and Linse, S. Deamidation and Disulfide Bridge Formation in Human Calbindin D-2gk with Effects on Calcium Binding. *Protein Science*. **2005**, *14*, 968-979.
19. Cournoyer, J.J., Pittman, J.L., Ivleva, V.B., Fallows, E., Waskell, L., Costello, C.E., and O'connor, P.B. Deamidation: Differentiation of Aspartyl from Isoaspartyl Products in Peptides by Electron Capture Dissociation. *Protein Science*. **2005**, *14*, 452-463.
20. Cournoyer, J.J., Lin, C., and O'connor, P.B. Detecting Deamidation Products in Proteins by Electron Capture Dissociation. *Analytical Chemistry*. **2006**, *78*, 1264-1271.
21. Sargaeva, N.P., Lin, C., and O'connor, P.B. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid Beta Peptides, Including a Beta 1-42, Using Electron-Ion Reactions. *Analytical Chemistry*. **2009**, *81*, 9778-9786.
22. Cournoyer, J.J., Lin, C., Bowman, M.J., and O'connor, P.B. Quantitating the Relative Abundance of Isoaspartyl Residues in Deamidated Proteins by Electron Capture Dissociation. *Journal of the American Society for Mass Spectrometry*. **2007**, *18*, 48-56.
23. Zischler, H., Hoss, M., Handt, O., Vonhaeseler, A., Vanderkuyl, A.C., Goudsmit, J., and Paabo, S. Detecting Dinosaur DNA. *Science*. **1995**, *268*, 1192-1193.
24. Pawlicki, R. Histochemical-Demonstration of DNA in Osteocytes from Dinosaur Bones. *Folia Histochemica Et Cytobiologica*. **1995**, *33*, 183-186.
25. Jedrezejewski, P.T., Girod, A., Tholey, A., Konig, N., Thullner, S., Kinzel, V., and Bossemeyer, D. A Conserved Deamidation Site at Asn 2 in the Catalytic Subunit of Mammalian Camp-Dependent Protein Kinase Detected by Capillary Lc/Ms and Tandem Mass Spectrometry. *Protein Science*. **1998**, *7*, 457-469.
26. Perez Hurtado, P. and O'connor, P.B. Deamidation of Collagen. *Analytical Chemistry*. **2012**, *84*, 3017-3025.
27. Schweitzer, M.H., Hill, C.L., Asara, J.M., Lane, W.S., and Pincus, S.H. Identification of Immunoreactive Material in Mammoth Fossils. *Journal of Molecular Evolution*. **2002**, *55*, 696-705.

28. Papayannopoulos, I.A. and Biemann, K. Fast-Atom-Bombardment and Tandem Mass-Spectrometry of Synthetic Peptides and by-Products. *Peptide Research*. **1992**, *5*, 83-90.
29. Carr, S.A., Hemling, M.E., Bean, M.F., and Roberts, G.D. Integration of Mass Spectrometry in Analytical Biotechnology. *Analytical Chemistry*. **1991**, *63*, 2802-2824.
30. Schindler, P., Muller, D., Marki, W., Grossenbacher, H., and Richter, W.J. Characterization of a Beta-Asp33 Isoform of Recombinant Hirudin Sequence Variant 1 by Low-Energy Collision-Induced Dissociation. *Journal of Mass Spectrometry*. **1996**, *31*, 967-974.
31. Li, X., Cournoyer, J.J., Lin, C., and O'connor, P.B. Use of O-18 Labels to Monitor Deamidation During Protein and Peptide Sample Processing. *Journal of the American Society for Mass Spectrometry*. **2008**, *19*, 855-864.
32. Zabrouskov, V., Han, X.M., Welker, E., Zhai, H.L., Lin, C., Van Wijk, K.J., Scheraga, H.A., and McLafferty, F.W. Stepwise Deamidation of Ribonuclease a at Five Sites Determined by Top Down Mass Spectrometry. *Biochemistry*. **2006**, *45*, 987-992.
33. Robinson, N.E. and Robinson, A.B., *Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*. Althouse Press: Cave Junction, OR, **2004**.

4.6. MASS ACCURACY AND PEPTIDE ASSIGNMENTS

TABLE 4.1: Molecular formulae and ion masses of the peptides assigned across the range of samples analyzed

Elemental composition	[M+H]¹⁺	[M+2H]²⁺	[M+3H]³⁺
C ₂₆ H ₄₅ N ₉ O ₁₀	644.33621	322.67175	215.45026
C ₃₅ H ₅₇ N ₁₃ O ₁₁	836.43730	418.72230	279.48395
C ₃₇ H ₅₈ N ₁₀ O ₁₃	851.42573	426.21652	284.48010
C ₃₆ H ₆₁ N ₁₃ O ₁₂	868.46352	434.73540	290.15936
C ₃₅ H ₅₉ N ₁₃ O ₁₄	886.43770	443.72250	296.15075
C ₃₈ H ₆₂ N ₁₂ O ₁₃	895.46317	448.23524	299.15924
C ₃₈ H ₆₇ N ₁₃ O ₁₂	898.51047	449.75888	300.17501
C ₃₈ H ₆₇ N ₁₃ O ₁₂	898.51049	449.75888	300.17502
C ₄₈ H ₇₃ N ₁₃ O ₁₆	1088.53700	544.77219	363.51722
C ₄₈ H ₇₃ N ₁₃ O ₁₆	1088.53710	544.77219	363.51722
C ₄₄ H ₇₄ N ₁₈ O ₁₅	1095.56538	548.28633	365.85998
C ₄₇ H ₇₆ N ₁₆ O ₁₅	1105.57488	553.29108	369.19648
C ₅₁ H ₈₀ N ₁₄ O ₁₆ S ₁	1177.56702	589.28715	393.19386
C ₅₁ H ₇₉ N ₁₃ O ₁₇ S ₁	1178.55104	589.77916	393.52186
C ₅₂ H ₈₅ N ₁₅ O ₁₇	1192.63206	596.81967	398.21554
C ₅₃ H ₈₀ N ₁₄ O ₁₈	1201.58478	601.29603	401.19978
C ₅₃ H ₇₉ N ₁₃ O ₁₉	1202.56879	601.78804	401.52778
C ₅₂ H ₈₉ N ₁₇ O ₁₆	1208.67460	604.84094	403.56305
C ₅₂ H ₈₈ N ₁₆ O ₁₇	1209.65861	605.33294	403.89106
C ₅₁ H ₈₃ N ₁₅ O ₂₂	1258.59099	629.79913	420.20185
C ₅₃ H ₉₀ N ₁₈ O ₁₈	1267.67533	634.34130	423.22996
C ₅₃ H ₉₀ N ₁₈ O ₁₈	1267.67533	634.34130	423.22996
C ₅₆ H ₈₆ N ₁₆ O ₁₉	1287.63279	644.32003	429.88245
C ₅₆ H ₈₅ N ₁₅ O ₂₀	1288.61681	644.81204	430.21045
C ₅₉ H ₈₉ N ₁₅ O ₂₀	1328.64811	664.82769	443.55422
C ₅₇ H ₈₉ N ₁₇ O ₂₀	1332.65422	666.83075	444.88959
C ₅₉ H ₈₉ N ₁₅ O ₂₁	1344.64302	672.82515	448.88586

TABLE 4.1: Molecular formulae and ion masses of the peptides assigned across the range of samples analyzed (Continued)

Elemental composition	[M+H]¹⁺	[M+2H]²⁺	[M+3H]³⁺
C ₆₃ H ₉₈ N ₁₈ O ₂₀	1427.72775	714.36752	476.58077
C ₆₀ H ₉₄ N ₁₈ O ₂₃	1435.68120	718.34424	479.23192
C ₆₁ H ₉₄ N ₂₀ O ₂₂	1459.69243	730.34985	487.23567
C ₆₁ H ₉₂ N ₂₀ O ₂₃	1473.67170	737.33949	491.89542
C ₆₅ H ₁₀₅ N ₂₁ O ₂₂	1532.78158	766.89443	511.59871
C ₆₇ H ₁₀₉ N ₂₁ O ₂₂	1560.81288	780.91008	520.94248
C ₆₇ H ₁₀₇ N ₁₉ O ₂₄	1562.78091	781.89409	521.59849
C ₆₅ H ₁₀₅ N ₂₁ O ₂₅	1580.76632	790.8868	527.593626
C ₆₇ H ₁₀₃ N ₂₁ O ₂₄	1586.75576	793.88152	529.59011
C ₇₀ H ₁₁₃ N ₂₁ O ₂₅	1648.82893	824.91810	550.28116
C ₆₉ H ₁₀₇ N ₂₃ O ₂₈	1706.77287	853.89007	569.59581
C ₇₈ H ₁₂₁ N ₂₃ O ₂₇	1812.88750	906.94739	604.967352
C ₇₉ H ₁₂₄ N ₂₂ O ₂₈	1829.90282	915.45505	610.639124
C ₇₉ H ₁₂₄ N ₂₂ O ₂₉	1845.89773	923.45251	615.970762
C ₈₂ H ₁₃₄ N ₂₈ O ₂₉	1975.99443	988.50085	659.33633
C ₈₄ H ₁₃₄ N ₂₆ O ₃₂	2019.97302	1010.49015	673.99586
C ₈₉ H ₁₄₆ N ₂₈ O ₂₈	2056.09341	1028.55035	686.03599
C ₈₅ H ₁₃₃ N ₂₉ O ₃₁	2056.97951	1028.99339	686.33135
C ₈₅ H ₁₃₂ N ₂₈ O ₃₂	2057.96352	1029.485399	686.659358
C ₈₉ H ₁₄₆ N ₂₈ O ₂₉	2072.08833	1036.54780	691.36682
C ₉₂ H ₁₅₁ N ₂₇ O ₃₁	2131.11421	1066.06074	711.04292
C ₁₀₅ H ₁₆₇ N ₃₁ O ₃₉	2487.21102	1244.10915	829.74186
C ₁₀₁ H ₁₆₃ N ₃₇ O ₃₇	2487.20834	1244.10781	829.74096
C ₁₁₅ H ₁₈₁ N ₃₉ O ₄₂	2781.32991	1391.16859	927.78155
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₃	2853.41258	1427.20993	951.80904
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₄	2869.40749	1432.207383	957.140681

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 8		Sample 9		Sample 10	
		Observed mass	Error ppm	Observed mass	Error ppm	Observed mass	Error ppm
C ₂₆ H ₄₅ N ₉ O ₁₀	GLPGER/304-309 (α 1)	644.33626	-0.08	-	-	644.33614	0.11
C ₃₅ H ₅₇ N ₁₃ O ₁₁	GPAGPQGPR/1084-1092 (α 1)	418.72229	0.03	418.72231	-0.02	418.72200	0.72
C ₃₇ H ₅₈ N ₁₀ O ₁₃	GFSGLDGAK/268-276 (α 1)	426.21685	-0.78	-	-	426.21651	0.02
C ₃₆ H ₆₁ N ₁₃ O ₁₂	GPSGPQGIR/994-1003 (α 2)	434.73506	0.78	434.73524	0.37	434.73509	0.71
C ₃₅ H ₅₉ N ₁₃ O ₁₄	GSEGPQGVR/361-369 (α 1)	443.72227	0.51	443.72235	0.33	443.72221	0.65
C ₃₈ H ₆₂ N ₁₂ O ₁₃	GPAGPSGPAGK/1052-1062 (α 2)	448.23508	0.36	448.23532	-0.17	448.23499	0.56
C ₃₈ H ₆₇ N ₁₃ O ₁₂	GVVGLPGQR/958-966 (α 1)	449.75865	0.50	449.75885	0.06	449.75891	-0.08
C ₃₈ H ₆₇ N ₁₃ O ₁₂	GVVGLPGQR/958-966 (α 1)	898.51045	0.05	-	-	898.51048	0.01
C ₄₈ H ₇₃ N ₁₃ O ₁₆	GFPGADGVAGPK/493-504 (α 1)	544.77215	0.07	544.77218	0.01	544.77225	-0.11
C ₄₈ H ₇₃ N ₁₃ O ₁₆	GFPGADGVAGPK(1oxi)/493-504 (α 1)	-	-	-	-	1088.53706	0.04
C ₄₄ H ₇₄ N ₁₈ O ₁₅	GPRGDQGPVGR/818-828 (α 2)	-	-	-	-	-	-
C ₄₇ H ₇₆ N ₁₆ O ₁₅	GVQPPGPAGPR(1oxi)/685-696 (α 1)	553.29105	0.05	553.29106	0.03	553.29108	0.00
C ₅₁ H ₈₀ N ₁₄ O ₁₆ S ₁	GQAGVMGFPGPK(2oxi)/574-585 (α 1)	589.28701	0.23	589.28706	0.15	589.28706	0.15
C ₅₁ H ₇₉ N ₁₃ O ₁₇ S ₁	GQAGVMGFPGPK/574-575 (α 1)	-	-	-	-	-	-
C ₅₂ H ₈₅ N ₁₅ O ₁₇	GVPGPPGAVGPAGK(2oxi)/598-611 (α 1)	596.81974	-0.12	596.81983	-0.27	596.81978	-0.18
C ₅₃ H ₈₀ N ₁₄ O ₁₈	GEPGNIGFPGPK/485-496 (α 2)	601.29630	-0.45	601.29615	-0.20	601.29644	-0.69
C ₅₃ H ₇₉ N ₁₃ O ₁₉	GEPGNIGFPGPK/485-496 (α 2)	-	-	-	-	-	-
C ₅₂ H ₈₉ N ₁₇ O ₁₆	IGQPGAVGPAGIR/1066-1078 (α 2)	604.84105	-0.19	604.84084	0.16	604.84093	0.01
C ₅₂ H ₈₈ N ₁₆ O ₁₇	IGQPGAVGPAGIR/1066-1078 (α 2)	-	-	-	-	-	-
C ₅₁ H ₈₃ N ₁₅ O ₂₂	GLTGSPPSPGPDGK/538-551(α 1)	629.79932	-0.30	629.79944	-0.49	629.79936	-0.36
C ₅₃ H ₉₀ N ₁₈ O ₁₈	GIPGPVGAAGATGAR(1oxi)/326-340 (α 2)	634.34119	0.17	634.34128	0.03	634.34122	0.13

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples (Continued)

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 8		Sample 9		Sample 10	
		Observed mass	Error ppm	Observed mass	Error ppm	Observed mass	Error ppm
C ₅₃ H ₉₀ N ₁₈ O ₁₈	GIPGPVGAAGATGAR(1oxi)/326-340 (α 2)	-	-	-	-	1267.67549	-0.13
C ₅₆ H ₈₆ N ₁₆ O ₁₉	GFPGSPGNIGPAGK/449-462 (α 2)	644.31982	0.33	644.32002	0.02	644.31982	0.33
C ₅₆ H ₈₅ N ₁₅ O ₂₀	GFPGSPGNIGPAGK/449-462 (α 2)	-	-	-	-	-	-
C ₅₉ H ₈₉ N ₁₅ O ₂₀	GFPGLPGPSGEPGK/970-983 (α 1)	664.82747	0.33	664.82790	-0.31	664.82761	0.12
C ₅₇ H ₈₉ N ₁₇ O ₂₀	GPSGPQGPSGPPGK/415-429 (α 1)	666.83079	-0.06	666.83100	-0.37	666.83119	-0.66
C ₅₉ H ₈₉ N ₁₅ O ₂₁	GFPGLPGPSGEPGK/970-983 (α 1)	672.82552	-0.55	672.82550	-0.52	672.82512	0.04
C ₆₃ H ₉₈ N ₁₈ O ₂₀	GIPGEFGLPGPAGAR(2oxi)/572-586 (α 2)	714.36755	-0.05	714.36773	-0.30	714.36739	0.17
C ₆₀ H ₉₄ N ₁₈ O ₂₃	GEPPAGLPGPSGER/472-486 (α 1)	718.34441	-0.24	718.34429	-0.07	718.34438	-0.20
C ₆₁ H ₉₄ N ₂₀ O ₂₂	GSAGPPGATGPPGAAGR/865-881 (α 1)	730.35013	-0.38	-	-	730.34982	0.05
C ₆₁ H ₉₂ N ₂₀ O ₂₃	GDGGPPGATGPPGAAGR/775-792 (α 2)	737.33936	0.17	737.33956	-0.10	737.33924	0.33
C ₆₅ H ₁₀₅ N ₂₁ O ₂₂	GEPPAGAVGPAGAVGPR/977-994 (α 2)	766.89477	-0.44	766.89433	0.13	766.89403	0.52
C ₆₇ H ₁₀₉ N ₂₁ O ₂₂	GETGPAGPAGPIGPVGAR/1066-1083 (α 1)	780.91051	-0.55	-	-	780.91006	0.02
C ₆₇ H ₁₀₇ N ₁₉ O ₂₄	DGLNGLPGPIGPPGPR/1152-1167 (α 1)	781.89451	-0.53	-	-	781.89385	0.31
C ₆₅ H ₁₀₅ N ₂₁ O ₂₅	GPPGESGAAGPTGPIGSR(1oxi)/590-607 (α 2)	-	-	-	-	790.88697	-0.21
C ₆₇ H ₁₀₃ N ₂₁ O ₂₄	GANGAPGIAGAPGPPGAR/397-414 (α 1)	793.88162	-0.13	793.88175	-0.29	793.88153	-0.02
C ₇₀ H ₁₁₃ N ₂₁ O ₂₅	GSTGEIGPAGPPGPPGLR(2oxi)/380-397 (α 2)	824.91815	-0.06	824.91850	-0.48	824.91764	0.56
C ₆₉ H ₁₀₇ N ₂₃ O ₂₈	DGEAGAQQGPPGAPAGER/612-630 (α 1)	853.89007	0.00	-	-	853.89008	-0.01
C ₇₈ H ₁₂₁ N ₂₃ O ₂₇	VGPPGPSNAGPPGPPGAGK(3oxi)/881-902 (α 1)	-	-	-	-	906.94742	-0.03
C ₇₉ H ₁₂₄ N ₂₂ O ₂₈	TGPPGPSGISGPPGPPGAGK(3oxi)/793-813 (α 2)	-	-	-	-	915.45463	0.46
C ₇₉ H ₁₂₄ N ₂₂ O ₂₉	TGPPGPSGISGPPGPPGAGK(4oxi)/793-813 (α 2)	-	-	-	-	923.45208	0.46
C ₈₂ H ₁₃₄ N ₂₈ O ₂₉	SGDRGETGPAGPAGPIGPVGAR/1062-1083 (α 1)	659.33634	-0.02	659.33646	-0.20	659.33671	-0.58

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples (Continued)

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 8		Sample 9		Sample 10	
		Observed mass	Error ppm	Observed mass	Error ppm	Observed mass	Error ppm
C ₈₄ H ₁₃₄ N ₂₆ O ₃₂	GEPGPTGIQGPPGAGEEGKR(2oxi)/448-468 (α 1)	673.99547	0.58	673.99640	-0.80	673.99604	-0.27
C ₈₉ H ₁₄₆ N ₂₈ O ₂₈	EGPVGLPGIDGRPGPIGPAGAR(1oxi)/463-484 (α 2)	686.03623	-0.35	686.03616	-0.25	686.03585	0.20
C ₈₅ H ₁₃₃ N ₂₉ O ₃₁	TGPPGPAGQDGRPGPPGPPGAR(4oxi)/552-573 (α 1)	686.33172	-0.53	-	-	686.33119	0.24
C ₈₅ H ₁₃₂ N ₂₈ O ₃₂	TGPPGPAGQDGRPGPPGPPGAR (4oxi)/552-573 (α 1)	-	-	-	-	-	-
C ₈₉ H ₁₄₆ N ₂₈ O ₂₉	EGPVGLPGIDGRPGPIGPAGAR(2oxi)/463-484 (α 2)	691.36742	0.30	691.36739	0.34	691.36742	-0.87
C ₉₂ H ₁₅₁ N ₂₇ O ₃₁	GLPGVAGSVGEPGPLGIAGPPGAR(3oxi)/881-904 (α 2)	-	-	-	-	1066.06137	-0.59
C ₁₀₅ H ₁₆₇ N ₃₁ O ₃₉	GPPGSAGSPGKDGLNGLPGPIGPGR/1141-1167 (α 1)	829.74214	-0.34	829.74226	-0.48	829.74176	0.12
C ₁₀₁ H ₁₆₃ N ₃₇ O ₃₇	AGEDGHPGKPRPGERGVVGPQGAR(3oxi)/139-163 (α 2)	-	-	-	-	829.74176	-0.96
C ₁₁₅ H ₁₈₁ N ₃₉ O ₄₂	GAPGAIGAPGPAGANGDRGEAGPAGPAGPR/674-706 (α 2)	927.78201	-0.50	927.78141	0.15	927.78113	0.45
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₃	GLTGPIGPPGAPAGPGDKGEAGPSGPAGPTGAR(2oxi)/763-795 (α 1)	951.80912	-0.08	951.80917	-0.13	951.80815	0.94
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₄	GLTGPIGPPGAPAGPGDKGEAGPSGPAGPTGAR(3oxi)/763-795 (α 1)	-	-	-	-	957.14073	-0.05

P: Hydroxyproline (Hyp), **N:** Deamidation of N (D), **Q:** Deamidation of Q (E)

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples (Continued)

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 12		Sample 14	
		Observed mass	Error ppm	Observed mass	Error ppm
C ₂₆ H ₄₅ N ₉ O ₁₀	GL P GER/304-309 (α 1)	644.33640	-0.29	-	-
C ₃₅ H ₅₇ N ₁₃ O ₁₁	GPAGPQGPR/1084-1092 (α 1)	418.72228	0.05	418.72227	0.07
C ₃₇ H ₅₈ N ₁₀ O ₁₃	GFSGLDGAK/268-276 (α 1)	426.21675	-0.55	-	-
C ₃₆ H ₆₁ N ₁₃ O ₁₂	GPSGPQGIR/994-1003 (α 2)	434.73523	0.39	-	-
C ₃₅ H ₅₉ N ₁₃ O ₁₄	GSEGPQGVR/361-369 (α 1)	-	-	-	-
C ₃₈ H ₆₂ N ₁₂ O ₁₃	GPAGPSGPAGK/1052-1062 (α 2)	448.23542	-0.40	-	-
C ₃₈ H ₆₇ N ₁₃ O ₁₂	G V VGL P QQR/958-966 (α 1)	449.75890	-0.06	449.75917	-0.66
C ₃₈ H ₆₇ N ₁₃ O ₁₂	G V VGL P QQR/958-966 (α 1)	-	-	-	-
C ₄₈ H ₇₃ N ₁₃ O ₁₆	G F P G ADGVAGPK/493-504 (α 1)	544.77215	0.07	544.77228	-0.17
C ₄₈ H ₇₃ N ₁₃ O ₁₆	G F P G ADGVAGPK(1oxi)/493-504 (α 1)	-	-	-	-
C ₄₄ H ₇₄ N ₁₈ O ₁₅	GPRGDQGPVGR/818-828 (α 2)	548.28642	-0.17	548.28597	0.65
C ₄₇ H ₇₆ N ₁₆ O ₁₅	GVQGP P GPAGPR(1oxi)/685-696 (α 1)	553.29085	0.41	553.29098	0.18
C ₅₁ H ₈₀ N ₁₄ O ₁₆ S ₁	GQAGVMGF P GP K (2oxi)/574-585 (α 1)	589.28665	0.85	-	-
C ₅₁ H ₇₉ N ₁₃ O ₁₇ S ₁	G Q AGVMGF P GP K /574-575 (α 1)	589.77878	0.64	589.77879	0.62
C ₅₂ H ₈₅ N ₁₅ O ₁₇	GVP G PPGAVGPAGK(2oxi)/598-611 (α 1)	596.81938	0.49	596.81957	0.17
C ₅₃ H ₈₀ N ₁₄ O ₁₈	G E PGNIG F GP K /485-496 (α 2)	601.29628	-0.42	-	-
C ₅₃ H ₇₉ N ₁₃ O ₁₉	G E PG N IG F GP K /485-496 (α 2)	601.78774	0.49	601.78769	0.57
C ₅₂ H ₈₉ N ₁₇ O ₁₆	IG Q PGAVGPAGIR/1066-1078 (α 2)	604.84056	0.62	604.84061	0.54
C ₅₂ H ₈₈ N ₁₆ O ₁₇	IG Q PGAVGPAGIR/1066-1078 (α 2)	605.33278	0.27	605.33289	0.09
C ₅₁ H ₈₃ N ₁₅ O ₂₂	GLTGS P GS P GP D GK/538-551(α 1)	629.79894	0.30	-	-
C ₅₃ H ₉₀ N ₁₈ O ₁₈	GIPGPVGAAGATGAR(1oxi)/326-340 (α 2)	634.34102	0.44	634.34088	0.66

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples (Continued)

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 12		Sample 14	
		Observed mass	Error ppm	Observed mass	Error ppm
C ₅₃ H ₉₀ N ₁₈ O ₁₈	GIPGPVGAAGATGAR(1oxi)/326-340 (α 2)	-	-	-	-
C ₅₆ H ₈₆ N ₁₆ O ₁₉	GFPGSPGNIGPAGK/449-462 (α 2)	644.31971	0.50	-	-
C ₅₆ H ₈₅ N ₁₅ O ₂₀	GFPGSPGNIGPAGK/449-462 (α 2)	644.81200	0.07	-	-
C ₅₉ H ₈₉ N ₁₅ O ₂₀	GFPGLPGPSGEPGK/970-983 (α 1)	664.82708	0.92	-	-
C ₅₇ H ₈₉ N ₁₇ O ₂₀	GPSGPQGPSGPPGPK/415-429 (α 1)	666.83097	-0.33	-	-
C ₅₉ H ₈₉ N ₁₅ O ₂₁	GFPGLPGPSGEPGK/970-983 (α 1)	-	-	-	-
C ₆₃ H ₉₈ N ₁₈ O ₂₀	GIPGEFGLPGPAGAR(2oxi)/572-586 (α 2)	714.36729	0.31	714.36723	0.40
C ₆₀ H ₉₄ N ₁₈ O ₂₃	GEPGPAGLPGPAGER/472-486 (α 1)	718.34424	0.00	718.34398	0.36
C ₆₁ H ₉₄ N ₂₀ O ₂₂	GSAGPPGATGFPGAAGR/865-881 (α 1)	730.34969	0.22	730.34915	0.96
C ₆₁ H ₉₂ N ₂₀ O ₂₃	GDGGPPGATGFPGAAGR/775-792 (α 2)	737.33949	-0.01	-	-
C ₆₅ H ₁₀₅ N ₂₁ O ₂₂	GEPGPAGAVGPAGAVGPR/977-994 (α 2)	-	-	-	-
C ₆₇ H ₁₀₉ N ₂₁ O ₂₂	GETGPAGPAGPIGPVGAR/1066-1083 (α 1)	-	-	-	-
C ₆₇ H ₁₀₇ N ₁₉ O ₂₄	DGLNGLPGPIGPPGPR/1152-1167 (α 1)	781.89471	-0.79	-	-
C ₆₅ H ₁₀₅ N ₂₁ O ₂₅	GPPGESGAAGPTGPIGSR(1oxi)/590-607 (α 2)	-	-	-	-
C ₆₇ H ₁₀₃ N ₂₁ O ₂₄	GANGAPGIAGAPGPPGAR/397-414 (α 1)	793.88181	-0.37	793.88114	0.48
C ₇₀ H ₁₁₃ N ₂₁ O ₂₅	GSTGEIGPAGPPGPPGLR(2oxi)/380-397 (α 2)	-	-	-	-
C ₆₉ H ₁₀₇ N ₂₃ O ₂₈	DGEAGAQQGPPGAPAGER/612-630 (α 1)	-	-	-	-
C ₇₈ H ₁₂₁ N ₂₃ O ₂₇	VGPPGPSGNAGPPGPPGAGK(3oxi)/881-902 (α 1)	-	-	-	-
C ₇₉ H ₁₂₄ N ₂₂ O ₂₈	TGPPGPSGISGPPGPPGAGK(3oxi)/793-813 (α 2)	-	-	-	-
C ₇₉ H ₁₂₄ N ₂₂ O ₂₉	TGPPGPSGISGPPGPPGAGK(4oxi)/793-813 (α 2)	-	-	-	-
C ₈₂ H ₁₃₄ N ₂₈ O ₂₉	SGDRGETGPAGPAGPIGPVGAR/1062-1083 (α 1)	659.33619	0.21	659.33575	0.88

TABLE 4.2: Sequences and mass error of the peptides assigned across the range of bone samples (Continued)

Elemental composition	Identified peptides/Length coll-I(chain)	Sample 12		Sample 14	
		Observed mass	Error ppm	Observed mass	Error ppm
C ₈₄ H ₁₃₄ N ₂₆ O ₃₂	GEPGPTGIQGGPPGAGEEGKR(2oxi)/448-468 (α 1)	673.99585	0.01	-	-
C ₈₉ H ₁₄₆ N ₂₈ O ₂₈	EGPVGLPGIDGRPGPIGPAGAR(1oxi)/463-484 (α 2)	686.03617	-0.26	686.03541	0.84
C ₈₅ H ₁₃₃ N ₂₉ O ₃₁	TGPPGPAGQDGRPGPPGPPGAR(4oxi)/552-573 (α 1)	686.33120	0.22	-	-
C ₈₅ H ₁₃₂ N ₂₈ O ₃₂	TGPPGPAGQDGRPGPPGPPGAR (4oxi)/552-573 (α 1)	-	-	686.66	0.67
C ₈₉ H ₁₄₆ N ₂₈ O ₂₉	EGPVGLPGIDGRPGPIGPAGAR(2oxi)/463-484 (α 2)	691.36769	-0.09	691.36729	0.49
C ₉₂ H ₁₅₁ N ₂₇ O ₃₁	GLPGVAGSVGEPGPLGIAGPPGAR(3oxi)/881-904 (α 2)	-	-	-	-
C ₁₀₅ H ₁₆₇ N ₃₁ O ₃₉	GPPGSAGSPGKDGLNGLPGPIGPPGPR/1141-1167 (α 1)	829.74217	-0.38	829.74265	-0.95
C ₁₀₁ H ₁₆₃ N ₃₇ O ₃₇	AGEDGHPGKPRPGERGVVGPQGAR(3oxi)/139-163 (α 2)	-	-	-	-
C ₁₁₅ H ₁₈₁ N ₃₉ O ₄₂	GAPGAIGAPGPAGANGDRGEAGPAGPAGPAGPR/674-706 (α 2)	927.78241	-0.99	927.78172	-0.25
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₃	GLTGPIGPPGAPAGPGDKGEAGPSGPAGPTGAR(2oxi)/763-795 (α 1)	951.80899	0.05	951.80918	-0.14
C ₁₂₁ H ₁₉₃ N ₃₇ O ₄₄	GLTGPIGPPGAPAGPGDKGEAGPSGPAGPTGAR(3oxi)/763-795 (α 1)	-	-	-	-

P: Hydroxyproline (Hyp), **N:** Deamidation of N (D), **Q:** Deamidation of Q (E)

TABLE 4.3: List of assigned ions from the ECD spectrum of m/z [793.88150]²⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
401.21431	401.21430	z ₄	-0.03	C_Pos_4
548.28272	548.28306	z ₅	0.62	C_Pos_5
788.38114	788.38145	z ₈ [•]	0.39	C_Pos_8
789.38897	789.38901	z ₈	0.05	C_Pos_8
793.88150	793.88160	[M+2H] ²⁺	0.13	Parent ion
845.40260	845.40271	z ₉ [•]	0.13	C_Pos_9
846.41043	846.41045	z ₉	0.02	C_Pos_9
861.42132	861.42106	y ₉	-0.30	C_Pos_9
916.43971	916.43972	z ₁₀ [•]	0.01	C_Pos_10
917.44754	917.44756	z ₁₀	0.02	C_Pos_10
1029.52377	1029.52371	z ₁₁ [•]	-0.06	C_Pos_11
1030.53160	1030.53162	z ₁₁	0.02	C_Pos_11
1086.54523	1086.54496	z ₁₂ [•]	-0.25	C_Pos_12
1087.55306	1087.55314	z ₁₂	0.07	C_Pos_12
1102.56395	1102.56374	y ₁₂	-0.19	C_Pos_12
1270.63002	1270.62995	z ₁₄ [•]	-0.06	C_Pos_14
1327.65148	1327.65136	z ₁₅ [•]	-0.09	C_Pos_15
1357.62565	1357.62495	c ₁₆ [•]	-0.51	N_Pos_16
1385.68020	1385.68101	z ₁₆ [•] -57	0.59	N_Pos_16
1428.66276	1428.66300	c ₁₇ [•]	0.17	N_Pos_17
1442.67842	1442.67843	z ₁₆ [•]	0.01	C_Pos_16
1513.71553	1513.71579	z ₁₇ [•]	0.17	C_Pos_17
1527.74297	1527.74192	[M+2H] ⁺ -60	0.69	Asp diagnostic ion
1526.74719	1526.74753	z ₁₈ [•] - CO ₂	-0.22	IsoAsp diagnostic ion
1570.73699	1570.73732	z ₁₈ [•]	-0.21	C_Pos_18
1586.75571	1586.75551	[M+H] ⁺	0.13	Single charge Parent
1587.76407	1587.76413	[M+2H] ⁺ •	-0.04	Charged reduced species
		* Mean:	0.05	
		* Std. Dev.:	0.27	

TABLE 4.4: List of assigned ions from the CAD spectrum of m/z [793.88150]²⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
354.14084	354.14080	b ₅ -H ₂ O	0.10	N_Pos_5
372.15136	372.15135	b ₅	-0.01	N_Pos_5
416.22520	416.22519	y ₄	-0.02	C_Pos_4
467.18847	467.18852	b ₆ -H ₂ O	0.11	N_Pos_6
524.20993	524.20993	b ₇ -H ₂ O	0.00	N_Pos_7
542.22050	542.22046	b ₇	-0.07	N_Pos_7
563.29361	563.29359	y ₅	-0.04	C_Pos_5
608.30945	608.30946	y ₁₃ ²⁺	0.01	C_Pos_13
619.28342	619.28343	b ₈ -2H ₂ O	0.01	N_Pos_8
620.31507	620.31508	y ₆	0.02	C_Pos_6
637.29399	637.29401	b ₈ -H ₂ O	0.03	N_Pos_8
643.82801	643.82802	y ₁₄ ²⁺	0.02	C_Pos_14
655.30456	655.30456	b ₈	0.01	N_Pos_8
672.33874	672.33879	y ₁₅ ²⁺	0.07	C_Pos_15
708.33110	708.33112	b ₉ -H ₂ O	0.03	N_Pos_9
726.34167	726.34169	b ₉	0.03	N_Pos_10
729.85221	729.85220	y ₁₆ ²⁺	-0.01	C_Pos_16
733.36275	733.36273	y ₇	-0.03	C_Pos_7
765.35256	765.35255	b ₁₀ -H ₂ O	-0.01	N_Pos_10
783.36312	783.36316	b ₁₀	0.05	N_Pos_10
784.87621	784.87623	b ₁₈ ²⁺	0.02	N_Pos_18
793.88150	793.88151	[M+2H]2+	0.01	Parent ion
804.39986	804.39989	y ₈	0.04	C_Pos_8
818.37910	818.37913	b ₁₁ -2H ₂ O	0.03	N_Pos_11
836.38967	836.38971	b ₁₁ -H ₂ O	0.05	N_Pos_11
854.40023	854.40025	b ₁₁	0.02	N_Pos_11
861.42132	861.42133	y ₉	0.01	C_Pos_9
932.45843	932.45844	y ₁₀	-0.01	C_Pos_10
1024.46937	1024.46963	b ₁₃	0.25	N_Pos_13
1045.54249	1045.54252	y ₁₁	0.03	C_Pos_11
1102.56395	1102.56402	y ₁₂	0.06	C_Pos_12
1135.51665	1135.51586	b ₁₄ -2H ₂ O	-0.70	N_Pos_14
1153.52722	1153.52718	b ₁₄ -H ₂ O	-0.04	N_Pos_14
1171.53679	1171.53778	b ₁₄	0.85	N_Pos_14
1215.61163	1215.61162	y ₁₃	0.01	C_Pos_13
		* Mean:	0.03	
		* Std. Dev.:	0.19	

TABLE 4.5: List of assigned ions from the ECD spectrum of m/z [829.74184]³⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
314.18227	314.18227	z_3	0.00	C_Pos_3
429.20921	429.20921	c_5	0.00	N_Pos_5
500.24632	500.24634	c_6	-0.04	N_Pos_6
557.26778	557.26778	c_7	0.00	N_Pos_7
814.36895	814.36898	c_{10}	-0.04	N_Pos_10
829.74184	829.74176	[M+3H] ³⁺	0.10	Parent ion
863.44955	863.44946	z_9^\bullet	0.10	C_Pos_9
864.45737	864.45738	z_9	-0.01	C_Pos_9
941.45609	941.45614	c_{11}^\bullet	-0.05	N_Pos_11
942.46391	942.46396	c_{11}	-0.05	N_Pos_11
1057.49085	1057.49075	c_{12}	0.09	N_Pos_12
1089.58129	1089.58114	z_{11}^\bullet	0.14	C_Pos_11
1114.51231	1114.51230	c_{13}	0.01	N_Pos_13
1146.60275	1146.60285	z_{12}^\bullet	-0.09	C_Pos_12
1147.61057	1147.61074	z_{12}	-0.15	C_Pos_12
1162.62147	1162.62149	y_{12}	-0.02	C_Pos_12
1184.59247	1184.59142	[M+3H] ²⁺ -60	0.89	Asp diag. ion
1227.59637	1227.59614	c_{14}	0.19	N_Pos_14
1236.09976	1236.09953	$z_{27}^{2+\bullet}$	0.19	C_Pos_27
1244.61331	1244.61287	[M+3H] ²⁺ $^\bullet$	0.35	Charged reduced species
1261.62969	1261.62967	z_{13}^\bullet	0.02	C_Pos_13
1284.59395	1284.59434	$c_{14}^\bullet + 58$	-0.30	N_Pos_14
1342.62331	1342.62331	c_{15}	0.00	N_Pos_15
1374.71375	1374.71376	z_{14}^\bullet	-0.01	C_Pos_14
1399.64477	1399.64428	c_{16}	0.35	N_Pos_16
1431.73521	1431.73512	z_{15}^\bullet	0.06	C_Pos_15
1447.75393	1447.75342	y_{15}	0.35	C_Pos_15
1546.76215	1546.76217	z_{16}^\bullet	-0.01	C_Pos_16
1624.76869	1624.76812	c_{18}^\bullet	0.35	N_Pos_18
1625.77651	1625.77690	c_{18}	-0.24	N_Pos_18
1931.95828	1931.95779	z_{20}^\bullet	0.25	C_Pos_20
		* Mean:	0.08	
		* Std. Dev.:	0.22	

TABLE 4.6: List of assigned ions from the CAD spectrum of m/z [829.74184]³⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
412.18266	412.18266	b ₅	0.00	N_Pos_5
465.20925	465.20920	b ₆ -H ₂ O	0.11	N_Pos_6
483.21977	483.21977	b ₆	0.00	N_Pos_6
496.76161	496.76160	y ₁₀ ²⁺	0.02	C_Pos_10
522.23072	522.23063	b ₇ -H ₂ O	0.16	N_Pos_7
540.24123	540.24123	b ₇	0.00	N_Pos_7
553.30365	553.30360	y ₁₁ ²⁺	0.08	C_Pos_11
555.28852	555.28850	a ₈ -2CO ₂	0.04	N_Pos_8
570.25181	570.25174	a ₈ -CH ₄ N ₂	0.11	N_Pos_8
591.25223	591.25215	b ₈ -2H ₂ O	0.14	N_Pos_8
596.78356	596.78337	[b ₁₄ -H ₂ O] ²⁺	0.32	N_Pos_14
605.78855	605.78849	b ₁₄ ²⁺	0.10	N_Pos_14
609.26275	609.26269	b ₈ -H ₂ O	0.09	N_Pos_8
627.27326	627.27325	b ₈	0.02	N_Pos_8
639.32785	639.32789	y ₁₃ ²⁺	-0.07	C_Pos_13
654.29703	654.29670	[b ₁₅ -H ₂ O] ²⁺	0.51	N_Pos_15
663.30202	663.30200	b ₁₅ ²⁺	0.03	N_Pos_15
682.80776	682.80750	[b ₁₆ -H ₂ O] ²⁺	0.38	N_Pos_16
721.33894	721.33894	[b ₁₇ -3H ₂ O] ²⁺	-0.01	N_Pos_17
724.38061	724.38062	y ₁₅ ²⁺	-0.02	C_Pos_15
730.34454	730.34422	[b ₁₇ -2H ₂ O] ²⁺	0.43	N_Pos_17
734.35733	734.35736	a ₁₇ ²⁺	-0.05	N_Pos_17
739.34979	739.34947	[b ₁₇ -H ₂ O] ²⁺	0.44	N_Pos_17
748.35478	748.35476	b ₁₇ ²⁺	0.03	N_Pos_17
822.44681	822.44683	y ₈	-0.02	C_Pos_8
823.73832	823.73827	b ₂₇ ³⁺	0.06	N_Pos_27
829.74184	829.74180	[M+3H] ³⁺	0.05	Parent ion
879.46827	879.46824	y ₉	0.03	C_Pos_9
925.43736	925.43739	b ₁₁	-0.03	N_Pos_11
930.976125	930.97607	y ₁₉ ²⁺	0.06	C_Pos_19
938.457765	938.45781	[a ₂₂ -C ₂ H ₃] ²⁺	-0.05	N_Pos_22
974.49214	974.49213	y ₂₀ ²⁺	0.01	C_Pos_20
992.51595	992.51593	y ₁₀	0.02	C_Pos_10
1003.00287	1003.0029	y ₂₁ ²⁺	-0.03	C_Pos_21
1038.521425	1038.52158	y ₂₂ ²⁺	-0.15	C_Pos_22
1040.4643	1040.46425	b ₁₂	0.05	N_Pos_12
1082.03744	1082.03745	y ₂₃ ²⁺	-0.01	C_Pos_23

TABLE 4.6: List of assigned ions from the CAD spectrum of m/z [829.74184]³⁺

(Continued)

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
1158.067023	1158.06675	[y ₂₅ -H ₂ O] ²⁺	0.24	C_Pos_25
1162.62147	1162.62143	y ₁₂	0.03	C_Pos_12
1167.07201	1167.07201	y ₂₅ ²⁺	0.00	C_Pos_25
1192.55931	1192.55918	b ₁₄ -H ₂ O	0.10	N_Pos_14
1210.56982	1210.56988	b ₁₄	-0.05	N_Pos_14
1277.64841	1277.64843	y ₁₃	-0.02	C_Pos_13
1325.59676	1325.59710	b ₁₅	-0.26	N_Pos_15
1382.61822	1382.61838	b ₁₆	-0.12	N_Pos_16
1447.75393	1447.75389	y ₁₅	0.03	C_Pos_15
1459.68125	1459.68110	b ₁₇ -2H ₂ O	0.10	N_Pos_17
1477.691765	1477.69171	b ₁₇ -H ₂ O	0.04	N_Pos_17
1495.70228	1495.70258	b ₁₇	-0.20	N_Pos_17
		* Mean:	-0.05	
		* Std. Dev.:	0.15	

TABLE 4.7: List of assigned ions from the ECD spectrum of m/z [927.78145]³⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
314.18228	314.18227	z_3	-0.03	C_Pos_3
385.21939	385.21937	z_4^\bullet	-0.05	C_Pos_4
538.28578	538.28585	z_6^\bullet	0.13	C_Pos_6
539.29359	539.29365	z_6	0.10	C_Pos_6
609.32289	609.32282	z_7^\bullet	-0.12	C_Pos_7
741.38895	741.38886	c_9	-0.12	N_Pos_9
763.39711	763.39712	z_9^\bullet	0.01	C_Pos_9
779.41583	779.41553	y_9	-0.39	C_Pos_9
834.43422	834.43398	z_{10}^\bullet	-0.29	C_Pos_10
895.46317	895.46314	c_{11}	-0.03	N_Pos_11
966.50028	966.50051	c_{12}	0.24	N_Pos_12
988.50844	988.50835	z_{12}^\bullet	-0.09	C_Pos_12
1023.52174	1023.52237	c_{13}	0.62	N_Pos_13
1059.54555	1059.54565	z_{13}^\bullet	0.09	C_Pos_13
1093.55103	1093.55058	c_{14}^\bullet	-0.41	N_Pos_14
1094.55885	1094.55897	c_{14}	0.11	N_Pos_14
1188.58814	1188.58816	z_{14}^\bullet	0.02	C_Pos_14
1266.60725	1266.60674	c_{16}	-0.40	N_Pos_16
1381.63419	1381.63394	c_{17}	-0.18	N_Pos_17
1383.15918	1383.15850	$z_{33}^{2+\bullet}$	-0.49	C_Pos_33
1391.67273	1391.67228	$[M+3H]^{2+\bullet}$	-0.32	Charged reduced species
1402.71854	1402.71801	z_{16}	-0.37	C_Pos_16
1517.74547	1517.74682	z_{17}	0.89	C_Pos_17
1594.75676	1594.75606	c_{19}	-0.44	N_Pos_19
1689.79388	1689.79273	z_{19}	-0.68	C_Pos_19
1794.83646	1794.83593	c_{21}	-0.30	N_Pos_21
* Mean:			-0.10	
* Std. Dev.:			0.33	

TABLE 4.8: List of assigned ions from the CAD-MS² spectrum of m/z [927.78145]³⁺

Theoretical mass	Experimental mass	Fragment	Error (ppm)	Position
370.1721	370.17209	b ₅	-0.01	N_Pos_5
455.26124	455.26119	a ₆	-0.11	N_Pos_6
465.24559	465.24552	b ₆ -H ₂ O	-0.15	N_Pos_6
483.25615	483.25617	b ₇	0.03	N_Pos_7
522.26705	522.26698	b ₇ -H ₂ O	-0.13	N_Pos_7
540.27761	540.27747	b ₇	-0.27	N_Pos_7
554.3045	554.30448	y ₆	-0.04	C_Pos_6
583.31981	583.31986	a ₈	0.09	N_Pos_8
593.30416	593.30413	b ₈ -H ₂ O	-0.05	N_Pos_8
611.31472	611.3147	b ₈	-0.04	N_Pos_8
709.36835	709.36836	y ₁₆ ²⁺	0.01	C_Pos_16
722.39437	722.39435	y ₈	-0.03	C_Pos_8
779.41583	779.41587	y ₉	0.05	C_Pos_9
786.38727	786.38709	[y ₁₈ -H ₂ O] ²⁺	-0.23	C_Pos_18
795.39256	795.39255	y ₁₈ ²⁺	-0.01	C_Pos_18
843.90074	843.90059	[y ₁₉ -H ₂ O] ²⁺	-0.18	C_Pos_19
852.90603	852.90597	y ₁₉ ²⁺	-0.06	C_Pos_19
885.09526	885.09521	y ₃₁ ³⁺	-0.06	C_Pos_31
916.93531	916.93543	y ₂₁ ²⁺	0.13	C_Pos_21
947.5057	947.5057	y ₁₁	0.00	C_Pos_11
1000.98025	1000.98024	y ₂₃ ²⁺	-0.01	C_Pos_23
1004.52716	1004.52719	y ₁₂	0.03	C_Pos_12
1020.48569	1020.4862	[y ₂₄ -H ₂ O] ²⁺	0.50	C_Pos_24
1029.49097	1029.49096	y ₂₄ ²⁺	-0.02	C_Pos_24
1068.00425	1068.00394	[y ₂₅ -2H ₂ O] ²⁺	-0.29	C_Pos_25
1075.56427	1075.56439	y ₁₃	0.11	C_Pos_13
1077.00953	1077.00971	[y ₂₅ -H ₂ O] ²⁺	0.17	C_Pos_25
1086.01482	1086.01483	y ₂₅ ²⁺	0.01	C_Pos_25
1112.52809	1112.52856	[y ₂₆ -H ₂ O] ²⁺	0.43	C_Pos_26
1121.53337	1121.53383	y ₂₆ ²⁺	0.41	C_Pos_26
1141.03882	1141.03887	[y ₂₇ -H ₂ O] ²⁺	0.05	C_Pos_27
1150.0441	1150.04406	y ₂₇ ²⁺	-0.04	C_Pos_27
1206.58613	1206.58618	y ₂₈ ²⁺	0.04	C_Pos_28
1242.10469	1242.10512	y ₂₉ ²⁺	0.35	C_Pos_29
1261.62832	1261.62848	y ₁₅	0.13	C_Pos_15
1270.61542	1270.61454	y ₃₀ ²⁺	-0.69	C_Pos_30
1400.70343	1400.70321	y ₁₆ -NH ₃	-0.16	C_Pos_16
1417.72943	1417.72947	y ₁₆	0.03	C_Pos_16
		* Mean:	-0.11	
		* Std. Dev.:	0.20	

Chapter 5: Conclusions and future directions

5.1. CONCLUSIONS

The work presented in this thesis illustrates the use of FT-ICR-MS to investigate Asn deamidated marker peptides identified in archaeological ancient bones and the determination of their extent of deamidation, which could be further implemented for relative age determination. As highlighted throughout the thesis deamidation and isomerisation of Asn has gained interest from the archaeological point of view. So far the progress in mass spectrometry fragmentation techniques has allowed the confident analysis of these types of modifications in samples such as bone, tissue, blood, etc. For example ECD was applied to characterize the different deamidated peptides present in ancient bone samples, and also was useful to differentiate the isomers produced in the deamidation reaction.

Deamidation of Asn in collagen represents a potential alternative to measure the relative age of ancient materials. However, so far the fact of artificially induced deamidation by sample prep has kept researchers away from using this modification for this purpose. In this thesis, a methodology was developed to address this issue by determine the extent of deamidation before sample preparation. Some of the most predominant deamidated peptides were observed throughout a range of different bovine bone samples, these peptides were selected as markers for bovine and were also chosen for deamidation studies. The extent of deamidation was calculated successfully for these peptides.

It is proposed that the deamidated peptides identified in the study are potential age markers in these types of samples. Their presence was further investigated in older samples. One particular peptide was consistently present in all the analysis. Perhaps, this peptide was continually used and referred to throughout the thesis.

A comparison of the extent of deamidation from young collagen (< 10 years) to ancient collagen ($\approx 1,000$ years) was established, the general trend was clearly pointing that the older collagen presented an elevated percent of deamidation (36%) compared to its younger version (6%). This preliminary study was discussed with collaborators in the area of archaeology, where they highlighted the importance of this finding towards age dating.

Since amino acid reactions are temperature dependent, this factor was also considered. The approach was to calculate the activation energy of the deamidation reaction, which has to be taken into account in future calculation of age involving deamidation of Asn. The value of activation energy obtained was considerably lower than previous reports, being a possible explanation that it was calculated for one peptide and not for a whole population.

An alternative way of calculating the extent of deamidation was introduced in Chapter 4. This method can only be used when IsoAsp/Asp diagnostic ions show up in the spectrum. The advantage of using IsoAsp/Asp peak intensities is that it can be performed when high resolution cannot be achieved, and the method is also useful for ion trap analysis using ETD.

In conclusion, the author is confident that the results of the work presented in this thesis have introduced a valuable contribution that can be further incorporated to calculations of relative age of ancient materials. Another major contribution, which was highlighted in numerous conferences by pharmaceutical companies, was that a way of calculating how much artificial deamidation introduced by sample preparation was addressed in the study.

5.2. FUTURE DIRECTIONS

The next step in this project will be to map out the extent of deamidation of the age marker peptides against the radio carbon age of the sample. This graph will give preliminary information of the relationship between age and deamidation of collagen.

Either the mass defect method or the IsoAsp+Asp peak intensities method using ECD can be useful to determine the extent of deamidation to build the graph. The preliminary studies exploring the use of ECD have been performed, but a larger number of samples would be necessary to completely characterize this last methodology.

Since amino acid reactions are temperature dependent the data obtained from the extent of deamidation needs to be further calibrated. In order to achieve this, more temperature measurements are needed to accurately calculate the activation energy of the reaction. The author proposes to calibrate the data by using an assumed temperature history of the sample. A plot of the adjusted extent of deamidation at 10 °C against the thermal age of the sample is needed. Thermal age is defined as the time taken to produce a given degree of DNA degradation when temperature is held at a constant 10 °C. The thermal age adjusts the chronological sites according to their individual thermal histories, using the known temperature dependence of DNA depurination estimated in aqueous solution. An estimate of the thermal age of the sample can be obtained from the following reference software <http://beta.thermal-age.eu/> developed by archaeologist Matthew Collins and collaborators.

Appendix

Use of High Resolution Mass Spectrometry in the Analysis of Polymeric Excipients in Drug Delivery Formulations

¹This appendix has been partially/entirely reproduced from
Perez P., Lam P., Kilgour D., Bristow A., McBride E., O'Connor P., Use of High
Resolution Mass Spectrometry for the Analysis of Polymeric Excipients in Drug
Formulations. *Analytical Chemistry*, **2012**. 84 (20), 8579–8586

A.1. INTRODUCTION

Improving the bioavailability of poorly water soluble drugs is one of the most significant challenges for the pharmaceutical industry, and a wide variety of different excipients are used to overcome this issue. Enabling formulations often employ polymeric excipients to enhance solubility and therefore bioavailability; two of which are Gelucire[®] 44/14 and polysorbate 80; these are known to improve solubility of poorly-water-soluble drugs, and hence, increase their bioactivities.¹⁻³ In addition to the use of Gelucire[®] 44/14 and polysorbate 80 as excipients in drugs, they are also widely used as cosmetic⁴ and food additives.⁵ Gelucire[®] 44/14 and polysorbate 80 can improve solubility of the drug within the formulation and *in vivo* through association with the drug molecule, for example by micelle formation in the intestinal environment. Therefore the excipient composition and variability can be critical to product performance and drug absorption. Gelucire[®] 44/14 is formed by polyglycolysis of hydrogenated palm kernel oil with polyethylene glycol (PEG) 1500. Gelucire[®] 44/14 is a semisolid wax, which is composed of approximately 72% mono- and di-fatty acid esters of polyethylene glycol (PEG) 1500; 20% mono-, di, and triglycerides and 8% free PEG (see Fig. A.1 for structures).^{1, 6} Different microscopy, NMR, x-ray diffraction, and HPLC-MS techniques have been used for characterisation of Gelucire[®] 44/14.^{1,7-9}

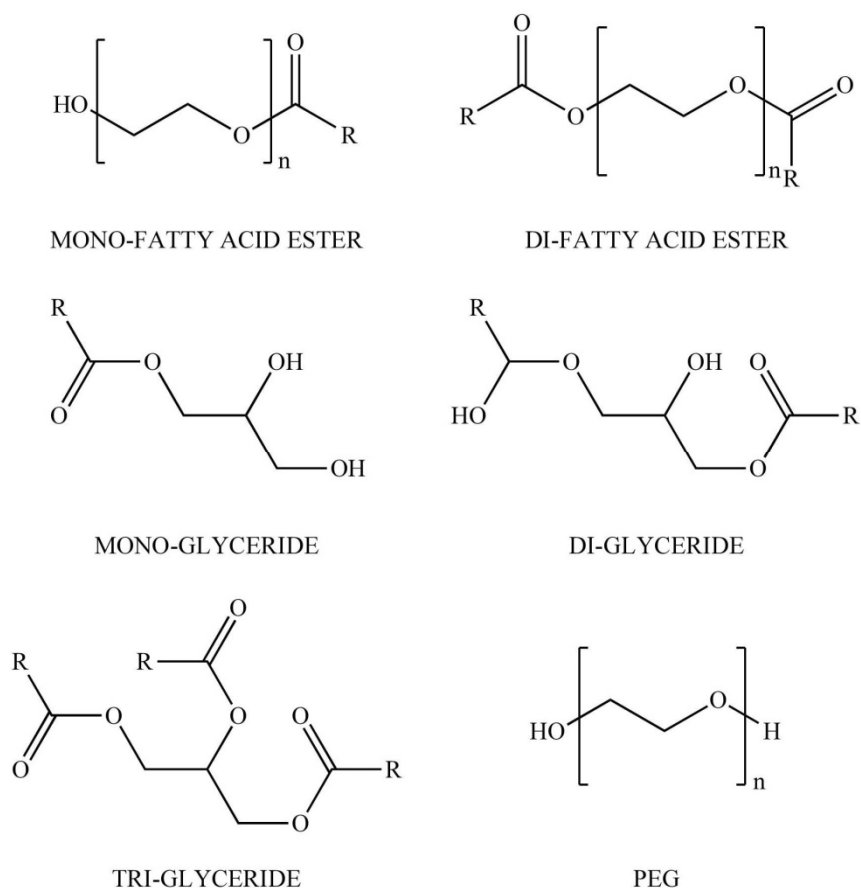
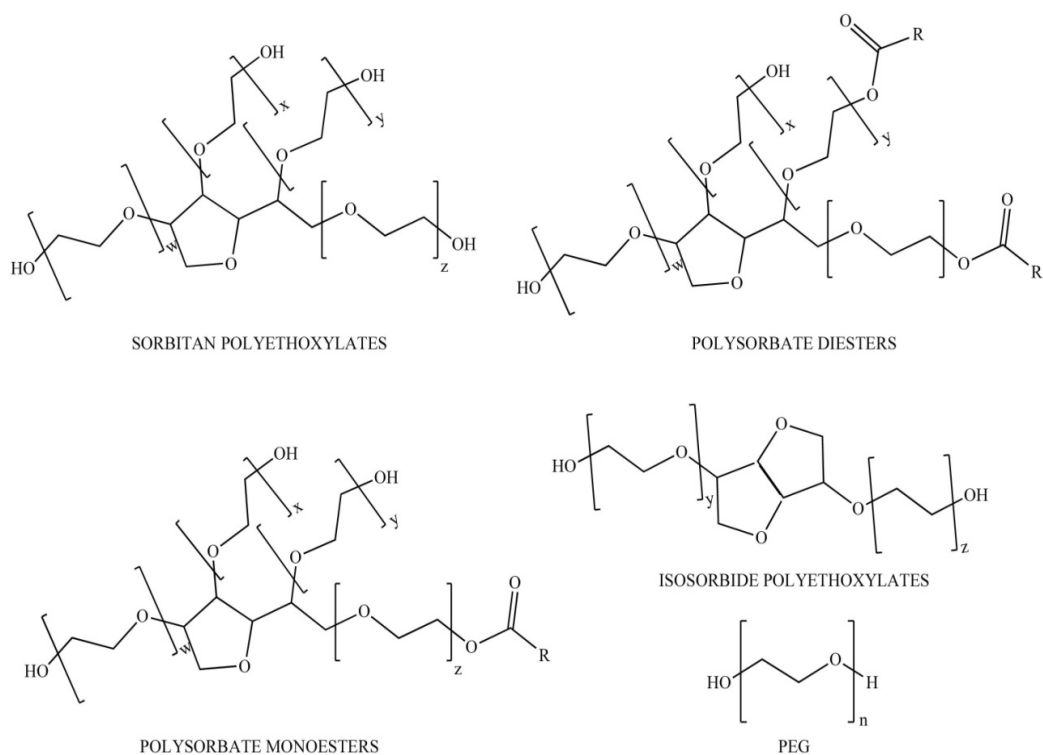


FIGURE A.1: Expected Compounds in Gelucire® 44/14.

Polysorbate 80 is formed by the reaction of fatty acids (commonly stearic ($\text{C}_{18}\text{H}_{36}\text{O}_2$), oleic ($\text{C}_{18}\text{H}_{34}\text{O}_2$) or linoleic ($\text{C}_{18}\text{H}_{32}\text{O}_2$) acid) with sorbitol in the presence of an acid catalyst.¹⁰ The products of this reaction are a mixture of sorbitan polyethoxylates, polysorbate diesters, polysorbate monoesters, isosorbide polyethoxylates and free PEG (see Fig. A.2 for structures). Infrared spectrophotometry (IR), thin-layer chromatography (TLC), gas chromatography (GC), and high performance liquid chromatography (HPLC) have been used for analysis of polysorbate 80.^{10, 11} In the 2000s, components of Polysorbate 80 were investigated by MALDI-TOF MS,¹² indicating the presence of polyoxyethylene isosorbide and sorbitan polyethoxylate monolinoleate. However, neither experiment

could differentiate sorbitan polyethoxylate (1507.84 m/z), sorbitol polyethoxylate monooleate (1507.93 m/z) and sorbitol polyethoxylate dioleate (1507.02 m/z) because of the proximity in mass (± 0.1 Da) (see Fig. A.2). Furthermore, the heterogeneity of polysorbates makes characterization of these molecules a significant analytical challenge. The use of LC-MS has enhanced the characterization of such molecules.⁹ However, to date, there is no study reported in the literature using high resolution mass spectrometry for the analysis of Gelucire[®] 44/14 and Polysorbate 80, although, FT-ICR has been commonly applied for structural analysis of other polymers.¹³⁻²² This study aims to identify the different components that can be found in the batches of polysorbate 80 and Gelucire[®] 44/14 respectively using FT-ICR-MS, to provide a better understanding about the potential impact of variability in these complex excipients on drug product performance.



A.2: Expected Compounds in Polysorbate 80.

A.2. EXPERIMENTAL SECTION

A.2.1. Materials

Two different batches of Gelucire[®] 44/14 (a product of Gattefosse S.A., St Priest, France), identified as (A and B), and polysorbate 80, identified as [Acros (A) and Fisher (B)], were provided by Astrazeneca. Chloroform, HPLC grade methanol and sodium hydroxide were obtained from Fisher Scientific, Leicestershire, UK.

A.2.2. Sample Preparation

Sample preparation was carried out by master student Yuko Lam. Gelucire[®] 44/14 samples were dissolved in 1:9 Chloroform/MeOH. Polysorbate 80 samples were dissolved in HPLC grade methanol. All solutions were then further diluted in 1 mM sodium hydroxide down to an approximate final concentration of (1 μ M).

A.2.3. FT-ICR-MS

The experiments were carried out on a Bruker Solarix 12 T FT-ICR-MS (Bruker Daltonics, Bremen, Germany). Solutions of Gelucire[®] 44/14 and Polysorbate 80 (described above) were electrosprayed at a flow rate of 300 μ L/h. An average of 400 scans using ESI in positive ion mode were obtained. The mass spectrometer was externally calibrated using a standard mixture (HP mix) within the range of interest. External calibration of mass spectra produced a mass accuracy of ≤ 5 ppm. Internal calibration was also performed using confidently assigned PEGs as internal calibrants, improving the mass accuracy to values < 1 ppm. Ion accumulation time varied from 0.1 to 0.5 sec depending on the sample. Front and back trap voltages were kept about 0.45 V. Electrospray voltage was kept at 4.5 kV. The mass range for broadband detection was (200 – 3000 m/z).

A.3. RESULTS AND DISCUSSION

As part of a master project the student Yuko Lam carried out most of the analysis and interpretation of the data shown in this chapter.

A.3.1. End-Group Determination

The most significant challenge in a mixture of polymeric components is the determination of the end-group mass. The analysis relies on distinguishing the end-group(s) from the repeat unit. The most common method to achieve this is by using the mass of the observed peak in the spectrum.^{16, 23} This method has been applied in polymeric compounds analyzed by FT-ICR and MALDI-TOF. The principle of the method is based on the following equation:

$$M_{\text{peak}} = n(M_{\text{mono}}) + M_{\text{end}} + M_{\text{adduct}} \quad (\text{A.1})$$

where M_{peak} is the mass of the observed peak in the spectrum, n is the number of the repeat unit, M_{mono} is the mass of monomer, M_{end} is the mass of end group, and M_{adduct} is the mass of the adduct added. In an experiment, the mass of the adduct ion, (eg. Fe^+ or Na^+), is usually known, and it will depend on the solution added to the sample during preparation. In order to simplify the equation, the mass of the adduct ion can be deducted. The simplified equation can be written as follows:

$$M_{(\text{peak-adduct})} = n(M_{\text{mono}}) + M_{\text{end}} \quad (\text{A.2})$$

where $M_{[\text{peak-adduct}]}$ is the mass of the molecule only. The mass of the monomer c and be determined from the spectrum which is the space unit between two peaks. However, M_{end} and n cannot be determined directly from the spectrum, and must be inferred. An equation is required for the connection between M_{end} and n . Therefore, the equation can be further developed:

$$M_{\text{end}} = (i-n)(M_{\text{mono}}) + M_{\text{res}} \quad (\text{A.3})$$

where i is the maximum integer value obtained by:

$$i = M_{(\text{peak-adduct})} / M_{\text{mono}} \quad (\text{A.4})$$

M_{res} is the smallest possible end-group mass. Taking the example of a mass spectrum of Gelucire® 44/14 (A), there is a peak at 1009.5772 with a mass error of 0.17 ppm (see Table A.27). The information for equat m/z ion A.4 can be extracted from the spectrum. $M_{\text{peak-adduct}}$ is calculated by subtracting M_{adduct} , the mass of the adduct is: Na = 22.989770 Da, so $M_{(\text{peak-adduct})} = 986.5875$. The monomer unit is (CH₂-CHO-) ($M_{\text{mono}} = 44.0262$), thus $i \leq 22$. Introducing these values in equation (6.3) the mass of the end group is (M_{res} is 18.0107). This corresponds to the mass of a water molecule. Therefore, the chemical formula related to the peak at m/z 1009.5772 is proposed to be [H(OCH₂CH₂)₂₂OHNa]⁺. This method can be directly applied to cases where the end-group mass is smaller than the mass of the repeat unit. In this particular case the M_{res} is directly equal to the mass of the end group.

In the cases where the end-group mass is bigger than the mass of the repeat unit, several steps are required for the end-group determination. An example for this case is the peak at m/z 1307.8277 with a mass error of 0.10 ppm, which is taken from PEG series extracted from the spectrum (See Table A.7). Using equation A.4, the value of the maximum integer value can be calculated by introducing the values corresponding to the monomer unit ($M_{\text{mono}} = 44.0262$), and the assumed adduct mass, Na⁺, $M_{[\text{peak-adduct}]} = 1284.8379$ m/z . The maximum integer value, i , is therefore estimated to be 29, leaving a residual mass (M_{res}) of 8.077718. There is no acceptable combination with mass corresponding to 8.077718. Therefore, in order to obtain a possible end-group mass, M_{mono} is added to M_{res} following equation A.3:

$$8.0777 + (1) * 44.0262 = 52.1034$$

$$8.0777 + (2) * 44.0262 = 96.1296$$

$$8.0777 + (3) * 44.0262 = 140.1558$$

$$8.0777 + (4) * 44.0262 = 184.1820$$

$$8.0777 + (5) * 44.0262 = 228.2082$$

$$8.0777 + (6) * 44.0262 = 272.2350$$

In the case of Gelucire[®] 44/14, fatty acids react with free PEG to form monoester or diester PEG. Therefore, the mass of the end-group is possibly coming from the mass of the fatty acid. Myristic acid (C14) gives the same molecular mass m/z 228.2089, which is close to an end group mass of 228.2082. The repeat unit is equal to $29 - 5 = 24$. The chemical formula of m/z 1307.8277 is therefore proposed to be $[H(OCH_2CH_2)_{24}C_{14}H_{27}O_2]$.

With the equation $M_{\text{peak}} = n(M_{\text{mono.}}) + M_{\text{end}} + M_{\text{adduct}}$, consistency of the end-group mass and the monomer mass can be shown graphically. If three free PEG series, from the mass spectrum of Gelucire[®] 44/14 (A), are taken as example, the peak masses can be plotted versus the repeat unit. A best fit line can be calculated through these points using linear least squares (see **Fig. A.3**). The three plots in **Fig. A.3** corresponds to the masses of members of three polymeric series: [PEG.Na], [PEG.Na₂] and [PEG.Na₃]. As expected, the mass difference between the linear regression lines in each case is M_{adduct} (Na = 22.9898). The gradient of each line can be used to determine the mass of the monomer unit and also to show that this unit is consistent along the length of the polymer chain. The intercept for each line corresponds to the mass of the end unit and the adduct $[M_{\text{res}} + M_{\text{adduct}}]$.

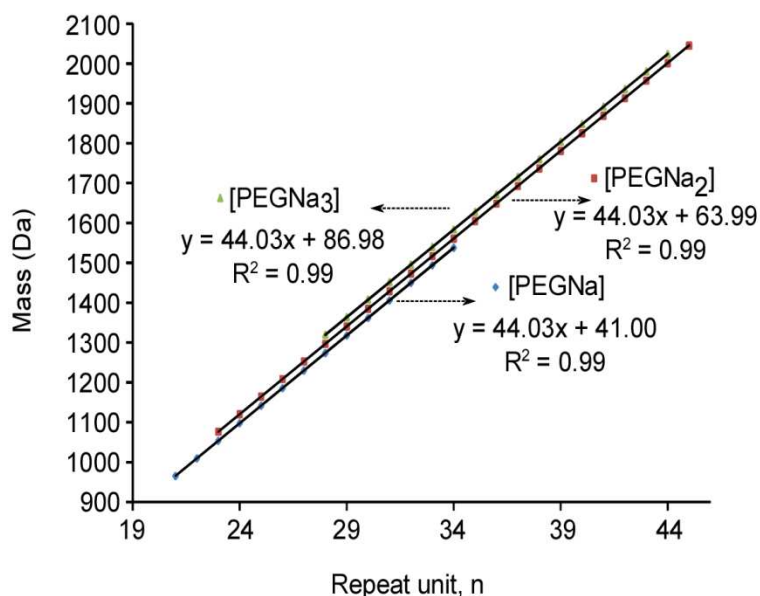


FIGURE A.3: Consistency of the end-group mass and monomer mass of free PEG present in Gelucire 44/14[®] (A).

A.3.2. Comparison between the batches of Gelucire[®] 44/14

With the end-group determination methods described above, different polymeric series and compounds were identified in the samples: A and B, corresponding to Gelucire[®] 44/14. The identification of at least 6 common polymeric series was successfully achieved. **Fig. A.4A** shows the various polymeric distribution patterns assigned to samples of (A). In the case of sample (B), a similar pattern is shown in **Fig. A.4B**, but the signals for the polymeric series are greatly reduced by the presence of some unknown single charge contaminants.

In both spectra, most of the expected compounds (**Fig. A.1**) were identified. In addition to the monoester components, various diester polymeric compounds were also identified in both samples. Interestingly, PEG stearate ($C_{18}H_{36}O_2 \cdot [OCH_2CH_2]_n$), which is proposed as a constituent by the manufacturer, Gattefosc, was not identified in either sample. Monoglycerides, diglycerides, and

triglycerides were also found in both samples between 400 – 800 m/z . The mass error for each peak assigned to these classes of molecules was less than 0.630 ppm. Around 36 glycerides peaks were found in both Gelucire[®] samples. They are usually formed by different combinations of fatty acids and glycerol. No mono caprylic acid (C8) glyceride was found in the spectra, which is normally detected in negative ion mode,²⁴ but mono stearic acid (C18) glyceride was found. Moreover, a small fraction of Polypropylene glycol [H(OCH₂CH₂CH₂)_nOH] was also found in both samples. In the spectra, there is a low intensity charge state distribution series in the 500-600 m/z region. Using internal calibration and matching the exact masses to possible elemental compositions and isotopic distributions, the chemical composition for these series was calculated to be [C₇₄H₁₄₇O₃₃Na₄]³⁺ with an average mass error of (0.119±0.092) ppm. This series corresponded to PEG monolaurate ester [C₁₂H₂₃O₂ • (OCH₂CH₂)_nNa] (see Fig. A.5). The components in the samples are similar to the structures proposed by previous literature¹ and the producer.

TABLE A.1. List of most common polymeric compounds found in the two batches of Gelucire[®] 44/14

Compound	Elemental Formulae	Gelucire [®] 44/14 Sample (A)	Gelucire [®] 44/14 Sample (B)
PEG monolaurate ester	$C_{12}H_{24}O_2 \cdot [OCH_2CH_2]_n$	X	X
PEG monoacrylate	$C_8H_{16}O_2 \cdot [OCH_2CH_2]_n$	X	X
PEG monocaprate	$C_{10}H_{20}O_2 \cdot [OCH_2CH_2]_n$	X	X
PEG monomystate	$C_{14}H_{28}O_2 \cdot [OCH_2CH_2]_n$	X	X
PEG monopalmitate	$C_{16}H_{32}O_2 \cdot [OCH_2CH_2]_n$	X	X
Free PEG	$H(OCH_2CH_2)_nOH$	X	-
	$[C_{56}H_{122}O_{10}Na]^+$	X	-
	$[C_{64}H_{122}O_{11}Na]^+$	X	-
PEG stearate	$C_{18}H_{36}O_2 \cdot [OCH_2CH_2]_n$	-	-
	$C_{20}H_{18}O_3 \cdot [OCH_2CH_2]_n$	X	X
	$C_{22}H_{42}O_3 \cdot [OCH_2CH_2]_n$	X	X
	$C_{24}H_{46}O_3 \cdot [OCH_2CH_2]_n$	X	X
	$C_{26}H_{50}O_3 \cdot [OCH_2CH_2]_n$	X	X
diesters	$C_{28}H_{54}O_3 \cdot [OCH_2CH_2]_n$	X	X
	$C_{30}H_{58}O_3 \cdot [OCH_2CH_2]_n$	X	X
	$C_{32}H_{62}O_2 \cdot [OCH_2CH_2]_n$	X	X
mono caprylic acid (C8) glyceride	$C_8H_{16}O_2$	-	-
mono stearic acid (C18) glyceride	$C_{18}H_{36}O_2$	X	X
Polypropylene glycol	$H(OCH_2CH_2CH_2)_nOH$	X	X

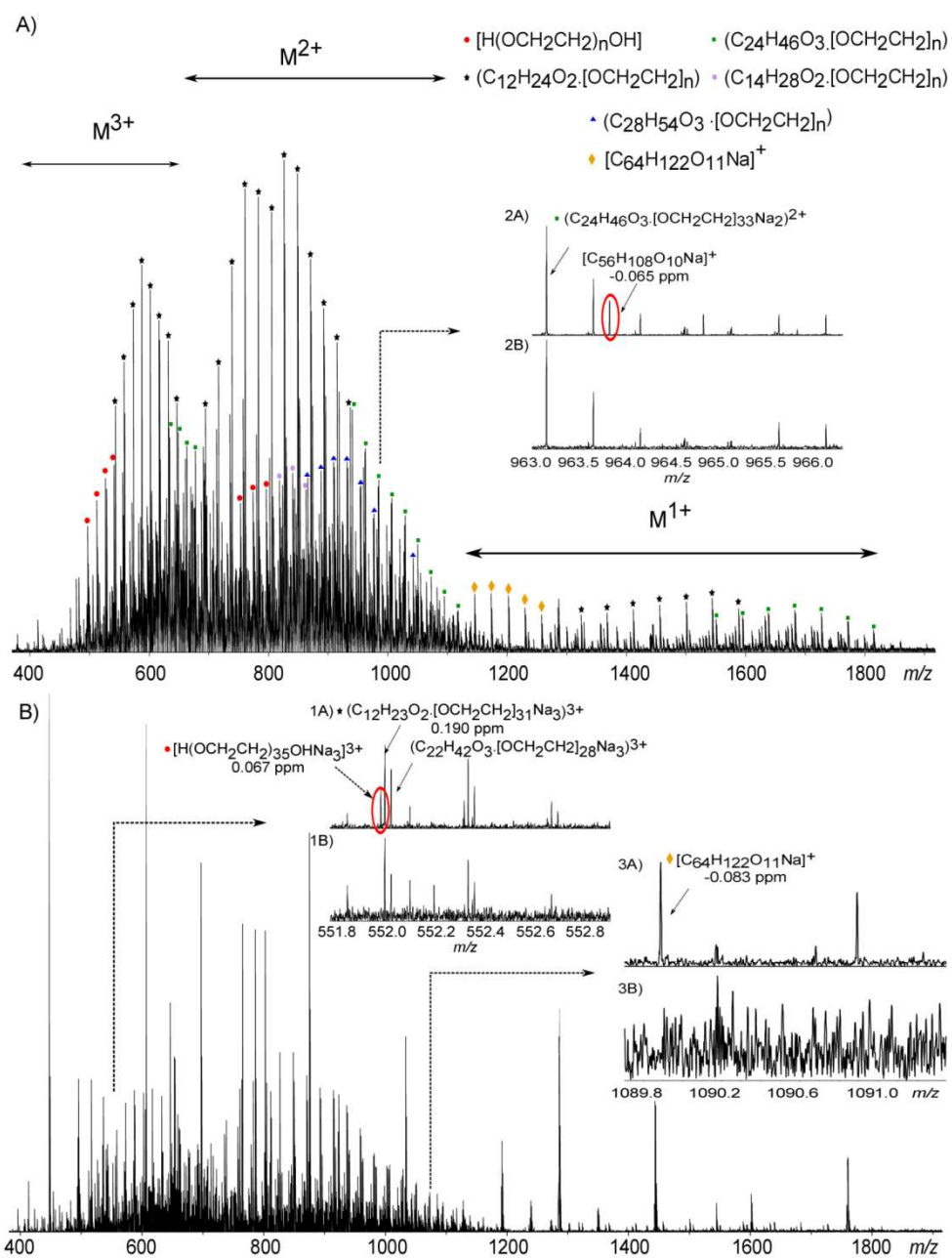


FIGURE A.4: A) Spectrum of polymeric distribution patterns in Gelucire[®] 44/14 (A). B) Spectrum of polymeric distribution patterns in Gelucire[®] 44/14 (B). Inserts 1A-1B, 2A-2B and 3A-3B show the differences between the two spectra.

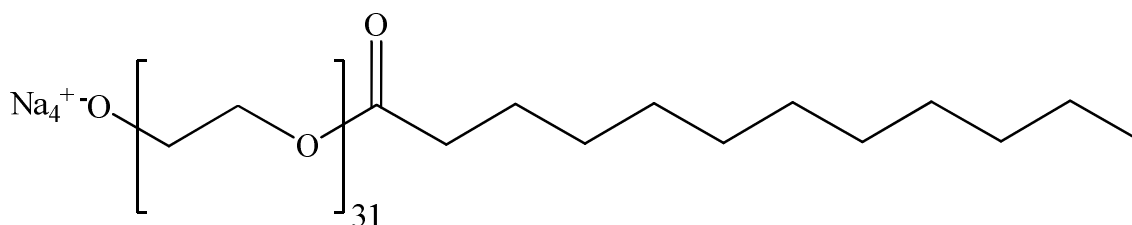


FIGURE A.5: Proposed chemical structure for $[\text{C}_{74}\text{H}_{147}\text{O}_{33}\text{Na}_4]^{3+}$ with neutralization on $-\text{OH}$ group.

The mass spectra of both batches are very similar. However, using the high resolution of FT-ICR-MS, three major polymeric differences can be identified. The

first difference is the presence of PEG [$\text{H}(\text{OCH}_2\text{CH}_2)_n\text{OH}$] (mass error of 0.067 ppm) in (A), but not in (B) (Fig. A.4, insert 1A-1B). This series is next to the series of $(\text{C}_{12}\text{H}_{23}\text{O}_2 \cdot [\text{OCH}_2\text{CH}_2]_n\text{Na}_3)^{3+}$ and $(\text{C}_{22}\text{H}_{42}\text{O}_3 \cdot [\text{OCH}_2\text{CH}_2]_n\text{Na}_3)^{3+}$ with 0.02 Da difference between the neighboring peak distributions.

The other two differences are the two polyethylene polymeric series, with a repeat unit of CH_2CH_2 , being inferred from the recorded mass difference of m/z 28.0313 (± 0.0002). These series are only present in sample (A). The polyethylene polymeric series, produces its most intense peak at m/z 963.7839 (1+). (Figure A.4, inset 2A-2B). The proposed chemical formula corresponding to this peak is $[\text{C}_{56}\text{H}_{108}\text{O}_{10}\text{Na}]^+$, with mass error of -0.065ppm. Polyethylene polymeric series 2 (Figure A.4, inset 3A-3B), produces its most intense peak at m/z 1089.8884 (1+). The proposed chemical formula in this peak is $[\text{C}_{64}\text{H}_{122}\text{O}_{11}\text{Na}]^+$, with mass error -0.083ppm. From the mass spectrum, only the information of the polymeric chemical composition is obtained leaving some structural ambiguity. Thus, it is challenging to elucidate the molecular structure of these compounds; which underlines the importance of using MS/MS as well as complementary techniques in the analysis of highly complex mixtures. Although in this particular case the use of MS/MS were relatively uninformative, showing only one fragment corresponding to $[\text{M}+\text{H}-\text{Na}]^+$. (See Fig. A.6).

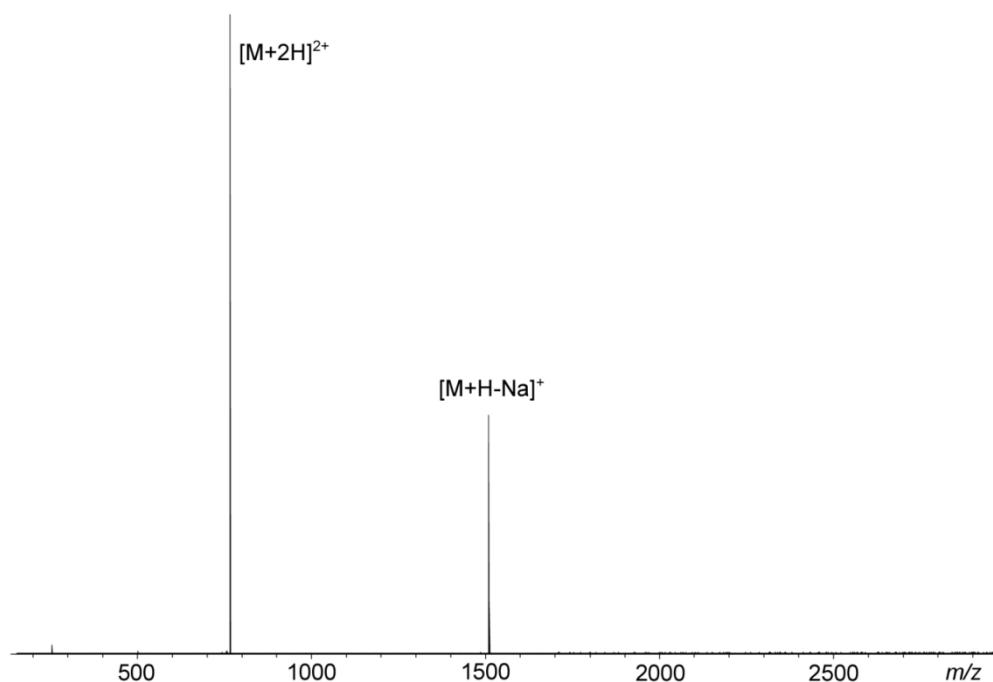
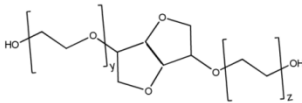
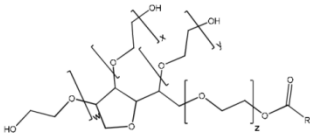


FIGURE A.6: CAD spectrum of 765.4194^{2+} from Polysorbate 80.

A.3.3. Comparison between batches of Polysorbate 80 from Acros (A) and Fisher (B)

Analysis indicates that both the Acros (A) and Fisher (B) samples are composed of sorbitan polyethoxylates, isosorbide polyethoxylates, and free PEG (Tables A.2). Based on previous publications, and the chemistry involved, these compounds are expected to be present in polysorbate 80.^{12, 25} The average mass errors of all charge states in the spectra reported herein was below 0.12 ppm. **Fig. A.7** shows a singly charged compound located between the singly charged isosorbide polyethoxylate peaks. This compound corresponds to sorbitan polyethoxylate ($n = 25$; m/z 1287.7138) with a molecular formula $[C_{52}H_{103}O_{26}Na_2]^+$ and a mass error of 0.2 ppm. The two possible chemical structures of this compound are shown in **Fig. A.8**. Polypropylene glycol $[H(OCH_2CH_2CH_2)_nOH]$ was also found in both samples. PEG monolinoleate ester was only found in the Fisher sample. All of the compounds indicated above are products of the reaction to produce Polysorbate 80.

TABLE A.2: List of most common polymeric compounds found in the two batches of Polysorbate 80

Compound	Elemental Formulae	Polysorbate 80 Sample (A)	Polysorbate 80 Sample (B)
Sorbitan polyethoxylates	$[C_{52}H_{103}O_{26}Na_2]^+$	X	X
Isosorbide polyethoxylates		X	X
Free PEG	$H(OCH_2CH_2)_nOH$	X	X
Polypropylene glycol	$H(OCH_2CH_2CH_2)_nOH$	X	X
Polysorbate monooleate ester		-	X

A closer look of the spectrum shows four main differences in the polymeric series between the two samples (See inserts in **Fig. A.7**). The specific mass errors for every peak are reported in section A.6 of this chapter. When comparing all the series across the spectrum, there is only 0.1 Da difference between them. In the particular case of the peak at m/z 917.5813, there are two possible chemical structures that have the same molecular weight with different repeat units. For example, isosorbide polyethoxylate monolinoleate ester repeat unit is 11; while PEG monolinolenate ester repeat unit is 14, but they both have the same mass. Compared with Acros sample (A), Fisher sample (B) components are a closer match with the expected products in Polysorbate 80 with some fatty acid reacted with sorbitan, isosorbide polyethoxylate, and free PEG. However, no stearic acid was detected.

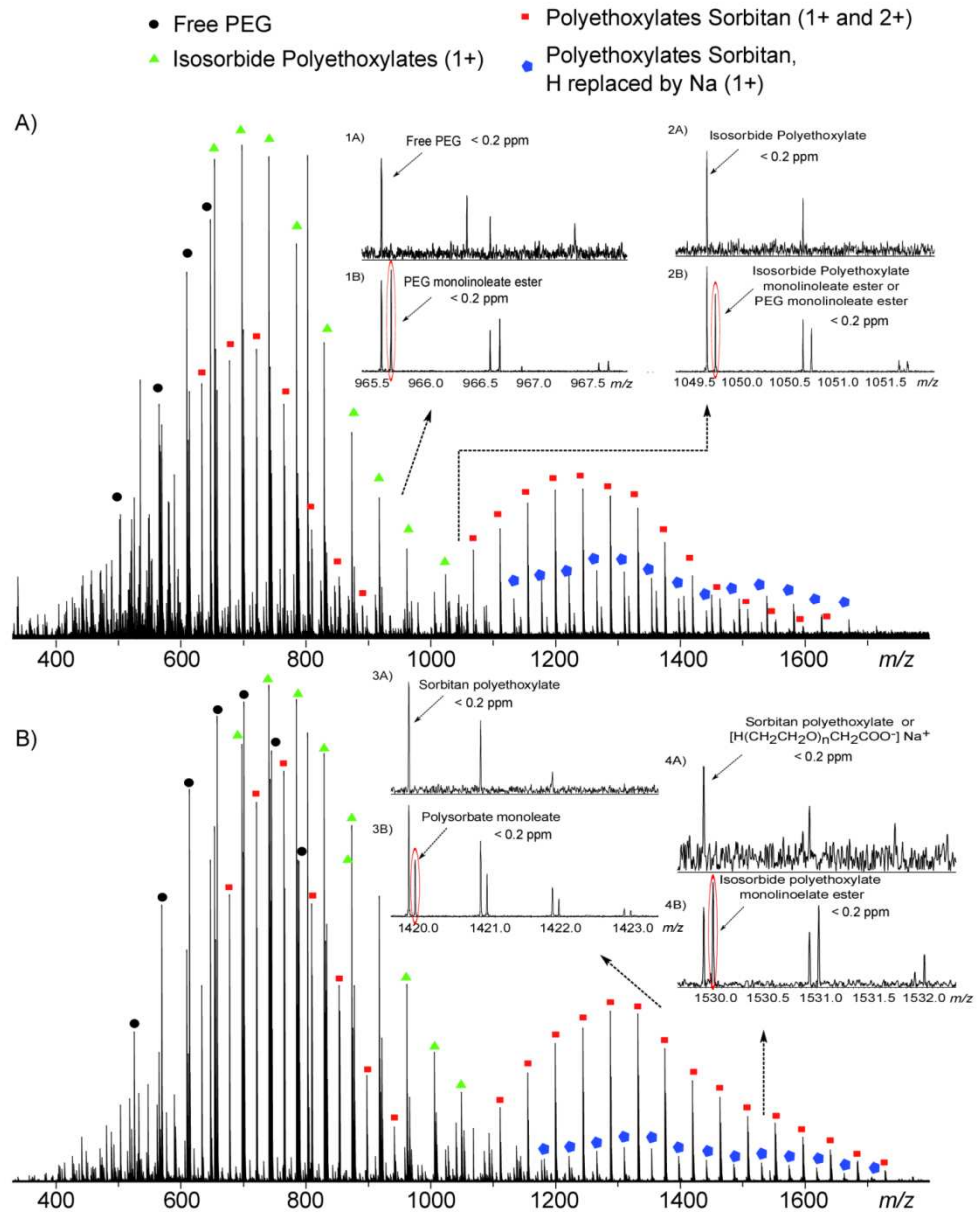


FIGURE A.7: A) Spectrum of polymeric distribution patterns in Acros sample (A). B) Spectrum of polymeric distribution patterns in Fisher sample (B). Inserts 1A-1B, 2A-2B, 3A-3B, and 4A-4B show the differences between the two spectra.

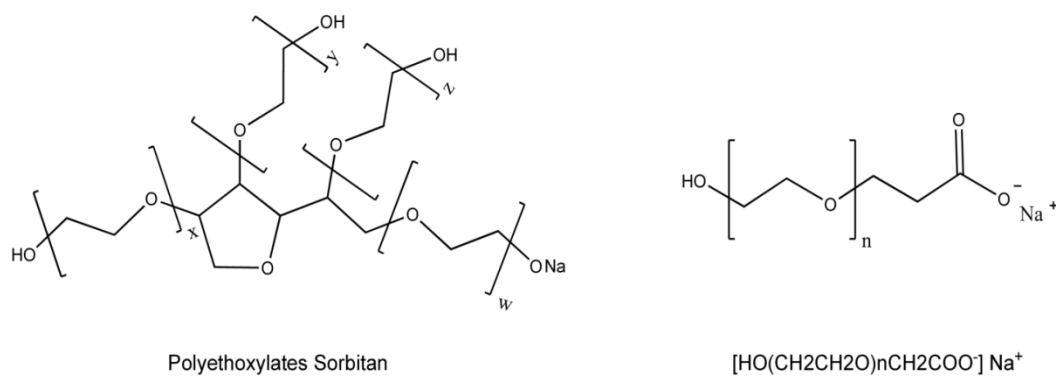


FIGURE A.8: Proposed structure for sorbitan polyethoxylate $[\text{C}_{52}\text{H}_{103}\text{O}_{26}\text{Na}_2]^+$; $[\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{CH}_2\text{COOH}]$ with a Na replacement.

A.4. CONCLUSION

The use of high resolution mass spectrometry revealed the presence of polymeric differences between two batch samples of Gelucire[®] 44/14 (A & B) and two batch samples of Polysorbate 80. The additional presence of free PEG $[\text{HO}(\text{OCH}_2\text{CH}_2)_n\text{OH}]$, polyethylene polymeric series $[\text{C}_{56}\text{H}_{108}\text{O}_{10}\text{Na}]^+$ and $[\text{C}_{64}\text{H}_{122}\text{O}_{11}\text{Na}]^+$ was determined in Gelucire[®] 44/14 (sample A). The analysis of these commercial Polysorbate formulations exposes the presence of PEG monolinoelate ester, isosorbide polyethoxylate or PEG monolinoleate ester, and Polysorbate monoleate. The results agreed with previous work on the Polysorbate species distribution of Polysorbate 80 using different techniques.^{8, 12} However, the use of FT-ICR-MS provided enough resolution to overcome the complexities in the spectra.

A.5. REFERENCES

1. Chambin, O. and Jannin, V. Interest of Multifunctional Lipid Excipients: Case of Gelucire (R) 44/14. *Drug Development and Industrial Pharmacy*. **2005**, *31*, 527-534.
2. Fernandez, S., Rodier, J.-D., Ritter, N., Mahler, B., Demarne, F., Carriere, F., and Jannin, V. Lipolysis of the Semi-Solid Self-Emulsifying Excipient Gelucire (R) 44/14 by Digestive Lipases. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids*. **2008**, *1781*, 367-375.
3. Sarpietro, M.G., Pitarresi, G., Ottimo, S., Giuffrida, M.C., Ognibene, M.C., Fiorica, C., Giammona, G., and Castellit, F. Interaction between Drug Loaded Polyaspartamide-Polylactide-Polysorbate Based Micelles and Cell Membrane Models: A Calorimetric Study. *Molecular Pharmaceutics*. **2011**, *8*, 642-650.
4. Lee, Y.H., Jeong, E.S., Cho, H.E., and Moon, D.-C. Separation and Determination of Polyethylene Glycol Fatty Acid Esters in Cosmetics by a Reversed-Phase Hplc/Elsd. *Talanta*. **2008**, *74*, 1615-1620.
5. Pauli, G.H. Chemistry of Food-Additives - Direct and Indirect Effects. *Journal of Chemical Education*. **1984**, *61*, 332-334.
6. Svensson, A., Neves, C., and Cabane, B. Hydration of an Amphiphilic Excipient, Gelucire (R) 44/14. *International Journal of Pharmaceutics*. **2004**, *281*, 107-118.
7. Singh, S.K., Verma, P.R.P., and Razdan, B. Atomic Force Microscopy, Transmission Electron Microscopy, and Photon Correlation Spectroscopy: Three Techniques for Rapid Characterization of Optimized Self-Nanoemulsifying Drug Delivery System of Glibenclamide, Carvedilol, and Lovastatin. *Journal of Dispersion Science and Technology*. **2011**, *32*, 538-545.
8. Abrar, S. and Trathnigg, B. Separation of Polysorbates by Liquid Chromatography on a Hilic Column and Identification of Peaks by Maldi-Tof-Ms. *Analytical and Bioanalytical Chemistry*. **2011**, *400*, 2119-2130.
9. Hvattum, E., Yip, W.L., Grace, D., and Dyrstad, K. Characterization of Polysorbate 80 with Liquid Chromatography Mass Spectrometry and Nuclear Magnetic Resonance Spectroscopy: Specific Determination of Oxidation Products of Thermally Oxidized Polysorbate 80. *Journal of Pharmaceutical and Biomedical Analysis*. **2012**, *62*, 7-16.
10. Kato, H., Nagai, Y., Yamamoto, K., and Sakabe, Y. Determination of Polysorbates in Foods by Colorimetry with Confirmation by Infrared Spectrophotometry, Thin-Layer Chromatography, and Gas-Chromatography. *Journal of the Association of Official Analytical Chemists*. **1989**, *72*, 27-29.
11. Mckean, D.L., Pesce, A.J., and Koo, W. Analysis of Polysorbate and Its Polyoxyethylated Metabolite. *Analytical Biochemistry*. **1987**, *161*, 348-351.
12. Ayorinde, F.O., Gelain, S.V., Johnson, J.H., and Wan, L.W. Analysis of Some Commercial Polysorbate Formulations Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Rapid Communications in Mass Spectrometry*. **2000**, *14*, 2116-2124.
13. Hart-Smith, G. and Barner-Kowollik, C. Contemporary Mass Spectrometry and the Analysis of Synthetic Polymers: Trends, Techniques and Untapped Potential. *Macromolecular Chemistry and Physics*. **2010**, *211*, 1507-1529.

14. Gruending, T., Weidner, S., Falkenhagen, J., and Barner-Kowollik, C. Mass Spectrometry in Polymer Chemistry: A State-of-the-Art up-Date. *Polymer Chemistry*. **2010**, *1*, 599-617.
15. O'connor, P.B. and Mclafferty, F.W. Oligomer Characterization of 4-23 Kda Polymers by Electrospray Fourier Transform Mass Spectrometry. *Journal of the American Chemical Society*. **1995**, *117*, 12826-12831.
16. Koster, S., Duursma, M.C., Boon, J.J., and Heeren, R.M.A. Endgroup Determination of Synthetic Polymers by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **2000**, *11*, 536-543.
17. Vanrooij, G.J., Duursma, M.C., Heeren, R.M.A., Boon, J.J., and Dekoster, C.G. High Resolution End Group Determination of Low Molecular Weight Polymers by Matrix-Assisted Laser Desorption Ionization on an External Ion Source Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Journal of the American Society for Mass Spectrometry*. **1996**, *7*, 449-457.
18. Mcewen, C.N., Simonsick, W.J., Larsen, B.S., Ute, K., and Hatada, K. The Fundamentals of Applying Electrospray-Ionization Mass-Spectrometry to Low-Mass Poly(Methyl Methacrylate) Polymers. *Journal of the American Society for Mass Spectrometry*. **1995**, *6*, 906-911.
19. Soeriyadi, A.H., Bennet, F., Whittaker, M.R., Barker, P.J., Barner-Kowollik, C., and Davis, T.P. Degradation of Poly(Butyl Methacrylate) Model Compounds Studied Via High-Resolution Electrospray Ionization Mass Spectrometry. *Journal of Polymer Science Part a-Polymer Chemistry*. **2011**, *49*, 848-861.
20. Bennet, F., Hart-Smith, G., Gruending, T., Davis, T.P., Barker, P.J., and Barner-Kowollik, C. Degradation of Poly(Methyl Methacrylate) Model Compounds under Extreme Environmental Conditions. *Macromolecular Chemistry and Physics*. **2010**, *211*, 1083-1097.
21. Benneto, F., Barker, P.J., Davis, T.P., Soeriyadi, A.H., and Barner-Kowollik, C. Degradation of Poly(Butyl Acrylate) and Poly(2-Hydroxyethyl Methacrylate) Model Compounds under Extreme Environmental Conditions. *Macromolecular Chemistry and Physics*. **2010**, *211*, 2034-2052.
22. Koster, S., Duursma, M.C., Boon, J.J., Nielen, M.W.F., De Koster, C.G., and Heeren, R.M.A. Structural Analysis of Synthetic Homo- and Copolyesters by Electrospray Ionization on a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *Journal of Mass Spectrometry*. **2000**, *35*, 739-748.
23. Li, Y., Hoskins, J.N., Sreerama, S.G., Grayson, M.A., and Grayson, S.M. The Identification of Synthetic Homopolymer End Groups and Verification of Their Transformations Using Maldi-Tof Mass Spectrometry. *Journal of Mass Spectrometry*. **2010**, *45*, 587-611.
24. Peris-Vicente, J., Garrido-Medina, R., Simo-Alfonso, E., Gimeno-Adelantado, J.V., and Domenech-Carbo, M.T. Infusion Mass Spectrometry as a Fingerprint to Characterize Varnishes in Oil Pictorial Artworks. *Rapid Communications in Mass Spectrometry*. **2007**, *21*, 851-856.
25. Frison-Norrie, S. and Sporns, P. Investigating the Molecular Heterogeneity of Polysorbate Emulsifiers by Maldi-Tof Ms. *Journal of Agricultural and Food Chemistry*. **2001**, *49*, 3335-3340.