Investigating the relationship between glutamine deamidation and collagen breakdown in bone

Joanna Paula Simpson

PhD

University of York

Chemistry

September 2015

Abstract

Bone can survive in the burial environment for millions of years and can provide direct information about individuals, during their lifetime and post mortem. However it is not currently possible to *directly* date bone material beyond the limits of radiocarbon dating ~50,000 years before present(BP). It was suggested by van Doorn *et al.*, (2012), that deamidation in bone collagen may be related to thermal age. However, a number of bones investigated from the same site resulted in a large range of glutamine deamidation measurements. This thesis therefore explores in detail the potential of glutamine deadmidation as a dating tool for bones older than 50,000 yeas BP.

The work presented in this thesis investigates a number of potential causes for the observed variability. It was found that the pre-treatment of bone with HCl not only increased levels of deamidation, but also caused significant levels of swelling and breakdown of the collagen fibril structure. Areas of bone exhibiting localised macroscopic degradation were shown to be correlated with an increase in the levels of glutamine deamidation. This indicates that the measurement of deamidation may be used as a method by which to investigate the preservation state of bone. The analysis of synthetic peptide mixtures, containing glutamine or glutamic acid showed that an ionisation bias occurs when using MALDI, with product peptide E consistently underestimated. However, analysis of the same peptide mixtures using ESI showed no ionisation bias between the two peptide products and is therefore preferable for the measurement of glutamine containing peptides.

Kinetic heating experiments performed at high temperatures (80 °C -140 °C), have proven unsuitable for the measurement of deamidation within the buffer soluble collagen fraction. Deamidation in the buffer soluble collagen fraction was successfully measured at low temperature (65 °C), and resulted in an increase in deamidation over time. However levels of deamidation in the mineral bound fraction extracted from the same bone chips heated at 65 °C did not appear to undergo deamidation over time; indicating that glutamine deamidation in this fraction is slower and therefore may be suitable for the analysis of preservation of older bone from the Late Pleistocene Period. Levels of deamidation in bones analysed from nine sites ranging in age from ~ 8000 - 30,000 years BP showed a correlation with levels of glutamine deamidation and thermal age. However, using the current methodology, it is not possible to identify the age of the sites using only the measured levels of deamidation. An adapted extraction method has been suggested to further investigate the relationship between deamidation and the thermal age of bone. The work carried out in this thesis has directly contributed to the broader knowledge of how collagen degrades in bone over time, and thus has wide reaching implications within the archaeological scientific community.

Table of contents

List of figures

List of tables

Acknowledgements

I would first like to thank my many supervisors for all their help and guidance over the past five years: Jane Thomas-Oates, Kirsty Penkman, Matthew Collins, Julie Wilson and Beatrice Demarchi.

Thank you Jane for all your help. You always make time for your students, despite your insanely busy schedule. I will remember our talks of cats and peptides with a smile. Kirsty, you ended up with me in your group despite the lack of AAR analysis in my thesis and have been my main supervisor. Thank you for making me so welcome in your group and for all your support over the past four years. I really appreciate everything you have done to help me on my PhD journey and I wouldn't be the researcher I am today without your help. Matthew, I'm not sure where to start. Many thanks to you for all your helpful ideas and our long conversations, your energy and enthusiasm; your stories of mummified chickens and stolen frozen dead Victorians certainly make a Chemistry PhD more interesting. Julie, thank you for your help with all things statistical and programming, you calm me down in my moments of panic and always make time to read/listen to my ideas.

I would like to thank the following people for providing samples: Mathias Stiller (Planck Institute for Evolutionary Anthropology), Duane Froese (University of Alberta) and Grant Zazula (Yukon Government), Colin Smith (La Trobe University), Tom Lord, Terry O'Connor, Andrew Chamberlin, Hervé bocherens, John Meadows and especially Ricardo Fernandes who helped me with the sampling of material in Kiel.

I would like to thank NERC for providing the funding for this PhD research, Adam Dowle, Ed Bergström, Meg Stark and Martin Fascione for their continuing technical support and advice. Thank you Sheila Taylor (the fairy godmother of our lab) for all your advice, support, encouragement and being an excellent coffee break companion. Thank you to The KP research group, especially Kirsty High for reading my thesis. I would also like to thank all my friends at city screen and all my fellow skaters at Auld Reekie Roller girls.

I would like to give special thanks to; Dave, my partner who has been so supportive despite my decisions usually involving me moving away from him. My two best friends, Ian and Phil- Ian a big thank you to you and the bairns for always being there for me, you have the ability to make any situation better, even those involving motorbikes, rain and purple hands. Phil you are just such a lovely human and always make me smile and I am so glad to have you as my friend. Last but not least I would like to thank my family, especially my mam, dad and grandma. I would like to thank Stimpy for listening to all my presentations and making wherever I moved to feel like home.

Declaration

I hereby certify that the work described in this thesis is my own, except where otherwise acknowledged, and has not been submitted previously for a degree at this, or any other, university. All sources are acknowledged as References.

Dedicated to the late Grandad Simpson, and my furry black companion Stimpy

Chapter 1 - Introduction

Thesis road map - *a brief introduction into the structure a layout of the thesis*

Chapter one – This thesis uses terminology from multiple scientific disciplines including: Chemistry, Biology, Archaeology and quaternary science; this chapter aims to introduce the relevant background in these areas as well as previous work within the field.

Chapter two – A large proportion of this study has included method development and therefore this chapter describes the key methodologies used.

Chapter three – Investigates the suitability of the method parameters employed by van Doorn *et al.,* (2012) and their effects on the measurement of glutamine deamidation in protein extracted from a Roman bone.

Chapter four – It was identified that in order to test how accurately we can measure levels of deamidation a reference material is required. In this chapter, two peptides were synthesised, one containing glutamine and the other containing glutamic acid in order to validate the current methodology, as well as investigate possible ionisation bias between two different ionisation methods utilised in Mass Spectrometry.

Chapter five – In order to compare levels of deamidation in bones from sites with different thermal histories, the relationship between temperature and glutamine deamidation in bone collagen is explored.

Chapter six – Looks at possible natural variability of deamidation in bone, by measuring sub-samples across the length of a Roman bovid metatarsal, as well as the effects of HCl and EDTA pre-treatment on the structure of collagen fibrils.

Chapter seven – This chapter compares levels of deamidation in protein extracted from bones of nine sites, ranging in age from ~ 800 years BP to ~30,000 years BP, in order to investigate the relationship between glutamine deamidation and thermal age.

Chapter eight- The final chapter is a summary of the conclusions from the work carried out, and suggested future work

1.1 Background information on proteins and amino acids

Amino acid residues join together to form chains of peptides and proteins though a condensation mechanism with removal of a water molecule. The in vivo process of amino acid and peptide synthesis involves the translation of mRNA using tRNA and ribosomes. These peptide chains can vary in length and composition. The structure of a protein is directly influenced by the number and type of residues present; for example in collagen every third residue is Gly (glycine), and the repetition of this small residue allows the chain to fold into its helical shape (Shoulders *et al.*, 2009)

The sequence of amino acids which make up a protein are conventionally labelled using a one letter code notation from left to right (N to C terminus). The order of residues in the protein chain is known as the primary structure, further folding of segments within the chain is described as the secondary structure. The overall folding of the chain into a compact three-dimensional structure is called the tertiary structure, sometimes referred to as the native structure because it is this structure that makes the protein biologically active. Some proteins form a quaternary structure, whereby there is an arrangement of different chains within the structure; an example of this would be haemoglobin (Amit and Ben-Tal, 2011). The structure of a protein affects its function within the body.

All amino acids except Gly have four different groups attached to the α carbon, and are therefore chiral and optically active. This results in each compound being able to exist in different forms (i.e. mirror images of one another) depending on how the groups are arranged. These are known as enantiomers (Figure 1).

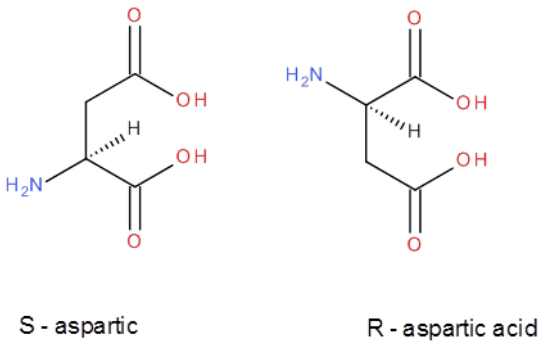


Figure 1: Two enantiomers of aspartic acid. The two forms R and S are mirror images of one another and are therefore non-superimposable

There are two common ways in which enantiomers are named. Generally in chemistry the Cahn-Inglod-Prelog nomenclature (CIP system) is used. This naming system denotes the isomer as R or S, this notation comes from the Latin words rectus (right) and sinister (left).

A second naming system is more commonly used in biology and biochemistry, as it was the way in which the stereochemistry of biological molecules were originally assigned. This system uses the notation L (laevo) and D (dextro) to name enantiomers. Amino acids are termed D and L on the basis of their similarity to the structure of glyceraldehyde. The discovery and naming of enantiomers of glyceraldehyde was determined by the way in which the molecule rotated plane polarised light.

Amino acids are in the L-configuration in most living tissues apart from some forms of bacteria (Bhattacharyya and Banerjee, 1974; Johnson and Miller, 2007). After death, L-amino acids undergo conversion to the D-form until a dynamic equilibrium is reached. This type of equilibrium is described as dynamic as the forward (K1) and backward rate (K2) is constant, therefore k1/k2 = 1. The term used to describe the conversion of one enantiomer to another is racemization. Factors which are known to increase the rate in which amino acids undergo racemization include extremes in pH, and temperature (Bada, 1972).

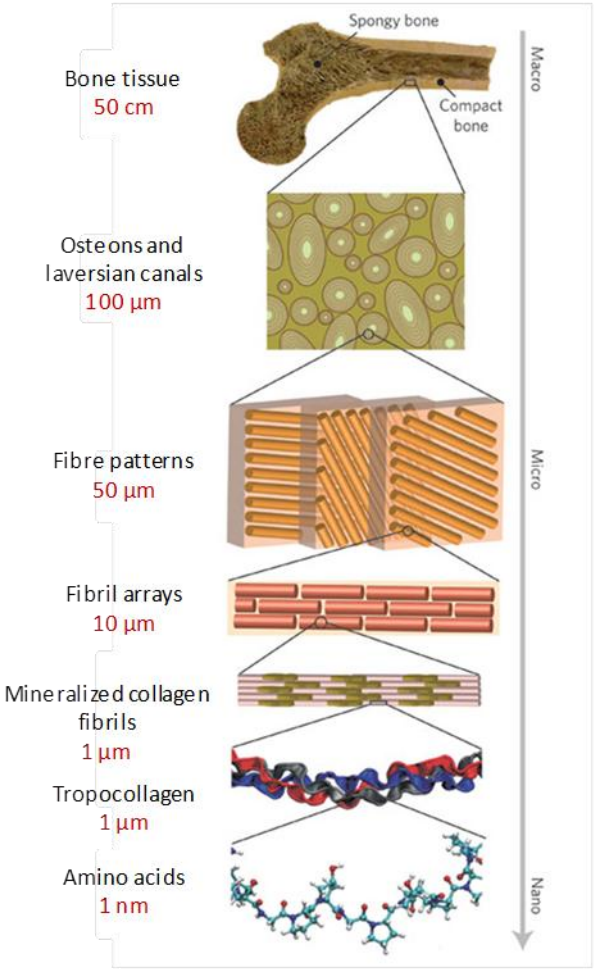
Aging of proteins in living tissue can result in the formation of D-enantiomers of amino acids through racemization. This degradation of residues has been linked with various medical conditions such as Alzheimer's disease and degradation of proteins within the eye lens (Radkiewicz *et al.*, 1996; Flaugh *et al.*, 2006). Throughout this thesis amino acid will be referred to by their full names, one and three letter codes (Table 1)

Table 2: amino acids and their three and one letter codes

Full name	three letter code	one letter code
glycine	Gly	G
alanine	Ala	A
serine	Ser	S
proline	Pro	P
valine	Val	V
threonine	Thr	T
cysteine	Cys	C
Leucine	Leu	L
isoleucine	Ile	I
asparagine	Asn	N
aspartic acid	Asp	D
glutamine	Glu	Q
lysine	Lys	K
glutamic acid	Glu	E
methionine	Met	M
histidine	His	H
phenylalanine	Phe	F
arginine	Arg	R
tyrosine	Tyr	Y
tryptophan	Trp	W

1.2 Structure of bone

Due to the complex composition of bone, chemical analysis can be a difficult. It is this complex structure that gives bone its strength and therefore ensures its preservation within the burial environment, and its prevalence in archaeological scientific research (Collins *et al.*, 1995; Collins *et al.*, 2002; Trueman and Martill, 2002; Turner-Walker, 2008; Dobberstein *et al.,* 2009) In order to understand the mechanical properties of this material it is important to consider its component phases and the structural relationships between them. Bone can be described as having five hierarchical structural levels (Rho *et al.*, 1998)(Figure 2).



1. **Macro Structure**, cancellous/Trabecular and cortical bone
2. **Micro Structure**, 10-500 µm (haversian systems, osteons, singular trabeculae)
3. **Sub-Micro structure**, 1 – 10 µm (lamellae)
4. **Nano structure**, ~ two hundred nm to 1 µm (fibular collagen and the embedded mineral)
5. **Sub-Nano structure**, below ~ two hundred nanometres, molecular structure of mineral, collagen and non-collagenous proteins

(Rho *et al.* 1998)

Figure 2: Structure of bone, detailing the five hierarchical levels from macro to sub macro structures. (Reproduced with permission from Ritchie (2011))

Degradation of bone can be described at various levels within the hierarchical structure. The level of preservation of bone can be assessed by studying the overall appearance as well as conducting histological analysis of the micro structure; however, this does not always relate directly to the preservation state of the collagen or the DNA (Nicholson, 1998). There are many factors other than age which can contribute to the level of preservation of organic material in bone; these include conditions of the burial environment such as: temperature, pH and fluctuation of the water table (Gordon and Buikstra, 1981). In order to understand the diagenetic process of collagen, degradation of the sub-nanostructures of bone needs to be investigated. Living bone material is mainly composed of; water, organic, and inorganic fractions. The inorganic fraction contains mineral salts called hydroxyapatite (HAP). The organic fraction includes proteins, lipids and water (Rho *et al.*, 1998).

Although this project will focus on the degradation of the protein, it is acknowledged that the inorganic mineral phase plays a vital role in the preservation of the organic fraction (Götherström *et al.,* 2002). It is thought that the way in which the mineral and collagen fraction are formed affects the strength of the bone material (Viguet-Carrin *et al.,* 2006). Both the tensile strength and

flexibility of this multiple strand structure has gained attention not only from research areas such as biology and chemistry but also from the field of mechanics and the built environment (Ottani *et al.,* 2001). The importance of bone mineral in the role of preservation of the organic fraction is highlighted by the correlation between hydroxyapatite preservation and successful DNA recovery from archaeological bone (Götherström *et al.,* 2002). Kasem *et al.* (2011) describe the HAP crystals as being embedded within the collagen matrix and plate-like in their morphology. This mineralisation process is thought to start in the smaller gaps between collagen fibrils, and is termed primary mineralisation. Secondary mineralisation follows with the filling of inter-fibrillar spaces. The presence of the hydroxyapatite (mineral) crystals, which embed and protect the protein, contribute to the stability and preservation of bone over geological timescales (Turner-Walker 2008; Covington *et al.,* 2010).The presence of the bone mineral has beento be correlated to collagen preservation, with an increased crystallinity in the hydroxyapatite found to correlate positively with DNA preservation (Götherström *et al.*, 2002). The presence of mineral is thought to have a powerful stabilizing effect, protecting bone against both microbial and chemical degradation (Turner-Walker, 2008).

Collagen makes up approximately 90% of the protein matrix (Marieb and Hoehn, 2007; Turner-Walker, 2008). This fibrous protein is made up of three polypeptide chains of similar size. The way in which these polypeptides can be arranged creates multiple variations of the protein. Type I collagen is the main protein found in bone material and is heterotrimeric, consisting of two identical chains identified as α 1, and a third identified as α-2 (Viguet-Carrin *et al.,* 2006; Whitford 2008; Shoulders *et al.,* 2009). The peptide sequence of collagen is repetitive, with every third residue being Gly (Rich and Crick 1961). Other residues in the sequence are described in their relation to Gly. For example residues will be referred to as being in either the X or Y position (Gly-Xaa-Yaa). It is common for x and y positions within collagen to contain high levels of proline (Pro) and lysine (Lys) (Berisio and Vitagliano, 2002). This repeating sequence helps to establish the topology required to form the triple helix characteristic of collagen, and is seen in many fibrous proteins found in living tissues (Orgel *et al.,* 2001). Gly, is an important part of the chain. This amino acid is the smallest in size and makes up around a third of the protein, enabling assemblage of adjacent chains, as well as being essential for the formation of the twist which is needed to create the helical shape (Viguet-Carrin *et al.,* 2006). Mutations where Gly is replaced with another amino acid are thought to be the most pathologically damaging (Shoulders *et al.,* 2009). Unlike most helical proteins, collagen can undergo post-translational modifications. Although it is thought that these may increase the strength of the protein (Burjanadze, 2000), little research has been carried out on the effects of these modifications on the whole collagen micro-fibril structure. These modifications can occur through both enzymatic and non-enzymatic degradation. During enzymatic degradation, an enzyme called lysyl oxidase forms cross-linked bonds between the polypeptide chains (Gorres and Raines, 2010).

1.2.1 Collagen microfibril structure

The shape and morphology of each bone will vary depending on its role within a given biological system; however the basic chemical structure remains the same, with type I collagen being ubiquitous throughout the animal kingdom (Orgel *et al.,* 2001). Procollagen is the immature form of the protein; at this stage both the N and C terminal propeptides are still attached at either end of the triple helix. These propeptides have different roles in collagen fibril formation. The main roles of the C-propeptide is to ensure that the procollagen remains soluble during its trafficking through the cell, whereas the N-propeptide ultimately influences the shape and diameter of the fibril (Viguet-Carrin *et al.* 2006). In order for the procollagen helices to form the microfibril, the C and N propeptides are removed enzymatically by procollagen C and N proteinases. Five collagen helices can then begin to fold, align and twist into a microfibril. The way in which these collagen helices line up is crucial to the formation of the microfibril.

Across the microfibril there are areas where the helices overlap with one another, as well as gap regions. Orgel *et al.* (2001) calculated the distances of these gaps and overlap regions within microfibrils using a range of analytical techniques such as XRD (X-Ray diffraction) and MIR (multiple isomorophous replacement) mapping. They describe the positioning of collagen molecules in relation to one another using both nanometers (nm) and the D period. The D period measures the stagger of molecules in the axial direction of the fibril (Figure 3).

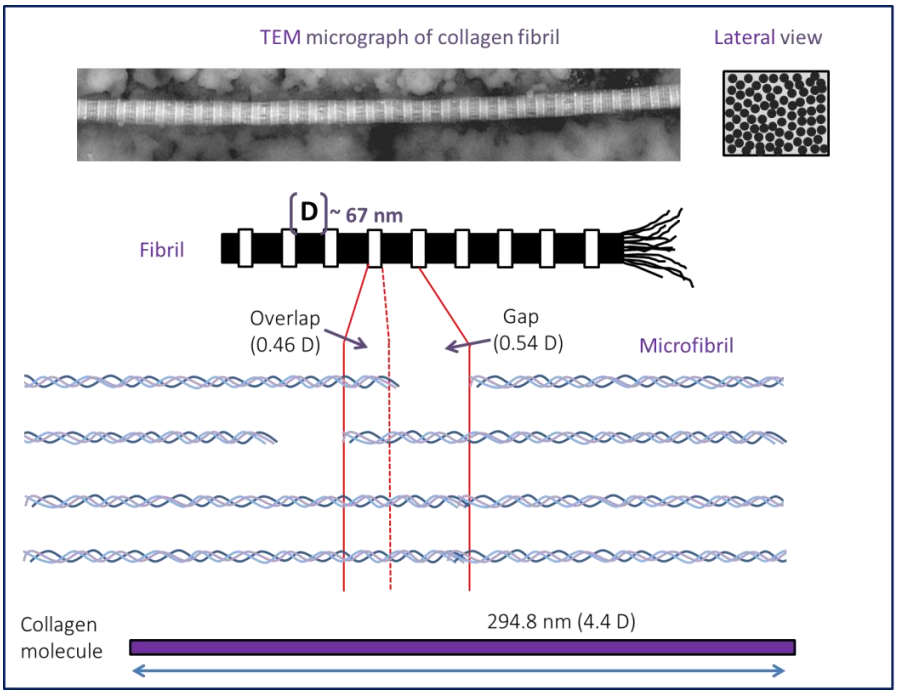


Figure 3: The D region spacing within a microfibril figure adapted from (Orgel *et al.* 2001)

Using these measurements it is possible to asses where collagen helices are in relation to one another in the microfibril, as 67 nm is equal to 234 amino acids (Orgel *et al.*, 2001). The overlap region is 0.46 D and therefore equal to 108 amino acids. The gap region is 0.54 D, equal to 126 amino acids in length.

1.2.2 The effects of hydrogen bonds and hydroxylation of proline on helix stability

One example of a post-translational modification which frequently occurs in collagen is the hydroxylation of proline. The degree or probability of proline hydroxylation is denoted by the one letter code of the amino acid in the sequence; P denotes a proline which has not been observed in collagen as hydroxyproline (Hyp) and O is always Hyp (Viguet-Carrin *et al.,* 2006). Interactions within the triple helix are enhanced by the hydrogen bonding between the amide groups and the hydroxyl group of hydroxyproline (Hyp). An indication of the importance of these hydrogen bonds has been demonstrated through heating/melting point experiments (Holmgren *et al.,* 1998).These experiments look at the transition mid-point temperature (Tm), which is the temperature at which half of the collagen has melted. A series of heating experiments with synthetic peptides has shown that the hydroxylation of proline in the Y position strengthens the helix, with hydrogen bonding to neighbouring chains resulting in an increase of Tm from 24 °C to 60 °C (Holmgren *et al.,* 1998). It should be noted that these experiments were carried out on short synthetic collagen peptides which are not subjected to intermolecular forces of the helix or surrounding chains, and do not therefore necessarily reflect the effects of increased hydrogen bonding within a biological fibril.

1.2.3 Non-collagenous bone proteins

Although collagen is the main protein found in bone (around 90% of the organic fraction), bone does contain a number of non-collagenous biomolecules, with osteocalcin being the second most abundant bone protein (Collins *et al.,* 2002). Other biomolcules found in bone include, blood proteins, cellular lips and DNA (Colins *et al.,* 2002; Kasem *et al.,* 2011). Although Osteocalcin has been found to persist in archaeological bone, work in this thesis is focused on deamidation occurring in collagen. It is likely that the methods described in this thesis also extract soluble non-collagenous proteins, however a specific set of peptides have been identified as collagenous using de-*neovo* sequencing and it is these peptides which are measured throughout the work presented in this thesis. It is also acknowledged that spectra main contain preparation induced artefacts or contaminants such as keratin, trypsin and CHCA matrix peaks and this is taken into consideration when interpreting spectral quality.

1.3 Deamidation

Deamidation is one of many pathways of degradation (Leo *et al.,* 2011) of asparagine and glutamine, both of which have been observed in collagen (Buckley *et al.*, 2009; Leo *et al.,* 2011). There are two alternative pathways of deamidation, either via internal cyclisation or via direct side-chain hydrolysis (Robinson and Robinson, 2004).

1.3.1 Mechanisms of glutamine deamidation

Deamidation of glutamine induces a negative charge followed by removal of ammonia (NH_3) and addition of water (H_2O) resulting in a mass increase of +0.984 Da. This mass change is readily measurable by mass spectrometry (Leo *et al.,* 2011; Li *et al.,* 2010). Glutamine deamidation has been found to occur at a slower rate than asparagine in both synthetic polypeptides and in longer constricted peptide chains (Geiger and Clarke, 1987). Glutamine can undergo deamidation via a 6 membered glutarimide intermediate (Figure 4).

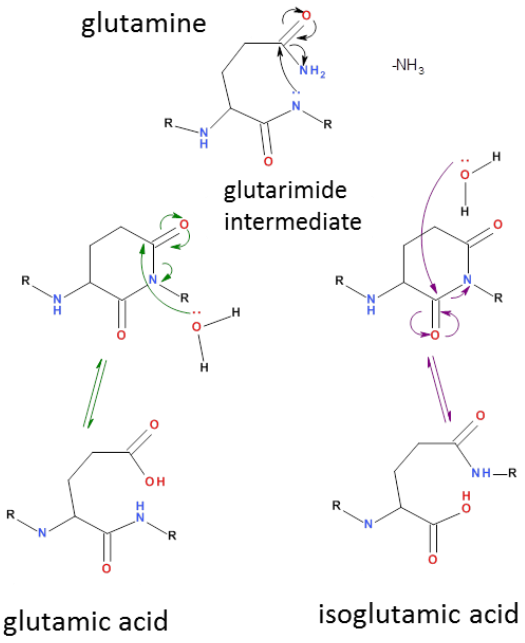


Figure 4: Glutamine undergoes deamidation via a 6 membered glutarimide intermediate, by which two isomers can form, glutamic or isoglutamic acid (Li *et al.,* 2010)

Although it is thought that glutamine has a significantly longer half-life than asparagine, the exact rate of glutamine deamidation is unknown. Capasso *et al.* (1991) stated that one of the reasons why glutarimide may undergo cyclization at a slower rate than that asparagine is due to the difficulty in obtaining a six-membered imide ring.

1.3.2 Mechanisms of asparagine deamidation

Asparagine can deamidate via a five membered succinimide intermediate (Figure 5). Asparagine deamidation has been investigated though the study of synthetic peptides. Geiger and Clarke (1987) found that asparagine had a half-life of just 1.4 days at 37 °C; the authors concluded that asparagine residues were found to have rapid deamidation rates in comparison to those measured in glutamine residues. This would suggest that asparagine would not be an ideal residue choice for measuring long term collagen degradation. It should be noted that deamidation of residues in short synthetic peptides maybe be increased due to increased flexibility, which may not be comparable to larger intact proteins (Geiger and Clarke 1987).

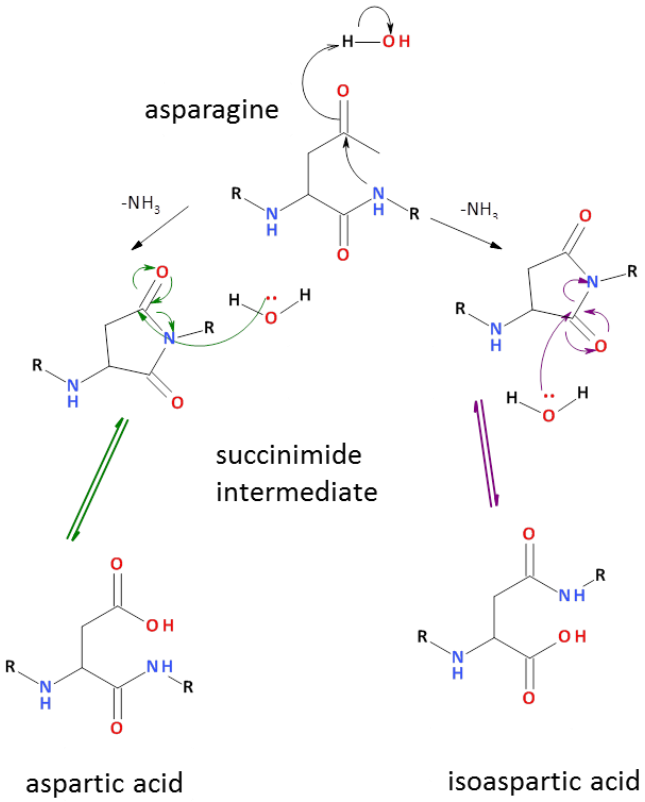


Figure 5: Asparagine undergoes deamidation via a 5 membered succinimide intermediate, by which two isomers can form, aspartic acid and isoaspartic acid (Hurtado and O'Connor 2012)

1.3.3 Previous studies on deamidation

The measurement of the non-enzymatic deamidation, of both glutamine and asparagine in human tissue has long been associated with cellular and organismal aging and degradation (Mckerrow, 1979; Geiger and Clark, 1987). When looking at deamidation in collagen, the structure of the microfilbril may be an important factor to consider, when looking at the ability of residues to undergo deamidation. The position of a residue, not only within the polypeptide but also within the

microfibril, will in turn influence what steric neighbours it has. For example, residues in the gap region may be more exposed than those in the overlap region (due to adjacent collagen helices).

This post translational modification has been the focus of a number of clinical studies, in particular those investigating deamidation in the α-crystallin eye protein (Takemoto and Boyle, 1998; Flaugh *et al.,* 2006). Deamidation in eye protein has been found to reduce the transparency of the eye lens, resulting in the formation of cataracts (van Kleef *et al.*, 1975; Masters *et al.*, 1978; Robinson, 2006; Moreau and King, 2012). As this modification has been found to occur over time it has also been an area of interest to those developing and storing protein based pharmaceutical formulations (Cleland *et al.*, 1993). Deamidation has also been of interest to those carrying out shotgun proteomic experiments, as the mass shift caused by deamidation can result in misidentification of peptides (Stevens *et al.*, 2008). The role of deamidation in cancerous cells has also been investigated. Takehara and Takahashi, (2003) found that deamidation in hepatocellular carcinomas was lower than in surrounding healthy liver tissue, indicating that the suppression of deamidation may play a role in the regulation of cell death. However, in order to better understand the impact of deamidation on human diseases the sites of deamidation need to be accurately identified (Piliang *et al.*, 2011).

Despite the large amount of clinical based research which has linked deamidation with aging, comparatively few studies have looked at deamidation occurring in protein post-mortem. Post-mortem deamidation has been investigated in a range of biological material such as: keratin in wool (Araki and Moini, 2011), collagen in bone (van Doorn *et al.*, 2012; Hurtado and O'Connor, 2012; Wilson *et al.*, 2012), protein binders in paint (Leo *et al.,* 2011), and keratin in mummified skin (Mikšík *et al.,* 2014). However there have been no robust studies which have evaluated deamidation in a range of dated material. Whilst analysing collagen peptides for the purpose of species identification, van Doorn *et al.*, (2012) reported that there may be a correlation between deamidation and the thermal age of bone. However large levels of variation were observed in samples from the same site, with the same thermal age. Araki and Moini (2011) measured the deamidation of asparagine in wool samples dating from the present to ~400 years BP and found the levels of deamidation of asparagine correlated with age; however, both of the peptides reported in this paper (YSCQLNQVQSLIVSVESQLAEIR and SQQQEPLVCPNYQSYFR) also contained multiple glutamine residues. The methodologies used in this these are not able to determine the origin of deamidadtion in samples which contain both Gln and Asn residues. Mikšík *et al.*, (2014) measured deamidation in samples of bone (rib) and muscle tissue taken from the mummified remains of the prince of Cangrande, Lord of Verona (1291–1329 AD). They found that up to 92 % of the measured asparaginyl and glutaminyl residues were deamidated, and the study concluded that deamidation is likely to be dependent on both the burial conditions and the thermal history, and that levels of deamidation cannot be used as a precise molecular clock.

Few studies have used short synthetic peptides to investigate the kinetics of glutamine deamidation although the reported half-lives of Gln obtained through kinetic studies is contradictory. Li *et al.* (2010) found that at a pH of 7.4 at 37 °C, the two peptides studied were found to have average half-lives for Gln deamidation of approximately 663 and 389 days, whereas Scotchler and Robinson (1974) found half-lives for glutamine at similar pH (6.6) of around 1,146 days. Hurtado and O'Connor (2012) measured asparagine deamidation in protein extracted from bone and found that in three of the 30 tryptic peptides identified, a first order kinetic mechanism was identified with half-lives of the three peptides at ~ 62 °C ranging from 2000 to 6000 s (pH 7.5).

1.4 Background to the Quaternary; a note on age terminology

The Quaternary spans the late part of the Cenozoic Era, and is used to describe the most recent of the geological time frames; spanning from the present day to ~2.5 million years ago (Figure 6). The Quaternary consists of two epochs: the Holocene, spanning the last 11.7 ka, and the Pleistocene (11.7 ka – 2.6 Ma) (dates reported by the International Commissions on Stratigraphy, v 2015/1). The Quaternary has been characterised by its climate change, with glacial (cold) and interglacial (warm) periods recorded in the oxygen isotope ratios of foraminifera, allowing the development of the marine oxygen isotope stratigraphy (Emiliani, 1955; Schreve, 2001; Railsbeck *et al.,* 2015). Glacial and interglacial periods are denoted by a marine oxygen isotope stage (MIS) with MIS 1 representing the present day. Within each stage there may also be sub-stages; these are now most commonly denoted by lower case letters, such as those in MIS 5 (warm: a, c and e, and cold: b and d). There have been some inconsistencies in the nomenclature used in this field, particularly when defining and naming sub-stages (Railsbeck et.al, 2015), so here the system of Railsbeck *et al.*, (2015) is followed.

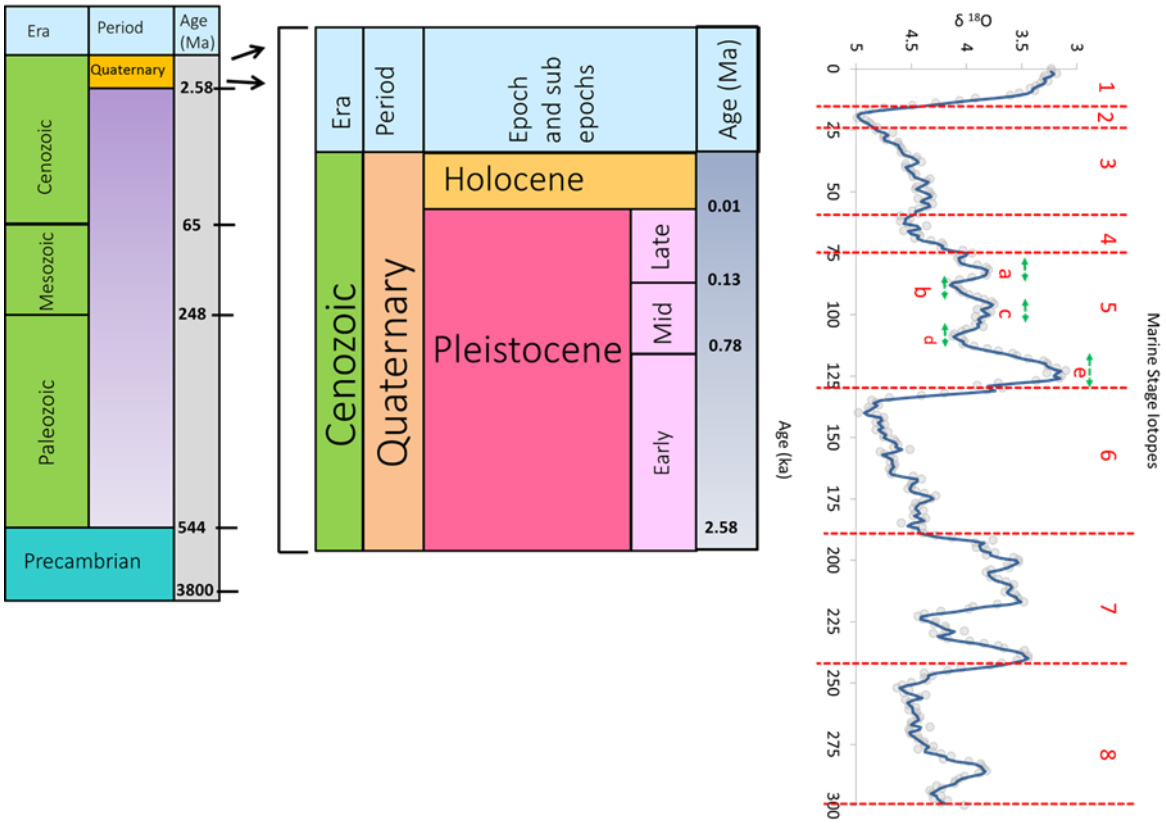


Figure 6: Schematic of the eras, periods and epochs which span the Quaternary (as reported by the International Commission on Stratigraphy, 2015) (left). MIS record from present day (MIS 1) to 300 ka BP (MIS 8), data taken from Lisiecki and Raymo (2005) (right)

1.5 Dating methodologies

There are a number of different strategies for estimating age in the Quaternary. The type of method used will depend on the type and estimated age of material in question, and each dating methodology has associated with it varying levels of analytical, statistical and human experimental error. The level of precision of a particular method is often limited by the available instrumentation and the understanding of the mechanism of the change being measured. The accuracy of a dating technique is a measure of how close the date generated is to the true calendar age of the artefact. When dating old sites, such as those from the Palaeolithic, it is often beneficial to use a multidisciplinary approach; this means where possible analysing more than one sample type, and/ or using more than one dating technique. The following dating methods are briefly introduced in this chapter: radiocarbon dating (Wood *et al.*, 2015), amino acid racemization (Bada *et al.*, 1973; Collins *et al.*, 2009), uranium-thorium dating (Towe, 1980; Sykes *et al.*, 1995), luminescence dating (Huntley *et al.*, 1985) and biostratigraphy (Schreve, 1997).

1.5.1 Methodologies which have been applied to the direct dating of bone

1.5.1.1 Radiocarbon dating

Radiocarbon dating is used to directly date bone material. It is a radiometric dating method which measures the conversion of unstable ^{14}C to stable ^{13}C over time in organic material. ^{14}C is produced in the atmosphere and is acquired by animals via either the consumption of plants, and/or other animals. Death of an organism stops the uptake of ^{14}C and therefore starts the radiocarbon clock. The amount of ^{14}C remaining in a bone sample is representative of the length of time which has occurred post mortem. This rate of decay from parent (^{14}C) to daughter (^{13}C) isotope is a constant, with a half-life of ~ 5700 years. This means that this technique could theoretically be used to date material from present day, to ~60 ka BP. In practise however, bone material can only be accurately dated to ~ 50 ka BP (van der Plicht and Palstra, 2014). One of the key problems with dating older material, near the radiocarbon limit is that there is very little ^{14}C remaining. This means that the introduction of modern contamination can greatly affect the calculated age; for example Wood *et al.* (2010) reported that the addition of 1% modern carbon to a sample which contained no ^{14}C would result in an estimated age of 37 ka BP. Advances in sample preparation such as ultrafiltration techniques (Brock *et al.*, 2007; Ramsey and Higham, 2007; Dobberstein *et al.*, 2009; Grün *et al.*, 2010; Wood *et al.*, 2010) have helped to reduce levels of modern contamination.

Although originally it was thought that the level of ^{14}C in the atmosphere was constant, it was later found that the levels of ^{14}C fluctuate and therefore a correction factor is needed to account for this (Renfrew and Clarke, 1974). This is done through a series of experimentally constructed calibration curves (Reimer *et al.*, 2013). The database used currently to calibrate radiocarbon ages is INTCAL13 (Reimer *et al.*, 2013).

The ability to determine the age of a bone using radiocarbon dating relies on the assumption that levels of ^{14}C measured in the sample are equal to those in the atmosphere. This assumption is true for terrestrial animals which feed on flora and fauna supported by carbon in the atmosphere. However, the source of carbon available to marine animals has a more complicated origin. This is due to the fact that the exchange of ^{14}C takes place at the water surface. Deeper waters, such as those that occur in the polar-regions, undergo gaseous exchange as well as cooling. The resultant increase in density causes the water to sink and therefore contact with the surface is reduced, which results in depleted levels of ^{14}C in more dense water ("marine reservoir effect"). Unlike in the atmosphere, the mixing in oceans is comparatively slow, resulting in both depth and geographical variations (Ascough *et al.,* 2005). The depletion of ^{14}C in marine samples can therefore result in radiocarbon dates that are older than the real age of the samples. It has also been found that the consumption of marine foods such as molluscs can also cause depleted levels of ^{14}C. This depletion is known as the reservoir effect (Fernandes and Bergemann, 2012).

1.5.1.2 Amino acid racemization

Amino acid racemization (AAR) is a biomolecular method that has been used to estimate the age of bone material (Bada *et al.,* 1973) although it is now widely accepted that this method is not suitable for bone material. Louis Pasteur first discovered that aspartic acid was optically active in 1851 (Robinson and Robinson, 2004). Since then there have been significant advances in the understanding of the mechanisms by which amino acids undergo racemization, in both synthetic peptides and structurally constrained chains (Geiger and Clark, 1987), such as those present in the triple helix of collagen. Progress has been made in deducing and understanding the mechanisms by which both glutamic acid and aspartic acid can undergo racemization.

It was previously thought that bone material could be dated by fitting a first order reversible kinetic model to the extent of racemization of aspartic acid in time (Bada and Schroeder, 1975), and calibrating the data using radiocarbon dated samples collected from the same site location. It is important that both the unknown-age samples and independently-dated bone have been subjected to the same temperature changes during their deposition, as temperature is known to directly affect the rate in which racemization occurs (Bada and Protsch, 1973). As temperature affects the rate of collagen degradation, the thermal history of a sample is a key influencing factor on collagen survival (Collins *et al.,* 2002).

The dating of bone material in the early 1970s was one of the initial applications of AAR to archaeological science. The technique was used to date skeletal remains found at La Jolla California (Bada *et al.*, 1974), yielding an exceptionally early (40,000 – 60,000 years old) date for hominin remains. It was later discovered that the age estimated for these skeletal remains was incorrect (Bada *et al.*, 1984; Taylor *et al.,* 1985), partly due to inaccurate ^{14}C dates obtained from the reference material used to calibrate the results (Bada *et al.,* 1984). These controversial claims at La Jolla did little to promote the successful application of the method (Johnson and Miller, 2007) and sparked a debate within the scientific community on the suitability of bone material for AAR analysis.

1.5.1.2.1 Open vs closed systems

In terms of using amino acid racemization for dating purposes, one area offocus has been on the issue of open and closed systems within biominerals (Towe, 1980; Sykes *et al.*, 1995). A closed system is one in which the organic material to be analysed is protected and cannot interact chemically or physically with the external environment (Penkman *et al.*, 2008). The only two variables which will affect the rate of racemization in a closed system are time and temperature (Collins *et al.*, 1999). Examples of closed system behaviour include the intra-crystalline fractions within mollusc shells and eggshell. In some species protein is encased within the intra-crystalline

fractions due to the biomineralisation mechanisms, and are therefore protected from the effects of environmental conditions such as fluctuating moisture and pH (Brooks *et al.*, 1990; Penkman *et al.,* 2008). Bone and other collagen based skeletal tissues (dentine, antler, ivory) are not closed systems. They represent a mineralized organic framework of collagen, which is open. As a consequence there is a slow exchange with the environment, resulting in loss of amino acids and uptake of contaminants (e.g. humic and fulvic acids) from the burial environment (Brooks *et al.,* 1990; Collins *et al.*, 1999; Marom *et al.*, 2012). This open system of bone also causes problems when trying to model how collagen degrades, and in turn using the degradation to estimate age. Mollusc shells are pH buffered, semi-closed systems and are therefore more likely to provide consistent diagenetic environments (Sykes *et al.,* (1995); Penkman *et al.,* (2008). Subsequent research has tended to agree arguing that conformational constraints on aspartic acid racemization in the collagen triple helix mean that racemization in bone is controlled by both the rates of denaturation and leaching (Collins *et al.,* 1999; 2009), although this view is not universal.

1.5.1.3 Uranium series

Uranium series dating looks at the equilibrium between Th^{230} and U^{234} as they decay into a series of unstable radioactive isotopes, This process continues to eventually form non-reactive lead isotopes (^{206}Pb, ^{207}Pb and ^{208}Pb) (Lowe and Walker, 1997). This decay pattern has been used to provide age information, and is commonly referred to as uranium series dating. This technique is most accurate when used to date chemically precipitated calcium carbonate minerals such as coral and speleothem (Schwarcz, 2005). This technique has been used to date bone material; however, the measurement of uranium in bone material is less reliable (Schwarcz, 2005) and has been described as being problematic, associated with large errors since its inception (Grün, 2006). This is largely due to the fact that bone is an open system, meaning that early uptake and leaching of uranium can occur (Pike *et al.*, 2002). In closed system carbonate materials the intermediate isotopes are described as being in a secular equilibrium, meaning each isotope decays at the same rate at which it is produced (Hajdas *et al.*, 2007). Natural processes result in the fractionation of these elements, for example in water Th is separated from U due to differences in solubility. This results in precipitated calcite material containing very low levels of Th, with the decay of ^{234}U to ^{230}Th increasing over time. Uranium series dating can been used to date carbonate materials such as stalagmites and flowstones to 500 ka BP (Edwards *et al.*, 1987), including a number of cave sites analysed in Chapter 7.

1.5.2 Dating methodologies which have been used to date the context in which they were buried

1.5.2.1 Luminescence dating

Luminescence measurements are made on mineral grains such as quartz and feldspar. These mineral grains have acquired energy in the form of trapped electrons though the exposure of light. Minerals which are able to acquire such energy are known as dosimeters. The process works as electrons are trapped within the crystalline structure and can then be released using light of a single wavelength. This causes the trapped electrons to re combine within the crystal to produce energy, which is then measured in the form of the emission of light of a different wavelength to that used to initially stimulate the grain. This technique is used to look at sedimentary deposits using quartz. As well as being resistant to chemical weathering quartz is also one of the most common detrital minerals at the earth's surface.

Grains of quartz and feldspar produce energy from the natural decay of yttrium, thorium and potassium. It is this energy which is stored as a trapped charge. This trapped charge can be emptied, and this is described as optical bleaching. Two common factors which can cause this are heat and light (Duller, 2004; Lian and Roberts, 2006). The charge is built up from the time of burial. This means that the total amount of charge a grain has is directly proportional to the length of time it has been buried. In order to measure the charge of an object it must be stimulated by light of a single wavelength; this results in the admittance of a different wavelength of light which is collected and measured and is termed optically stimulated luminance (OSL) (Huntley *et al.,* 1985). Because the intensity of the OSL signal is related to amount of trapped charge it can be used to calculate age (Duller, 2008). OSL dating has been used to date material up to around 150 ka BP (Duller, 2008), including sites in Africa dated s up to ~ 151 ka BP (for example Jacobs *et al.*, 2006).

1.5.2.2 Biostratigraphy

Biostratigraphical evidence from mammalian fossil assemblages has proven to be a useful tool by which to investigate the nature of climatic changes across the Pleistocene period in Britain (Schreve, 1998). The principles behind this technique utilise the extinction and rapid turnover of mammalian linages which occurred throughout the Pleistocene period, due to the intensely fluctuating environmental conditions identified by past interglacial periods. It is therefore thought that each climatic cycle has with it an associated unique suite of mammals. An example of this would be the presence of hippopotamus in the last interglacial (Stuart, 1976). This biostratigrahical framework is established by gathering information on the mammalian remains present at different sites which have with them an associated interglacial period. Data from multiple sites can then be used to reconstruct the inhabitants of interglacial throughout the quaternary (Shreve, 2001). It

should be noted that the absence of evidence (i.e. the absence of a particular mammal) is not evidence of absence (Lister, 1992; Currant and Jacobi, 2011). A good example of an established biostratigraphic framework is that obtained from the Thames Valley sequence, which has been described in the literature as being the most reliably-dated long terrestrial sequence in Britain (Bridgland, 1994; Schreve, 2001).

1.6 Analytical instrumentation

Appropriate analytical instrumentation is crucial to the successful analysis of modifications in collagen. Both chromatographic and mass spectrometric techniques have been widely reported in relation to the study of peptide mixtures, specifically for the measurement of deamidation in degraded proteins (Buckley *et al.,* 2009; Araki and Moini, 2011; van Doorn *et al.,* 2012; Wilson *et al.*, 2012; Welker *et al.,* 2015)

1.6.1 Mass spectrometry

Mass spectrometry can be used to analyse a wide range of both organic and inorganic compounds. The main advantage of the technique is its capability of high mass accuracy measurements, in samples of very low concentration. This type of instrumentation has been widely applied to a number of 'omics' fields, such as the study and function of proteins, in particular via the analysis of peptides. Simply put, mass spectrometers are made up of three main components, an ionisation source, mass analyser and ion detector. The sample analytes are converted into a gas phase ions (anions or cations, depending on the mode being used) in the ionisation source. The mass analyser then separates ions based on their mass to charge ratios (*m/z*). Depending on the type of instrument used, the ions can then be passed to the detector; for example when using a (TOF-MS) each time an ion reaches the detector a signal is generated. The type of signal generated is dependent on the mass analyser. Other types of instrumentation, such as (FT-ICR-MS) have a slightly more complex set up, whereby the signal is transposed before being presented as a mass spectrum. Each of the instrument components comes in a range of types, and can be coupled in different combinations depending on the requirements of the analysis and the chemistry of analytes being analysed. As well as determining the accurate mass of compounds, coupling of two mass analysers to perform tandem MS (MS/MS) can provide information about the structure of molecules (such as peptides), by looking at fragments produced following, most typically, collision induced dissociation and recording of a product ion spectrum.

1.6.1.1 Ionisation sources

In order for analytes to be guided through the mass spectrometer by electrical or magnetic fields they need to be charged. Samples can be ionised via a number of routes and the choice of ionisation source is largely dependent on the sample type and the chemical properties of the analytes, such as polarity, volatility, or how thermally labile they are. In the work presented in this thesis, two soft ionisation methods have been used; matrix-assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI). Soft ionisation methods are so called because they induce minimal fragmentation of the analytes during ionisation.

1.6.1.1.1 Matrix-assisted laser desorption/ionisation

The use of lasers to generate ions dates back to the 1960s (Jagtap and Ambre, 2005). Originally one of the main limitations of laser desorption was that it was not possible to ionise samples with a mass greater than 2000 Da (Bahr *et al*., 1994), because the increase in laser energy needed to ionise the larger molecules also induced fragmentation. As well as the fragmentation of analytes there were also issues with analyte mixtures of components with different light absorption characteristics. These problems were solved with the introduction of a light absorbing matrix added to the sample. When preparing samples for MALDI-MS analysis, the sample solution is mixed with a solution of matrix, either before or after spotting the sample onto the MALDI target plate, to produce a sample-matrix co-crystal. Different types of target plate formats are available depending on the amount and type of sample being analysed. There are a number of different matrices commercially available (Zenobi and Knochenmuss 1999). The sample is mixed with the matrix solution, with the matrix being in excess over the analyte. The mixture is then left to air dry on the target plate. As the solvent evaporates, the sample co-crystallises with the matrix. Mixing the sample with matrix is not only thought to allow the analytes to be desorbed and ionised irrespective of their individual light absorbing characteristics (Bahr *et al*., 1994), but also to protect the sample from the intense laser pulses (commonly nitrogen or Nd:Yag lasers) used during desorption (Karas *et al*., 1985; 1987), and thus reducing unwanted fragmentation of the analytes. The laser excites the matrix molecules which causes heating of the crystals. Once excited, the matrix molecules cause both the sample and matrix to be desorbed from the plate into an expanding gas plume (Jagtap and Ambre, 2005). Although there are many theories which explain how analytes may become charged (Zenobi and Knochenmuss 1998), proton transfer in the gas phase is the most widely accepted (Cole, 2010). It is thought that the matrix absorbs energy from the laser and transfers it to the analyte, also acting as a proton donor or acceptor (Figure 7). When using MALDI often only singly charged ions are formed. The reasons for this are not fully understood, with the most recent explanation put forward by Karas *et al*. (2000). This theory has

been termed the 'lucky survivor' hypothesis. Karas *et al*. (2000) propose that during MALDI, highly efficient ionisation occurs resulting in survival only of singly charged ions.

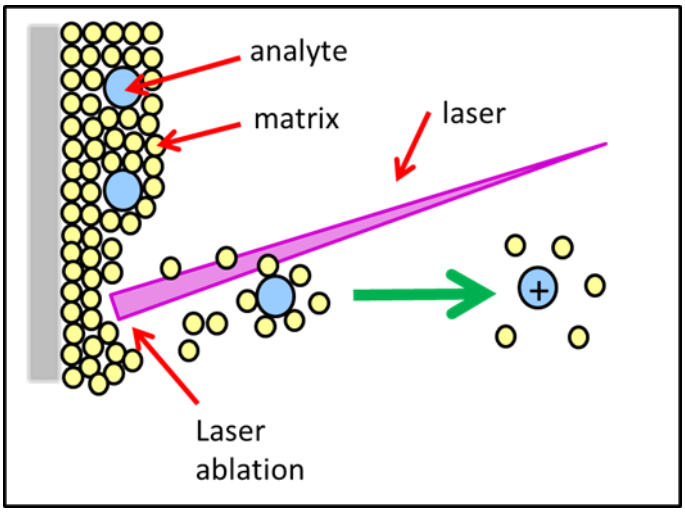


Figure 7: Schematic showing the ionisation of analytes using matrix-assisted laser/desorption ionisation. The sample is introduced into the ion source on a steel plate, under vacuum. Intense laser impulses cause the sample and matrix to be ablated from the plate. The analyte is then charged by proton transfer from the matrix in the expanding gas phase. The charged analyte is then passed to the mass analyser

1.6.1.1.2 Electrospray ionisation

Electrospray ionisation is described as being the 'brainchild' of Malcolm Dole (Kebarle and Verkerk, 2009). Although Dole did not apply this technology to mass spectrometry, he is noted as first discovering the possibilities of electrospray as an ionisation source in the 1960s (Dole *et al*., 1968). Interestingly Dole came across this technology by accident while visiting a car manufacturing company which was using electrospray to paint cars. Dole found that the small droplets of paint were attracted to the charged metal surface of the car (Kebarle and Verkerk 2009). One of the main limitations early on in the field of mass spectrometry was its inability to analyse large biomolecules, as the energy required to ionise them also caused significant fragmentation (Fenn, 2002). Work carried out by Fenn and Yamashita in the 1980s (Yamashita and Fenn 1984) addressed this problem and their work enabled ESI to be used for large molecules. Today one of the main benefits of this type of ionisation is that it can be used to analyse large biomolecules. This is due to the ability to create multiply-charged ions – since mass spectrometers measure *m/z*, high mass (m) analytes receiving many charges (z) are ionised without having very high *m/z* values. The premise of ESI is that it converts liquid samples into a charged gas phase sample, without fragmenting the analytes. During ESI a liquid sample is introduced into the source via a needle or a capillary. A positive charge is applied to the capillary, as the potential gradient increases the sample solution is forced out of the capillary mouth into a cone-like shape, termed a Taylor cone (Figure 8) (Zeng

2011). Opposite the capillary is a skimmer cone, to which a negative charge is applied. The increasing potential gradient causes an increase in repulsion of the positive ions until the Rayleigh limit is reached, resulting in the formation of a spray of charged droplets. Evaporation of the solvent from the droplets causes the droplets to become smaller and the charges to repel each other more. These droplets are guided towards the skimmer cone, before being converted into gas phase ions. There are two proposed routes by which the analytes in the charged droplets can be converted into gas phase ions; either via the charge residue mechanism, (CRM; (Dole *et al.*, 1968), or via the ion evaporation mechanism (IEM; Iribarne and Thomson, 1976).

During the CRM ionisation mechanism, it is hypothesised that the solvent in the charged droplets evaporates and the droplets become smaller, whilst maintaining the same number of charges. This continues until enough evaporation occurs that the surface tension of the droplet can no longer sustain the charge; at this point the droplet breaks into smaller droplets with reduced charge, eventually resulting in charged gas phase ions. During the IEM mechanism, it is proposed that ions are ejected from the surface of the droplet due to the high concentration of charge on the surface. Once the gas phase ions have been generated they are introduced into the mass spectrometer via the skimmer cone.

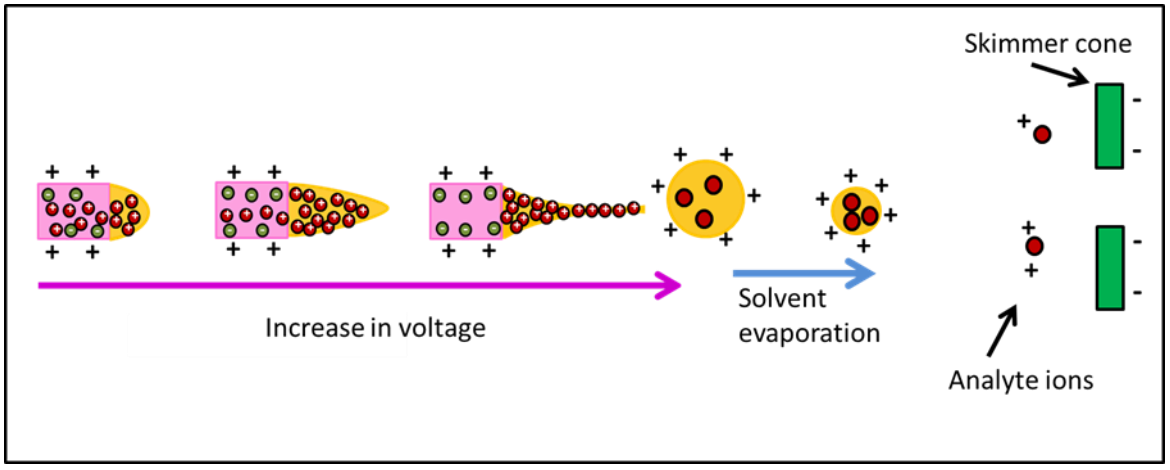


Figure 8: Schematic of analyte ionisation via ESI in positive ion mode. The sample is introduced as a liquid in a capillary. A voltage is applied to the capillary and the resulting potential gradient causes the sample solution to be forced out of the capillary mouth, this cone shaped formed by the sample solution is termed the Taylor cone formation. This continues until the Rayleigh limit is reached, at which point the formation and expulsion of small charged droplets occurs. These droplets are guided towards the skimmer cone by the increasing potential gradient and are convert into gas phase ions via either CRM or IEM mechanisms

1.6.1.2 Mass analysers

Once the analyte has been ionised it is guided to the mass analyser where ions are separated based on their *m/z*. When choosing a mass analyser it is important to consider the level of mass accuracy and resolution required of the data.

1.6.1.2.1 Time of flight

The concept of the time of flight mass spectrometer was first introduced by William Stephens at the American Physical Society meeting in Cambridge in 1946 (Stephens 1946). Shortly after this there were a number of advances in the field, with the first published design of a linear TOF in 1948 (Cameron and Eggers Jr., 1948). In 1955 Wiley and McLaren (Wiley and McLaren 1955) published a linear TOF-MS design which was to become the first commercial instrument.

The time of flight mass analyser, as the name suggests, separates ions based on the length of time they take to travel along a field-free drift tube. After the sample is ionised, the ions are accelerated by a voltage pulse into the flight tube in which they are separated. Because all the ions of the same charge are accelerated with the same amount of energy, the different mass ions travel at different velocities. This means the lighter ions travel faster than the heavier ones and thus reach the detector first (Figure 9).

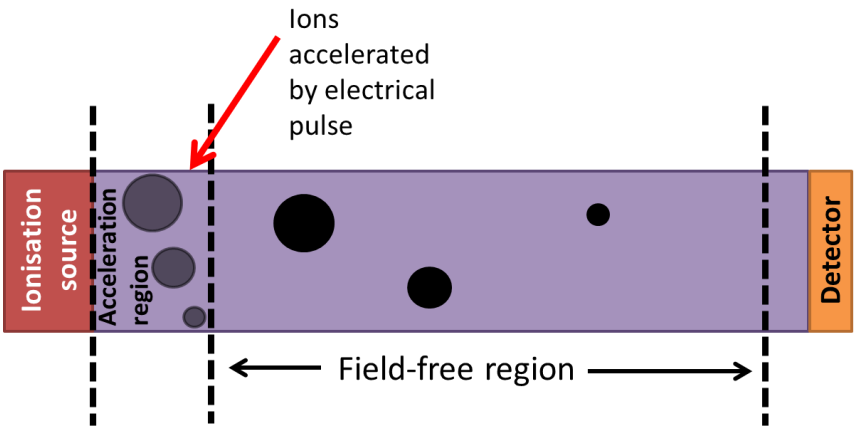


Figure 9: Schematic of a time of flight mass analyser separating ions with difference *m/z* ratios

An original drawback of the TOF was that it didn't produce very high resolution. Since the original design in 1948 there have been many improvements made in order to obtain higher mass accuracy results; one of these was the introduction of the mass reflectron (Figure 10) in 1973 by Mamyrin *et al*. (1973). Sometimes ions with the same *m/z* values can have slightly different kinetic energies, which in turn can alter the time it takes for them to travel along the flight tube. In order to correct for such slight differences in kinetic energy a reflectron can be used. When using a time of flight instrument coupled with a reflectron to correct for this spread, analytes of the same *m/z* but with a

higher kinetic energy spend longer in the reflectron than those with lower kinetic energy. If appropriate voltages to apply to the reflectron plates are chosen this results in analytes with higher kinetic energies having an extended path length, and therefore reaching the detector at the same time as those with lower kinetic energies, so correcting for the initial energy spread. The TOF-MS used in this thesis was the Bruker ultraflex III which is fitted with a reflectron. Since the introduction of the reflectron there have been a number of other advances in TOF design with the introduction of the orthogonal TOF (oaTOF) and TOF/TOF instruments.

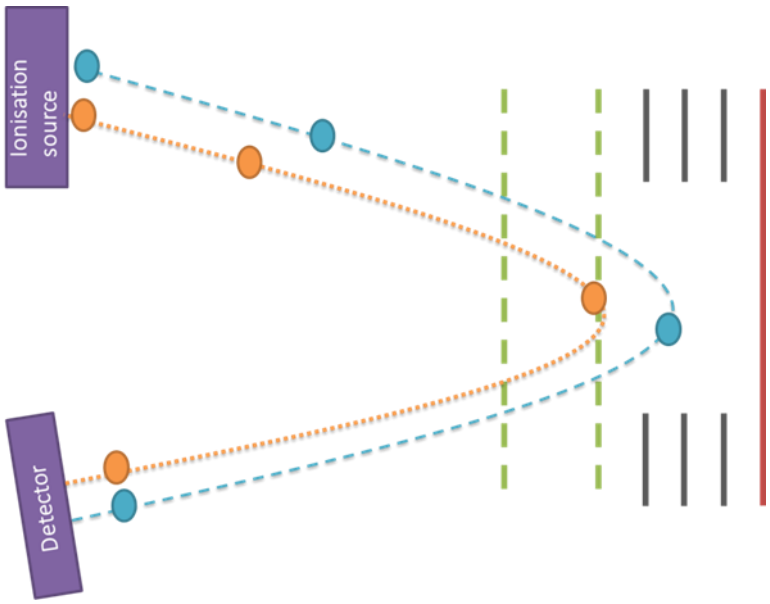


Figure 10: Schematic showing the use of a reflectron to correct for analytes with the same *m/z*, but different in kinetic energies. The analyte with the higher kinetic energy penetrates deeper into the refelectron, thus extending the path length. This results in increased resolution as both analytes arrive at the detector simultaneously

One of the main instruments used throughout the work presented in this thesis is the Bruker ultraflex III TOF/TOF mass spectrometer. This instrument was chosen for a number of reasons: it is widely commercially available, has a user-friendly interface and the analysis costs are lower than those for ultra-high resolution instruments such as a FT-ICR-MS, which makes it more accessible to archaeologists (rather than mass spectrometry specialists). The ultraflex is capable of automated MS and MS/MS analysis and is therefore described as a high throughput instrument. An advantage of the ultraflex III analyser is the presence of a lift cell which is located in the flight tube. The lift cell allows the acquisition of a full fragment ion spectrum with a single set of V settings (Figure 11).

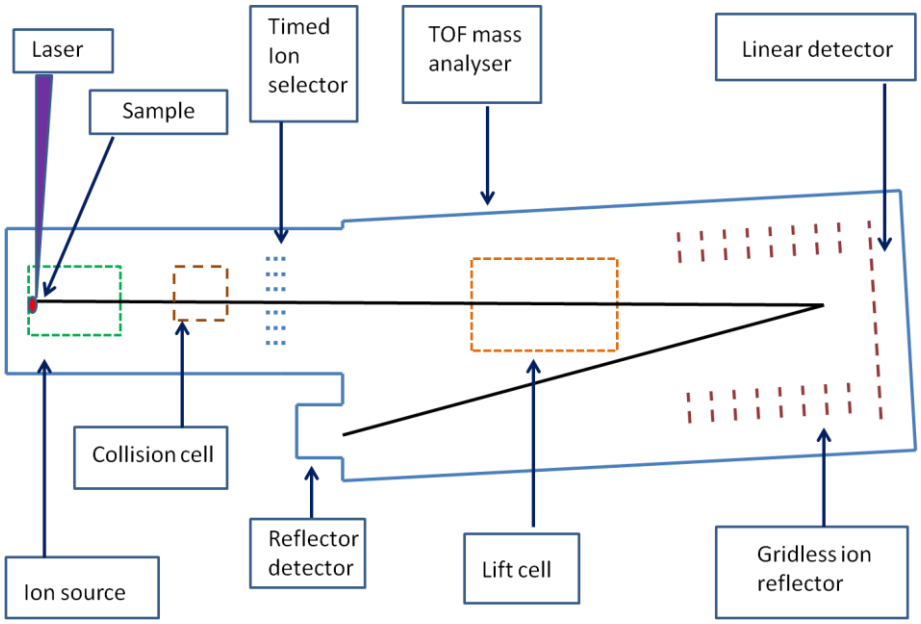


Figure 11: Schematic of the Bruker ultraflex TOF/TOF mass spectrometer

This ultrflex TOF/TOF instriment has a fixed MALDI source. In order for the MALDI TOF/TOF to overcome the issue of ions with broad KEs, the ions need to be focused prior to detection. This was not possible using the classic reflectron directly. During ionisation there are a number of factors which can lead to poor resolution. Not all of the ions are desorbed from the plate and ionised at the same time; in addition the analytes can have different starting velocities, can be accelerated in different directions. There is also the issue of the anlaytes being in different positions within the source when the extraction voltage is turned on, even if all of these factors are eliminated collisions within the field free region of the plume can cause issues as these fragments maintain the same velocity as the precursor ion but have a wide range of KEs. All of these things can result in a reduced level of resolution. In order to try and correct for this, delayed extraction can be used. Inside the ion source there are three components which are used during delayed extraction (Figure 12): the first is the target plate on which the analyte is spotted (Figure 12, P1), the second is the voltage plate with an electrode which sits opposite the sample plate (Figure 12, P2), and the third component is a grounded acceleration electrode. During delayed extraction the laser is fired which starts the ionisation process and sets the analyte molecules/ions in motion. A potential is pulsed between P1 to P2 generating an electric field. This electric field causes the ions to move towards P2. During this process the ions with higher energetic energy move more quickly towards P2 and are therefore exposed to less of an electrical potential than the slower ions. This results in the slower ions being accelerated, and the faster ions moving more slowly. The voltage potential between P1 and P2 can be adjusted so that ions with the same mass, but different starting velocities, will reach the detector at the same time thus improving the resolution.

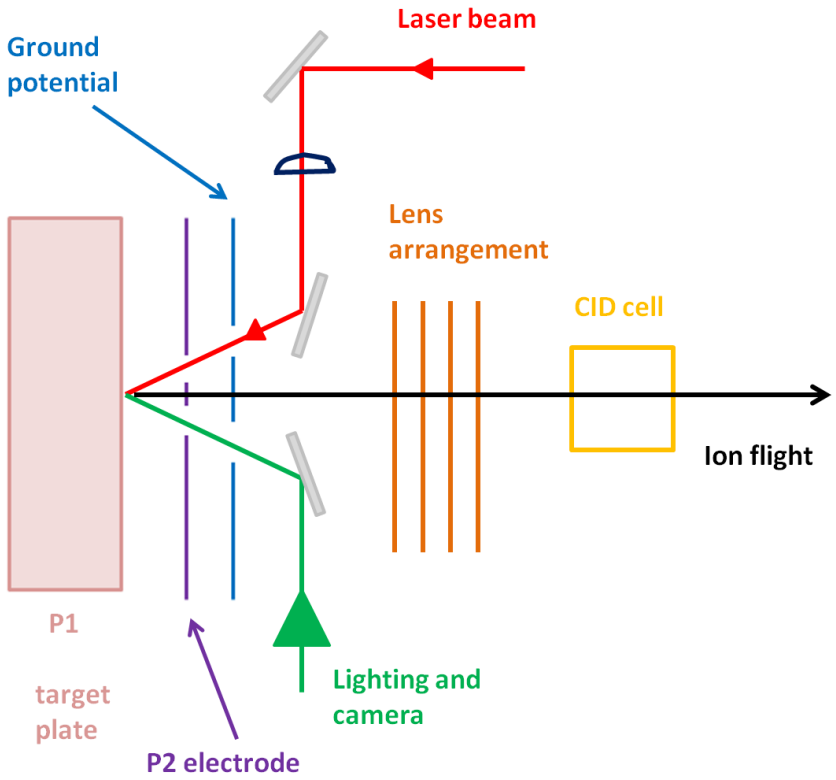


Figure 12: Schematic showing the basic components used during delayed ion extraction

To record a product ion spectrum, precursor ions may be selected on the basis of their *m/z* ratio. These ions can then be fragmented, and the resulting product ions used to deduce the chemical structure of the precursor. The precursor ion selector contains deflector plates which are arranged in vertical layers; known as a TIS gate (Timed Ion Selector gate). Electrodes are coupled to a high voltage supply with alternating polarity. The potential difference between the plates results in an electrostatic field which is perpendicular to the direction of the ion flight path. This electrostatic filed is used to deflect ions. Once the precursor ions and fragments leave the TIS gate they enter the Lift cell.

The lift cell is made up of four electrodes that which form three chambers within the cell, the lift chamber, pulsed acceleration stage and the acceleration stage. Electrodes one and two are always connected and form a cell which shields the ions. Precursor and fragment ions formed in the drift region of the instrument travel with the same velocity though the time of flight tube. The LIFT cell enables a complete product ion spectrum to be obtained from a single analysis by enabling the separation of precursor and fragment ions with the same velocity. The LIFT cell does this by raising the kinetic energy of the precursor and fragment ions so that the energy difference between the two does not exceed 30%. The reflectron is then able to successfully focus and reflect both the fragment and the precursor ions.

The ultraflex also has a second deflection unit called the post lift metastable suppressor (PLMS). This deflection unit is situated between the LIFT cell and the reflector and can be used to remove

fragment ions that have formed after the LIFT procedure so that only those formed between the source and the LIFT cell are separated and passed to the detector.

1.6.1.2.2 Fourier-transform ion cyclotron resonance

Fourier-transform ion cyclotron resonance mass spectrometers (FT-ICR-MS) separate ions based on their cyclotron frequency in a constant magnetic field (Marshall *et al.*, 1998). The technique was first developed by Comisarow and Marshall in 1974 (Comisarow and Marshall, 1974). Today FT-ICR-MS is well known for its high resolving power with extremely high mass accuracy. Modern FT-ICR mass analysers have three main components: a super conducting magnet, ICR cell and a high vacuum system. The ICR cell sits inside the superconducting magnet. Once the ions have been generated by the ionisation source they pass through a series of pumping stages in order to reach high vacuum. Once at high vacuum the ions enter the cell. There are trapping plates at each end of the cell to prevent the ions leaving (Figure 13).When the ions enter the ICR cell, the magnetic field confines them to a circular orbital motion, perpendicular to the field. The frequency of the ions' orbit is the ion cyclotron frequency and is dependent on the *m/z* of the ion. Each individual *m/z* ion packet is excited by RF pulses from excitation plates outside the ICR cell. Once excited, the higher orbits of the ions induce alternating currents in the detector plates, with the frequency of the induced current the same as the cyclotron frequency. Each packet of ions at any given *m/z* value is only excited within a narrow RF resonance window. In modern instruments all of the ion packets can be excited and analysed simultaneously. The cyclotron frequency is measured as the ions pass detection plates. Fourier transformation of the free induction decay is converted from the time domain to frequency, which is then converted into a mass spectrum.

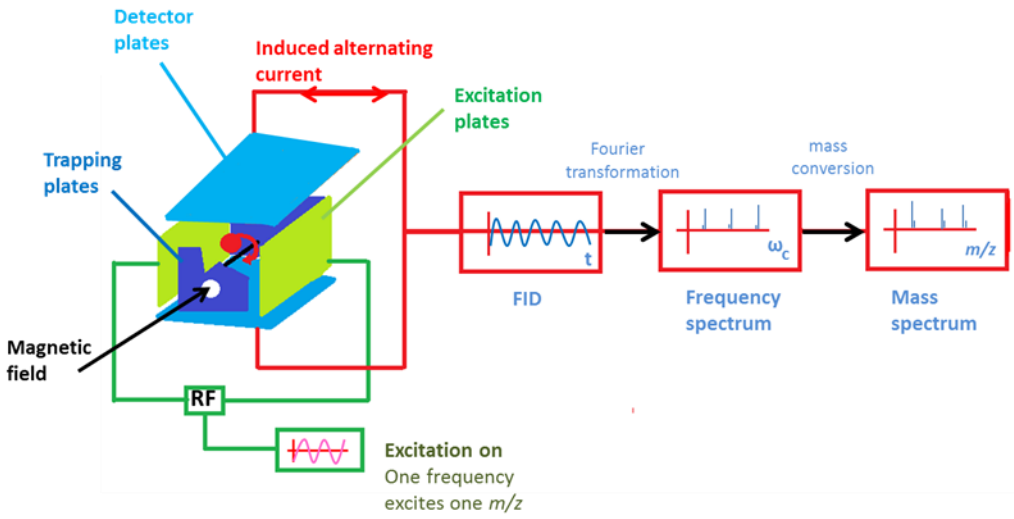


Figure 13: Schematic of a Fourier-transform Ion cyclone resonance mass analyser

1.6.2 Liquid chromatography

Mass spectrometry systems are often coupled with chromatography separation. The coupling of these techniques can aid separation of analytes before they are analysed by MS, as well as helping to identify peaks observed in chromatogram by their mass spectrum. The word chromatography has long been associated with colour, since the 19th century when it was largely associated with artists and pigments (Abraham, 2004). Chromatography in the analytical sense was pioneered in the early 1900s by the work of Mikhail Tswett, with his early studies which separated a series of extracted leaf pigments in a column packed with solid particles (Ettre and Sakodynskii, 1993). Today chromatography is a powerful analytical tool which allows the separation, detection and quantification of analytes within complex mixtures. The form of liquid chromatography used most widely today is high performance liquid chromatography (HPLC), originally termed high pressure liquid chromatography. HPLC systems have five main components: pump, degasser, injector, column and detector (Figure 14). The analytes are introduced and carried around the system by a liquid mobile phase.

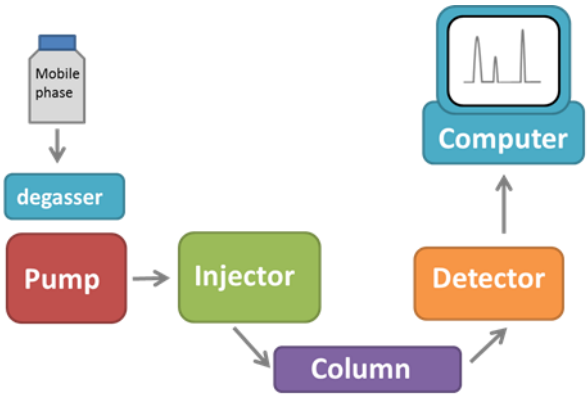


Figure 14: Simple schematic of the components of an HPLC system

During HPLC, analytes are separated based on their relative affinities for both the mobile and stationary phases (column). The choice of column and mobile phase depends on the chemical properties of the analytes. In this thesis, reversed phase HPLC was used, whereby the stationary phase is hydrophobic and the mobile phases are polar aqueous and, or organic solvents. This results in hydrophobic analytes being retained on the stationary phase of the column for longer than more hydrophilic analytes, with the hydrophilic compounds eluting first. The ratio of the percentage of organic solvent in the mobile phase is adjusted over time (gradient elution) in order to optimise separation of the analytes.

1.7 Discussion and thesis aims

As discussed here, bone material can only be directly dated using radiocarbon analysis to ~50 ka BP (van der Plicht and Palstra, 2014). The secure direct dating of bone samples older than this has proved challenging. As bone material ages, its major protein, collagen undergoes chemical diagenesis. Currently bone is not a suitable sample for many dating techniques due to the open system nature of its structure. It is hoped that isolating intact protein for the measurement of post mortem changes, may offer an insight into both the age and the preservation state of the bone. One of the ways in which collagen can degrade is via deamidation of the amino acid residues asparagine and glutamine, to aspartic acid and glutamic acid respectively. The comparatively short half-life observed for asparagine means that it is unlikely to be useful when measuring deamidation in bone material from the Pleistocene Period. It has been suggested by van Doorn *et al*. (2012) that glutamine deamidation may be an ideal marker of thermal age.

In this thesis, approaches have been developed using mass spectrometry (MS) to measure the conversion of glutamine to glutamic acid in both synthetic, and tryptic collagen peptides. Two mass spectrometers that have been used, a Bruker ultraflex III and solariX. The ultraflex instrument is a MALDI-TOF-MS, and was initially chosen for a number of reasons such as: its high sample throughput, low cost, user friendly interface and the fact that it is more commercially available to the archaeological scientific community. The solariX is a high resolution FT-ICR-MS. This instrument can be used with either an ESI or a MALDI source, and was used to investigate possible ionisation bias of Gln, or Glu-containing peptides, as well as to investigate the accuracy of the genetic algorithm used to calculate levels of deamidation in data obtained from the MALDI-TOF instrument. In addition to mass spectrometric techniques, chromatography (an Agilent 1100 HPLC.) was also employed to determine the purity of the peptides synthesised in Chapter 4.

The aims of this project are therefore to test whether the extent of non-enzymatic glutamine deamidation is a useful diagnostic tool for measuring the diagenetic history of a bone; and investigate the possible correlation between glutamine deamidation and time/thermal age, for the purpose of dating bone across the Pleistocene Period. In parallel, both high (80 °C-140 °C) and low (65 °C) temperature time course (kinetic) experiments have been performed on Roman cow bone (~800-1000 years BP) material to assess the rates of deamidation at different temperatures. With the data obtained we hope to assess the potential of deamidation as a novel method for dating Middle and Lower Palaeolithic bone material.

1.8 Calculating deamidation from TOF/TOF data

One disadvantage of the TOF ultraflex instrument used in the work presented in this thesis, is that it is not able to resolve deamidated and undeamidated peaks. This is due to insufficient resolving power of the mass analyzer. In order to calculate the extent of deamidation of glutamine, a program called *Q2E* (taken from Wilson *et al.* 2012) was used. *Q2E* uses a genetic algorithm to deconvolute the overlapping distributions. The deamidation of glutamine results in an overall mass increase of 0.984 Da, the n^{th} peak of the deamidated peptide signal (typically the mono-isotopic signal) overlaps the $(n+1)^{th}$ peak of the undeamidated form (typically the signal for the species containing one ^{13}C atom) Figure 15.

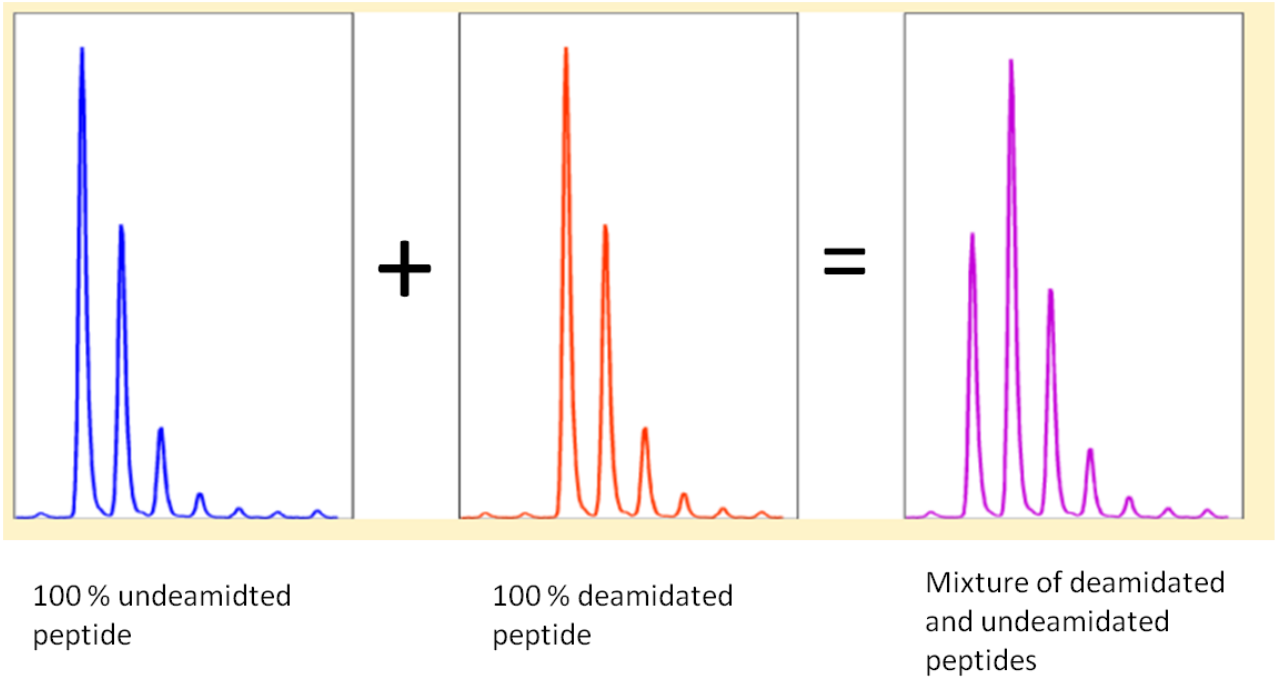


Figure 15: Isotopic distributions of undeamidated and deamidated versions of the same peptide (blue and red) plus an example of a mixture of the two peptide forms adapted from Wilson *et al.* 2012

For a peptide containing just one glutamine residue, a value between zero and one (referred to as the α-value) denotes the proportion of glutamine that is deamidated, and is determined by optimizing the fit of overlapping theoretical distributions with the experimental distributions. An α-value of 1 indicates no deamidation, while a value of 0 results from complete deamidation. The method can be extended to peptides with more than one glutamine residue. Each sample was analysed in triplicate by MALDI-MS and the α-value obtained from a weighted average of the three spectra, where the weights reflect the signal to noise ratio (S/N) of each peptide. Full details are given in Wilson *et al.* (2012). The code used to calculate deamidation levels is available as an R package from GitHub (https://github.com/franticspider/q2e.git).

Each peptide has an isotopic distribution which is dependent on the average terrestrial abundances of the stable isotopes. The theoretical isotopic distributions of peptides can be calculated using

their elemental composition. In order to calculate these distributions precisely, the product of polynomials, each representing the naturally occurring isotopes for an element in the peptide can be used (Yergey, 1983). When assessing the extent of deamidation, *Q2E* only considers the terms related to the first few peaks of the distribution, which simplifies the calculation. With the theoretical distribution calculated for the undeamidated peptide, the problem is simply to find the proportion (the alpha value) of this distribution and of the overlapping distribution shifted by 1 Da (at the resolution of the MALDI data) that gives the optimal fit to the observed experimental data. Various optimisation algorithms could be employed for this purpose, but most would provide a single optimal value for alpha. The genetic algorithm used is fast and easy to implement. Improved solutions are obtained from a population of random starting solutions, in a process that mimics evolution, until convergence is reached. In theory, different "optimal" solutions could be obtained so the method can also provide a measure of stability from multiple runs. The mean squared error between the calculated and experimental distributions is used as the fitness criterion and the program rejects those isotope distributions for which the final fitness measure is too low. In addition, the user can select the minimum level of signal to noise level, providing another criterion for rejection of poor spectra. In this case the minimum signal to noise (SNR) level was set to 3.0. To calculate a single alpha value the sample is spotted onto the MALDI plate three times, and a weighted average of these triplicate spectra are used to calculate the alpha value (the triplicate spectra are weighted according to the SNR level of each replicate). An example of the theoretical distributions of three of the peptides measured in this thesis (*m/z* 1105, 2056 and 2705, are shown below in Figures 16, 17 and 18, as well as the distributions observed in three real samples (Figures 19-23).

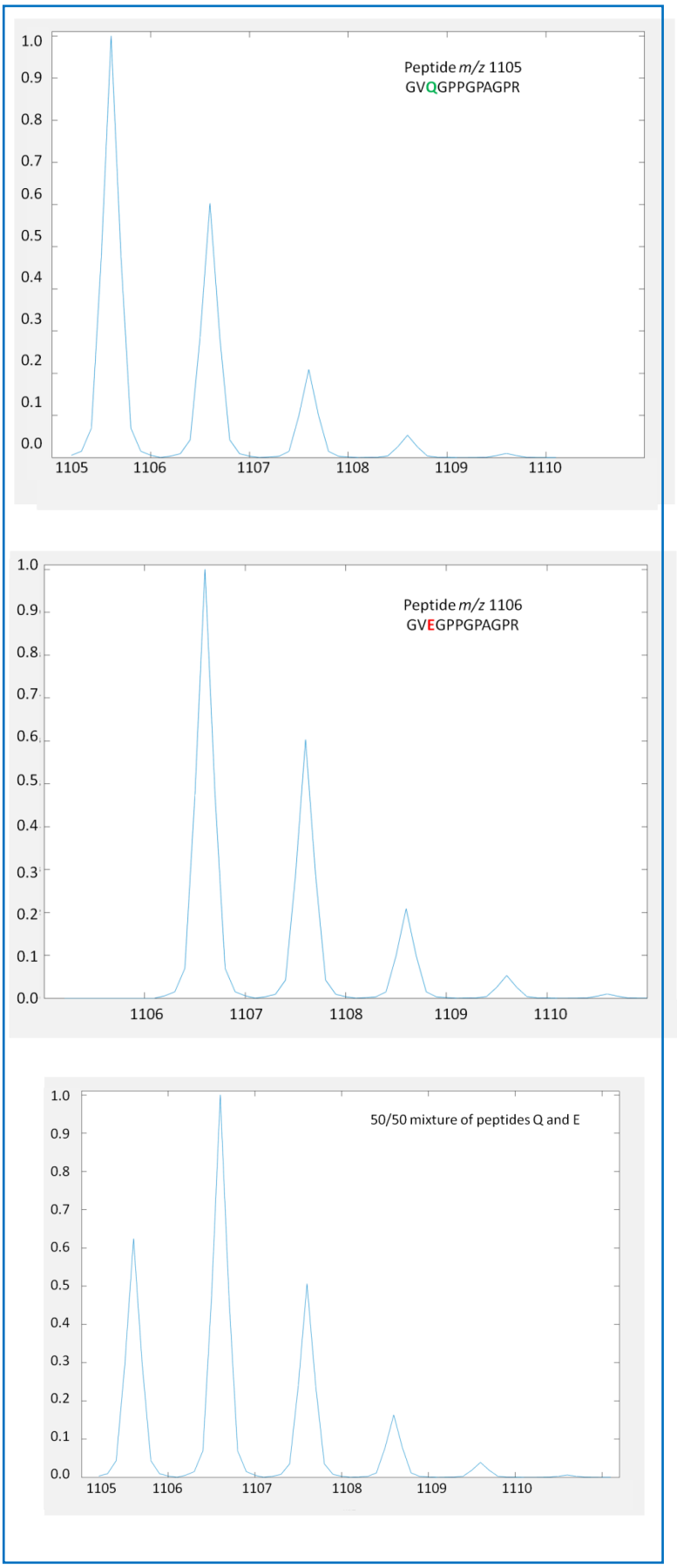


Figure 16: An example of the theoretical distributions of peptide *m/z* 1105 in various states of deamidation, (100%, 0% and 50%)

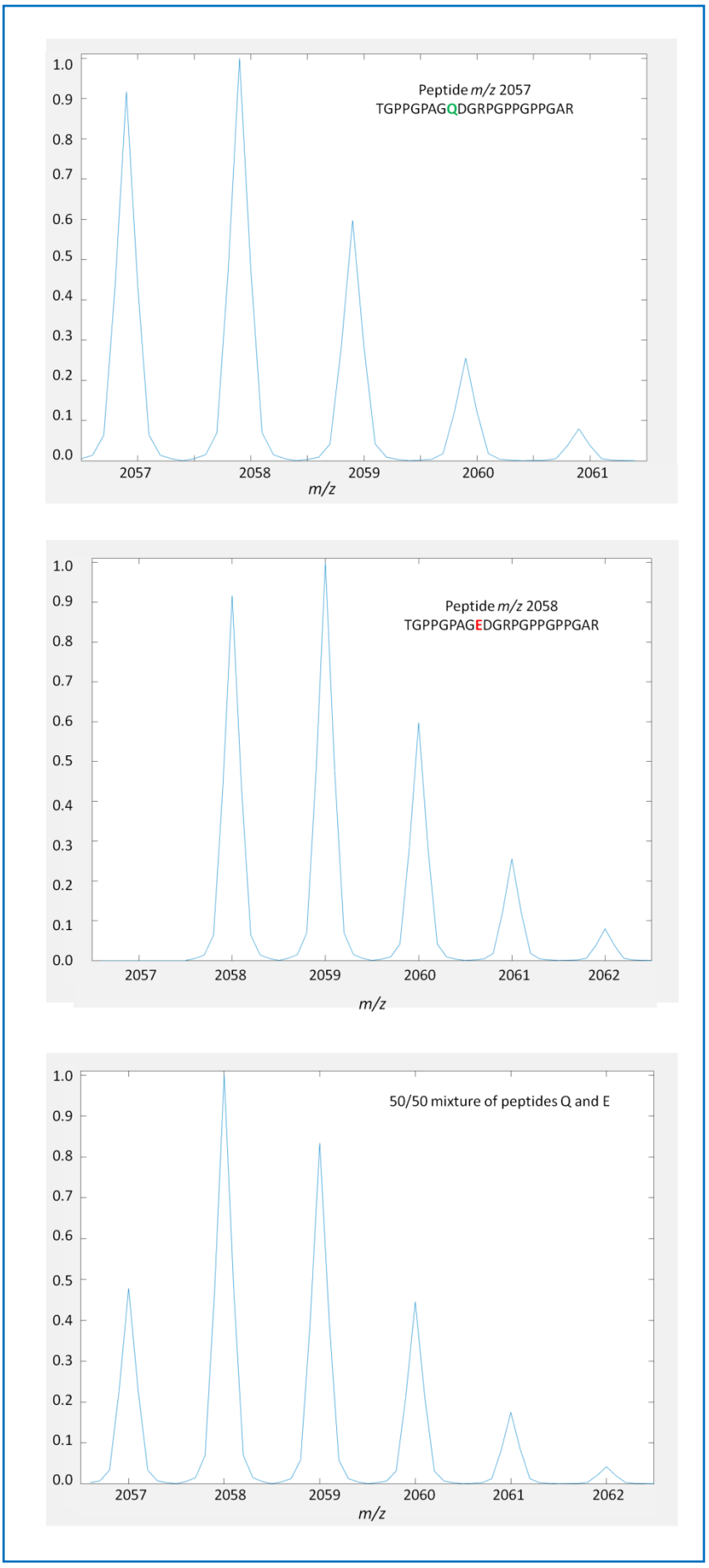


Figure 17: An example of the theoretical distributions of peptide *m/z* 2057 in various states of deamidation, (100%, 0% and 50%)

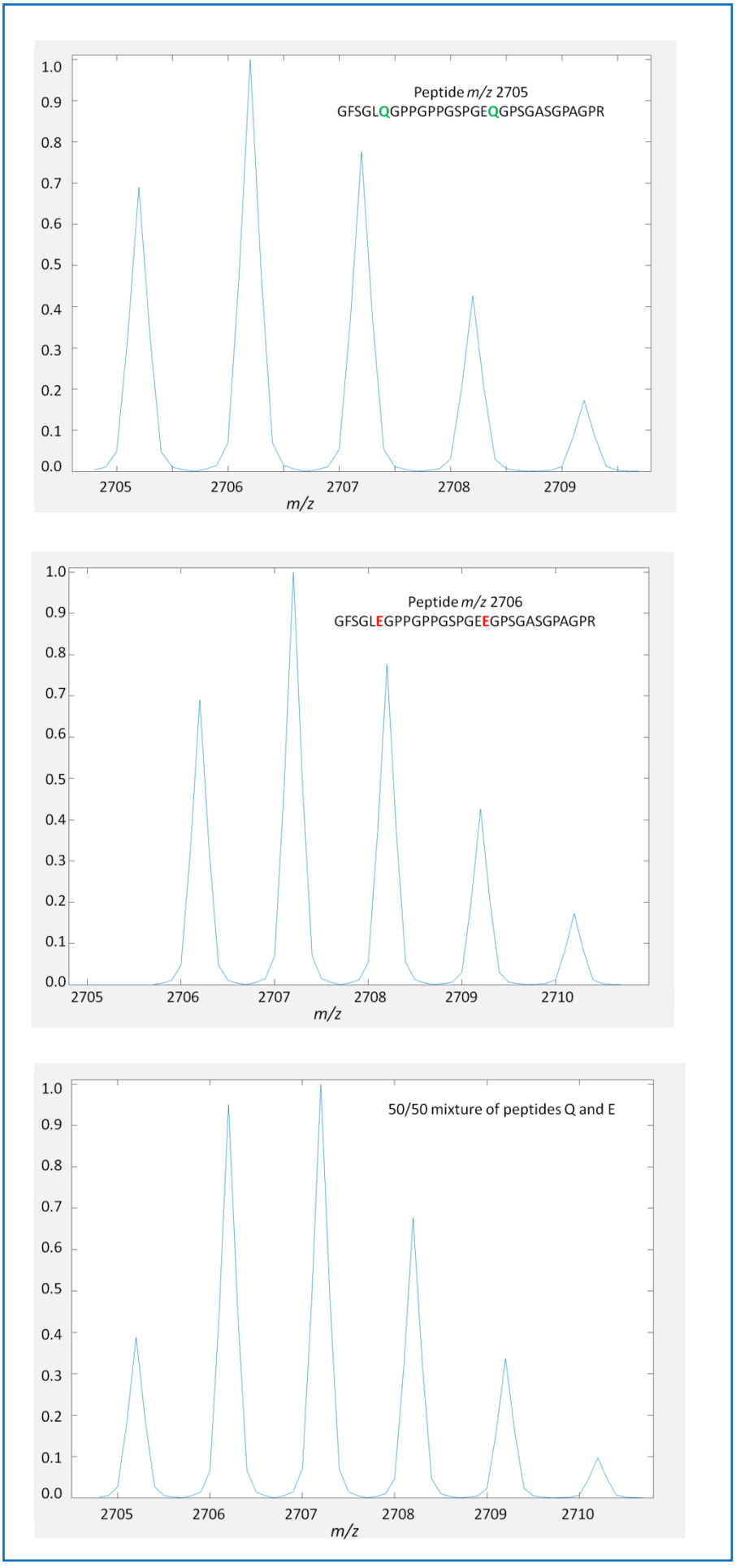


Figure 18: An example of the theoretical distributions of peptide *m/z* 2706 in various states of deamidation, (100%, 0% and 50%)

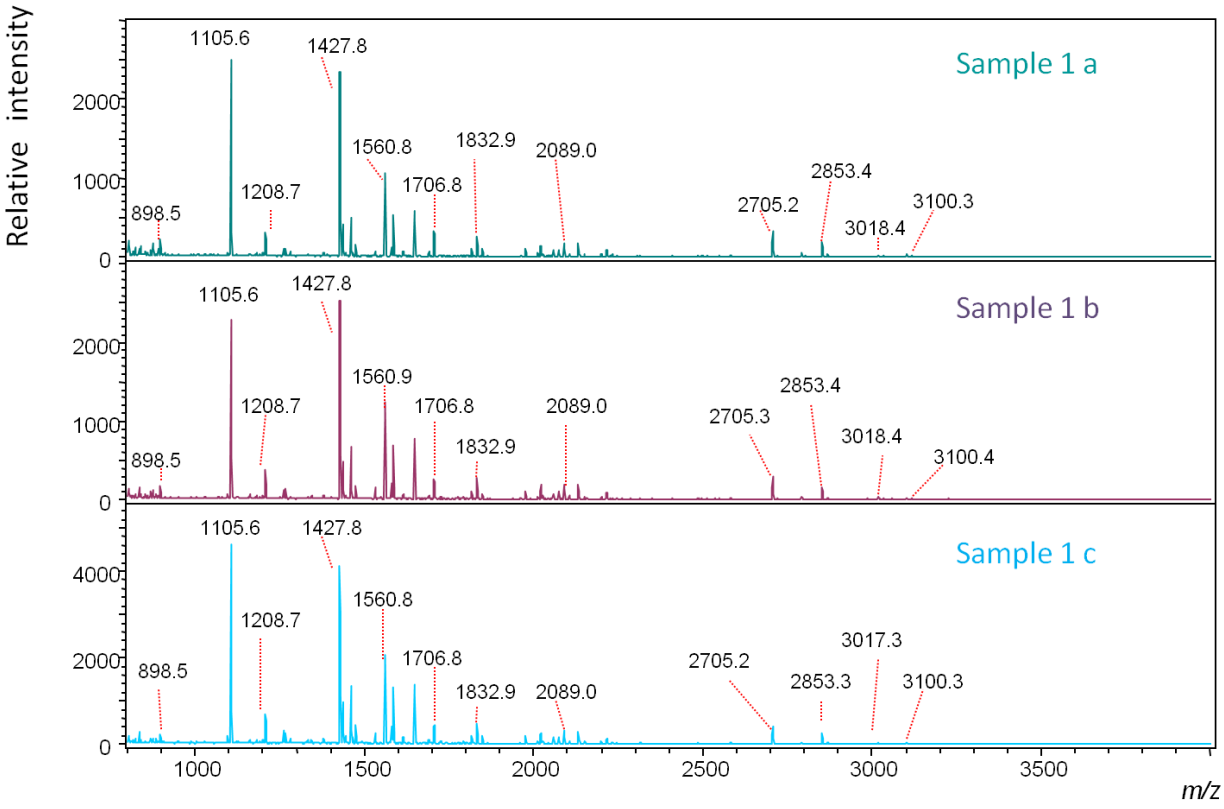


Figure 19 Three examples of spectra obtained from the analysis of three triplicate protein extracts from a piece of Roman *bovid* bone (samples 1a, 1b and 12c)

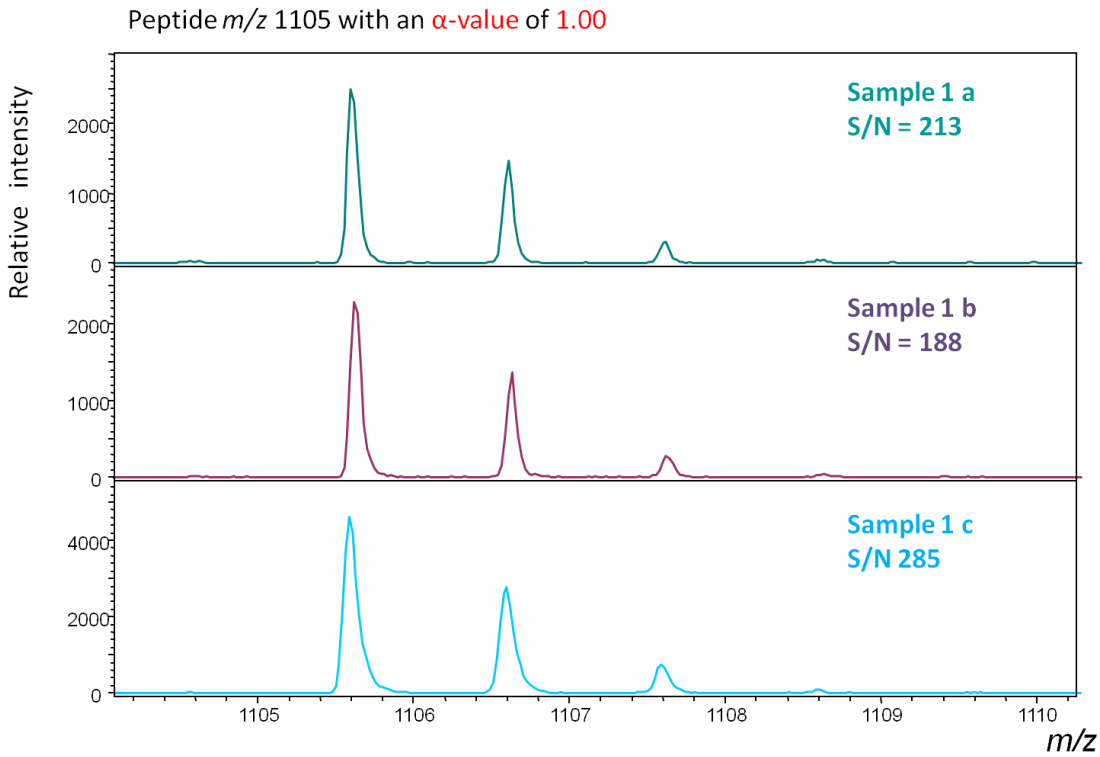


Figure 20: A close up of the isotopic distributions observed for peptide *m/z* 1105, with a calculated α-value of 1.00, indicating no deamidation has occurred.

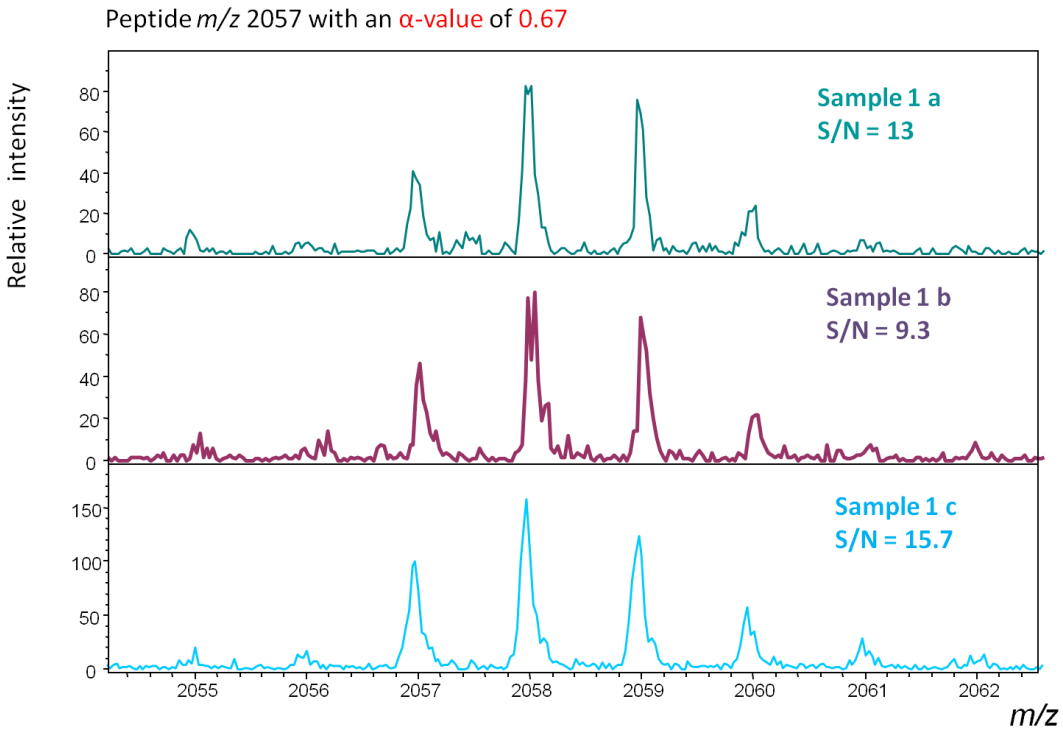


Figure 21: A close up of the isotopic distributions observed for peptide *m/z* 2057, with a calculated α-value of 0.67, indicating some deamidation has occurred.

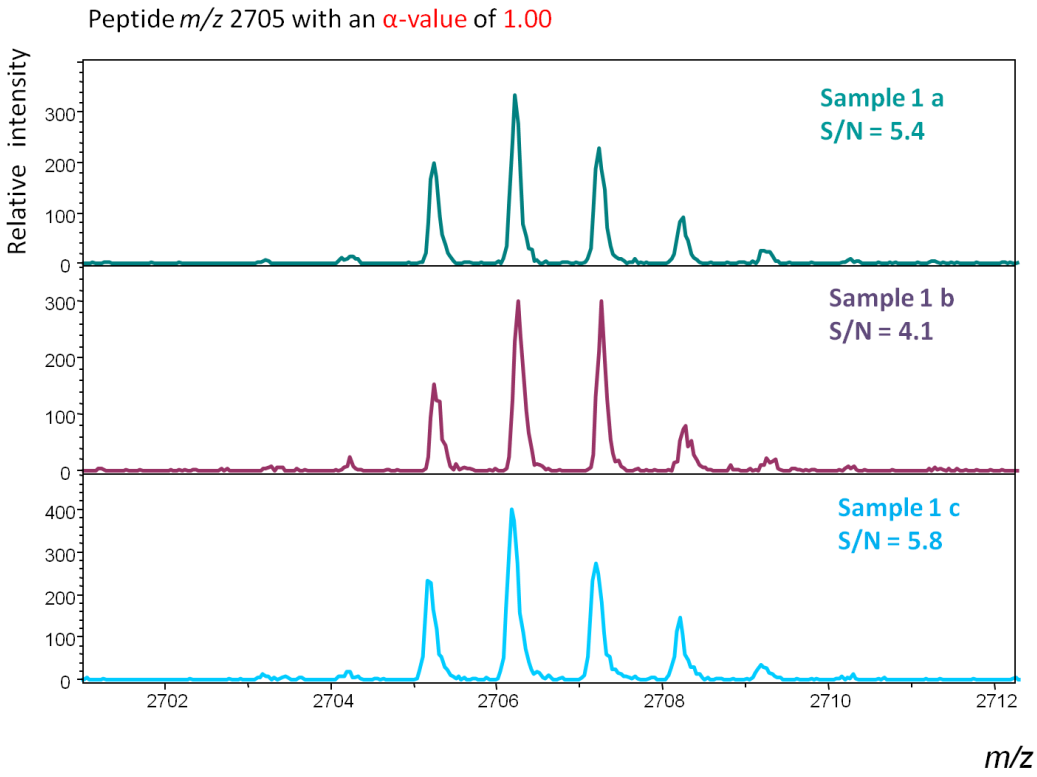


Figure 22: A close up of the isotopic distributions observed for peptide *m/z* 2705, with a calculated α-value of 1.00, indicating no deamidation has occurred.

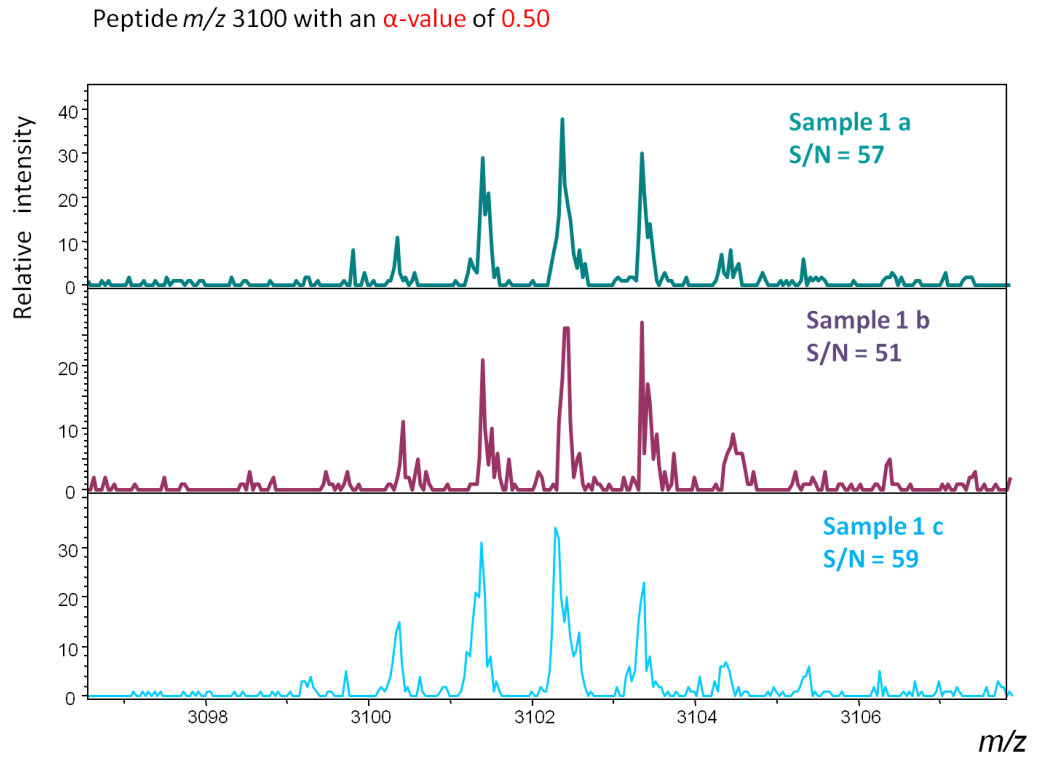


Figure 23: A close up of the isotopic distributions observed for peptide *m/z* 3100, with a calculated α-value of 0.50, indicating that around half of the peptide has demidated.

Chapter 2 - Methods

Introduction

The work presented in this thesis is heavily method development based. The key methodologies are described here, with further detail in chapters, as and when these methods are deviated from. The following methods have been used throughout this thesis to analyse protein extracted from bone, following the work of van Doorn *et al.* (2012) and Koon *et al.* (2012).

2.1 Sample preparation protocol for the extraction and preparation of protein extracted from bone to be analysed by mass spectrometry.

The sample preparation protocol for analysis of collagen tryptic peptides by mass spectrometry can be split into five main sections (Figure 13): 1) bone preparation and cleaning of the bone, 2) extraction of soluble proteins, 3) digestion of proteins into peptides, 4) purification of peptide mixtures, and 5) mass measurement of the peptides by mass spectrometry.

2.1.1 Sample preparation and cleaning

Step one, Figure 13: bones were cleaned at room temperature (~22 ºC) by soaking in 50 mM ammonium bicarbonate solution (pH 8.0) to minimise remove contaminants associated with the burial environment such as humic acids. Large bones were sliced using a diamond-edged water-cooled band saw in order to minimise heating of the sample, as deamidation may be induced thermally (van Doorn *et al.*, 2012). Slices were then either powdered or chipped. As well as bone chips, bone powder was also used in some experiments. Bone chips were powdered under liquid nitrogen using a SPEX freezer mill. Archaeological or sub fossil samples - which were already powdered or in chip form were cleaned at room temperature (~22 ºC) by soaking in 50 mM ammonium bicarbonate solution (pH 8.0) for one hour. The ammonium bicarbonate solution was then removed and discarded before performing the protein extraction.

2.1.2 Extraction of proteins using ammonium bicarbonate

Step two, Figure 24: 50 mM ammonium bicarbonate solution (pH 8) was added to each bone sample (approximately 100 µL per 30 mg of bone). The sample was then warmed for one hour at 65 ºC (adaptation of extraction procedures described in van Doorn *et al.* (2012).

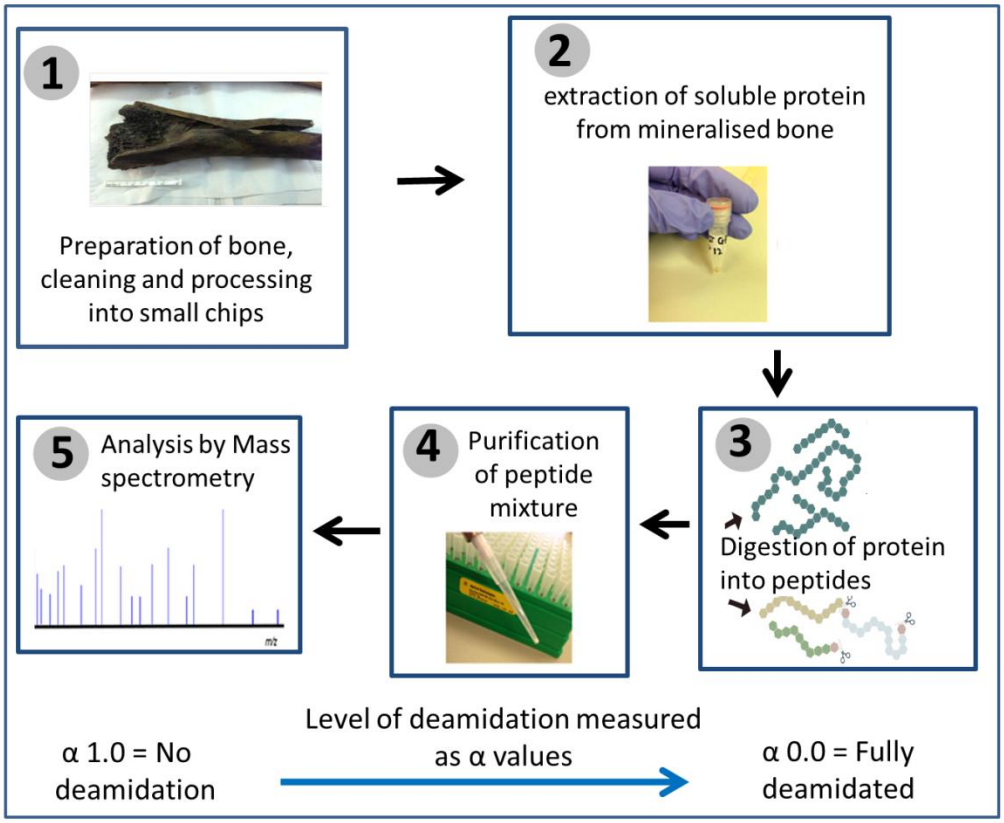


Figure 24: A schematic of sample preparation protocols. (1) Samples are cleaned in 50 mM ammonium bicarbonate at room temperature overnight. The sample is then cut into small pieces as required; (2) Soluble protein is extracted directly from the mineralised bone by warming in ammonium bicarbonate solution (at 65 °C) for one hour; (3) A tryptic digestion of the extracted protein is carried out overnight in ammonium bicarbonate solution at 37 °C ; (4) The resulting peptide mixture is purified using solid phase ZipTips; (5) the peptide mixture is analysed by MALDI-MS (section 2.5); the spectrum is used to estimate the level of deamidation occurring in specific peptides (section 2.6). The calculated glutamine deamidation level is given by the α-value, with a value of 1.0 representing no deamidation and 0.0 indicating complete deamidation of glutamine to glutamic acid

2.1.3 Digestion and purification of peptide mixtures

Steps three and four Figure 24; each of the 50 µL protein extracts was digested with 1 µL of porcine trypsin (0.4 µg/µL) overnight at 37 °C. The digests were purified using 100 µL C18 solid-phase tips (Millipore ZipTips) and desalted by washing with 0.1% trifluoroacetic acid (TFA) solution. Peptide mixtures were eluted in 50 µL of 50:50 (v/v) acetonitrile: water (water purified 18.0 MΩ) containing 0.1% TFA.

2.1.4 Analysis of peptide mixtures using mass spectrometry

Step five, Figure 24: the proteolytic peptide mixtures, consisting predominantly of tryptic peptides, were analysed using matrix-assisted laser/desorption ionisation with either the ultraflex (MALDI-

TOF-MS) or the solariX (MALDI or ESI FT-ICR-MS) instruments. For MALDI sample preparation a volume of 1 µL of sample solution was spotted on a ground steel MALDI target plate, followed by 1 µL of α-cyano-4-hydroxycinnamic acid matrix solution (1% in 50% ACN/0.1% TFA (w/v/v)). The sample and the matrix solutions were mixed together on the plate and allowed to air-dry. Each sample was spotted on to the MALDI target plate in triplicate. For analysis using the ultraflex the data was acquired using FlexControl softwere version 3.3 (Bruker Daltonics). Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in reflectron mode, equipped with a Nd:YAG smart beam laser. Spectra were acquired over a range of 800-5000 *m/z*. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg^1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu^1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). .). Monoisotopic masses were obtained using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a S/N threshold of 2. . Fragmentation was performed in LIFT mode without the introduction of a collision gas. The default calibration was used for MS/MS spectra, which were baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 m/z, cycles 4); monoisotopic peak detection used a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N of 6. Bruker flexAnalysis software (version 3.3) was used to perform spectral processing and peak list generation.

Majority of the samples were analysed using an automated run with calibrants spots analysed after every 8 acquisitions. Before starting the run the laser intensity was first optimised so that maximum signal intensity was obtained, without compromising the baseline of the spectrum. The average laser intensity used was around 43%, however this changed depending on the sample type, with extracts of low concentration typically requiring a higher laser intensity than concentrated extracts. The optimum number of laser shots was investigated at the beginning of each sample run with an average of 800 shots used. It was found that when analysing majority of the samples, the spectra obtained did not increase above 800 laser shots. The resolution was set to 2. In addition, before each run, the plate was aligned so that the laser started in the centre of each of each spot with the spot movement for each sample set to random.

When using the Bruker solariX instrument the same plate was used. The samples were analysed in positive mode with a smartbeam™ nitrogen laser (337 nm). Spectra were acquired using the solariXcontrol software and processed with DataAnalysis version 4.2 (Bruker Daltonics). The analysis carried out on the solariX was not automated and the instrument required only one calibration (using the same peptide mixture as the ultraflex) at the beginning of the analysis. Each spot was analysed manually with 800 laser shots used and an average of 8 spectra acquired for each analysis. Unlike the ultraflex the laser was moved manually during acquisition as an automated method was not available.

2.2 Determining the level of deamidation in a peptide in data obtained using the ultraflex (TOF instrument).

The deamidation of glutamine results in an overall mass increase of 0.984 Da. One disadvantage of the TOF instrumentation used in this work is that due to the insufficient resolving power of the mass analyser, it was not possible to resolve the deamidated and undeamidated signals: the nth peak of the deamidated peptide signal (typically the mono-isotopic signal) overlaps the (n+1)th peak of the undeamidated form (typically the signal for the species containing one ^{13}C atom). The extent of deamidation of glutamine (Q), converting it to glutamic acid (E) can be estimated by deconvolution of the two overlapping distributions as described in Wilson *et al.* (2012). For a peptide containing just one glutamine residue, a value between zero and one (referred to as the α-value) denotes the proportion of glutamine that is deamidated, and is determined by optimizing the fit of overlapping theoretical distributions with the experimental distributions. An α-value of 1 indicates no deamidation, while a value of 0 results from complete deamidation. The method can be extended to peptides with more than one glutamine residue. Each sample was analysed in triplicate by MALDI-MS and the α-value obtained from a weighted average of the three spectra, where the weights reflect the signal to noise ratio (S/N) of each peptide. Full details are given in Wilson *et al.,* (2012). The code used to calculate deamidation levels is available as an R package from GitHub (https://github.com/franticspider/q2e.git)

Chapter 3 - Optimising the extraction of glutamine containing collagen peptides from bone

Introduction

Protein can be extracted from bone in order to answer a number of questions relating to the past. Advances in genetics and analytical chemistry have helped to gain a better understanding of the structure of collagen and its function as a protein; as well as building libraries, containing the collagen amino acid sequences of hundreds of animals.

Although the amino acid sequences which make up all collagens are similar, there are localised variations between species. Previous studies (Buckley *et al.,* 2008; Buckley *et al.,* 2009; ; Buckley and Collins 2011; Kirby *et al.*, 2013) have looked at these variations by comparing different tryptic peptide mass fingerprints. The observed peptide mass fingerprints can be used for species identification (ZooMS) if the amino acid sequence of collagen from the particular species is known, or if an authentic sample of collagen from that species is available for comparison. van Doorn *et al.* (2012) developed a minimally destructive extraction procedure, whereby collagen is extracted by warming the bone in ammonium bicarbonate solution. The solution is then subjected to an overnight trypsin digestion. The subsequent collagen peptide mixture is purified on C18 ZipTips, and analysed using a Time of Flight Mass Spectrometer (TOF-MS) to generate a peptide fingerprint. This minimally destructive ammonium bicarbonate extraction method has been adapted by van Doorn *et al*. (2012) to extract and analyse levels of glutamine deamidation in collagen peptides. Hurtado and O'Connor (2012) have used a similar ammonium bicarbonate extraction to measure deamidation of asparagine (Asn) to aspartic acid (Asp) in collagen using a Fourier Transform Ion Resonance Mass Spectrometer (FT-ICR-MS). Originally the protocol by Buckley *et al.* (2009) was designed to be purely qualitative, with the interpretation of the data looking only at the absence or presence of specific peptides. The protocol was designed to extract and recover as many tryptic peptides as possible with the expectation that some of them would be species specific. In order to adapt this method to measure the conversion of glutamine (Gln) to glutamic acid (Glu) accurately, the method must be semi-quantitative to measure the relative ratios of the Gln and Glu-containing peptides without altering them. Each step within the method should be investigated to ensure that minimal deamidation is caused during the sample preparation, and that there is no bias in the extraction of either of the two peptide forms, as this would alter the final measurement. The methods employed also aim to maximise the recovery of Gln-containing peptides. Lastly in order to measure deamidation using MS, both the Gln and Glu-containing peptides need to give the same mass spectrometric response. This is investigated in Chapter 4. The methods used by Buckley *et al.* (2009) and van Doorn *et al.* (2012) can be divided into four main steps (Figure 25).

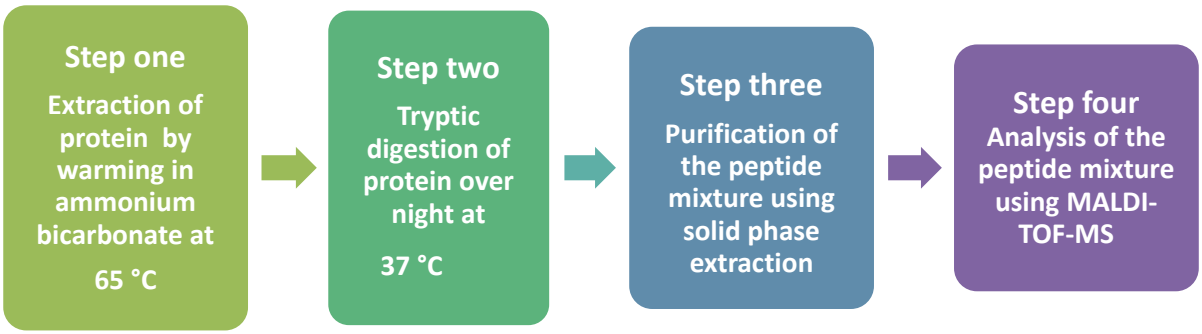


Figure 25: Flow diagram of steps 1-4 used extract and analyse glutamine-containing peptides in collagen from bone. The whole process from extraction to obtaining spectra takes approximately three days

In order to assess the suitability and optimise the method for the measurement of glutamine deamidation, the following were investigated as part of this thesis:

(1) The effect of different temperatures (37 °C, 65 °C, 110 °C and 140 °C) for the ammonium bicarbonate extraction of glutamine deamidation
(2) The effect of trypsin digestion time on deamidation
(3) Comparison of spectra obtained using either standard MALDI, or anchor chip MALDI plates
(4) Reproducibility of α-values.

3.1 Investigating the relationship between temperature and deamidation in collagen peptides.

The effect of temperature and time on both glutamine and asparagine deamidation have been investigated using both natural and synthetic polypeptides (Scotchler and Robinson 1974; Stratton *et al.,* 2001; Hurtado and O'Connor 2012; Robinson 2004), with studies of glutamine in pentapeptides showing a decrease in half-life with increased temperature (Scotchler and Robinson 1974). Analysis of glutamine deamidation has been carried out on a wide range of sample types, such as wool (Araki and Moini 2011) and protein-based binders found in paints (Leo *et al.,* 2011). Glutamine deamidation has also been investigated in the crystalline protein within the eye lens (Flaugh *et al.* 2006). One common conclusion from these studies is that deamidation appears to be temperature dependent; however at present kinetic data on the rate of this reaction is not consistent and it is difficult to compare the data obtained from different peptide sequences using different kinetic methods. Most of the kinetic studies of Gln and Asn deamidation have been carried out on short synthetic peptides, and therefore may not be representative of mechanisms of deamidation within larger, more complex structures such as proteins. Both glutamine and

asparagine have two alternate deamidation pathways, either via a cyclic intermediate, or via direct side chain hydrolysis, due to the lack of flexibility necessary for the protein backbone to adopt the appropriate interatomic distance needed for the formation of the cyclic intermediates (Van Duin and Collins 1998; van Doorn *et al.,* 2012) . Therefore, in proteins such as collagen it is easier to envisage that deamidation occurs via direct hydrolysis. If this mechanism path is correct it would mean that, in collagen the two residues are likely to be equally stable; however once in solution, gelatine - the soluble form of collagen - no longer has the same structural constraints, with the protein forming random coils. In these conditions at neutral pH asparagine has found to deamidate rapidly, and therefore differences in levels of deamidation between Gln and Asn are likely to be preparation induced.

In this study, 12 peptides (Table 2) were measured in protein extracted from a bovine bone excavated from Tanner Row, York. The bone was found in an un-stratified context and therefore the exact age is not known; from archaeological interpretation of the site however it is thought to be most likely from the Roman period. The bone appeared to be well preserved with a waxy feel to the outer surface.

Temperatures of 37 °C, 65 °C, 110 °C and 140 °C were used to investigate the temperature dependence of Gln-deamidation. The first two temperatures are used in different steps of the sample preparation protocol described in Figure 24. Although 37 °C is relatively low in terms of kinetic studies, this temperature is used during the tryptic digestion step, during which samples are incubated at 37 °C, for up to 16 hours. Higher temperatures (80 °C, 110 °C and 140 °C) have been routinely used to study the kinetic reactions of amino acid racemisation in protein (Penkman *et al.,* 2008; Crisp e*t al.,* 2013; Demarchi *et al.* 2013). As Gln deamidation is thought to be relatively slow, it was hypothesised that exposure to the temperatures of 37 °C and 65 °C would cause no or little deamidation.

Table 3: 12 peptides that are observed in MALDI mass spectra, of tryptic digests, of bovine type I collagen and contain at least one glutamine residue. Where possible the theoretical amino acid sequence of the peptides has been demonstrated by product ion analysis. For peptides where this was not possible, due to poor spectral quality, sequences were taken from published data (Wilson *et al.,* 2012) and assigned on the basis of the peptides' accurate *m/z* values

M + [H]	peptide sequence	Number of proline hydroxylations	Collagen chain	position in collagen chain
836.44	GPAGPQGPR*	none	COLL 1A1	[1084-1092]
1105.57	GVQGPPGPAGPR*	1 Oxidation	COLL 1A1	[685-696]
1690.77	DGEAGAQGPPGPAGPAGER	none	COLL 1A1	[612-630]
1706.77	DGEAGAQGPPGPAGPAGER	1 Oxidation	COLL 1A1	[612-630]
2056.98	TGPPGPAGQDGRPGPPGPPGAR*	4 x Oxidation	COLL 1A2	[552-573]
2073.01	GAPGADGPAGAPGTPGPQGIAGQR	1 x Oxidation	COLL 1A1	[934-957]
2089.01	GAPGADGPAGAPGTPGPQGIAGQR	2 x Oxidation	COLL 1A1	[934-957]
2689.26	GFSGLQGPPGPPGPSGEQGPSGASGPAGPR	2 x Oxidation	COLL 1A1	[1111-1140]
2705.26	GFSGLQGPPGPPGSPGEQGPSGASGPAGPR*	3 x Oxidation	COLL 1A1	[1111-1140]
3001.50	GPSGEPGTAGPPGTPGPQGLLGAPGFLGLPGSR	3 x Oxidation	COLL 1A2	[845-877]
3100.41	GLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPR*	6 x Oxidation	COLL 1A1	[187-219]
3665.54	GSQGSQGPAGPPGPPGPPGPPGPSGGGYEFGFDGDFYR*	4 x Oxidation	COLL 1A2	[1079-1116]

* Assignment of sequence demonstrated using product ion spectrum.

3.1.1 Methods

The bone was sliced using a diamond edged water cooled band saw in order to minimise introduction of heat, powdered using a freezer mill (SPEX) under liquid nitrogen. The resulting powder was left to air-dry for three days.

Triplicate samples for each time point and temperature were prepared, resulting in a total of 96 samples (eight time points, at four temperatures x three replicates). For each sample ~10 mg aliquots of bone powder were weighed. Samples to be extracted at 37 °C and 65 °C were prepared by first transferring the sample to individual plastic polypropylene (PP) microfuge tubes; those heated at 110 °C and 140 °C were transferred to 2 mL glass vials with plastic screw tops. 100 µL of 50 mM ammonium bicarbonate solution (pH 8.0) was added to each sample. Samples heated at 37 °C, 65 °C and 110 °C were heated for 0.5, 1, 2, 4, 6, 7, 8 and 24 hours. Samples heated at 140 °C were heated for 0.25, 0.45, 0.50, 1, 1.5, 2, 2.5 and 3 hours. After heating, 50 µL of supernatant from each sample was transferred to a new polypropylene (PP) microfuge tube. The samples were digested overnight, the resulting peptide mixtures consisting predominantly of tryptic peptides were analysed using MALDI-TOF-MS as described in Chapter 1.

3.1.2 Results

As no reference standards (material with known levels of deamidation) were available, biological material (bovine bone) was used. This means that values of deamidation present in the sample prior to heat treatment are unknown; it is therefore not possible to quantify the absolute levels of deamidation produced from heating. However the levels of deamidation observed at the four temperatures can be compared. The peptide with an *m/z* value of 1105.57 (GVQGPPGPAGPR) was chosen to compare deamidation at different temperatures. This peptide has been described as deamidating slowly, with reported half-life of 640 days (van Doorn *et al.,* 2012). In this experiment it was the most frequently observed out of the 12 Gln-containing peptides that were measured (Table 2). α-values for *m/z* 1105.57 were calculated for each sample as described in Chapter 2 (section 2.2). α-values measured in *m/z* 1105.57 were compared across the four temperatures, with α-values plotted against time. Samples heated at 37 °C showed little or no deamidation in the peptide at *m/z* 1105.57 over 24 hours, with only one sample out of the eight time points showing deamidation (with an α-value of 0.99) (Figure 26, top left). Samples heated at 65 °C over 24 hours show lower α-values than those heated at 37 °C (Figure 26, top right). Heating for 30 minutes at 65 °C yielded a mean α-value of 0.96, and after 24 hours the average α-value was 0.94. α-values calculated for peptides heated at this temperature ranged from 0.92 to 1.00.

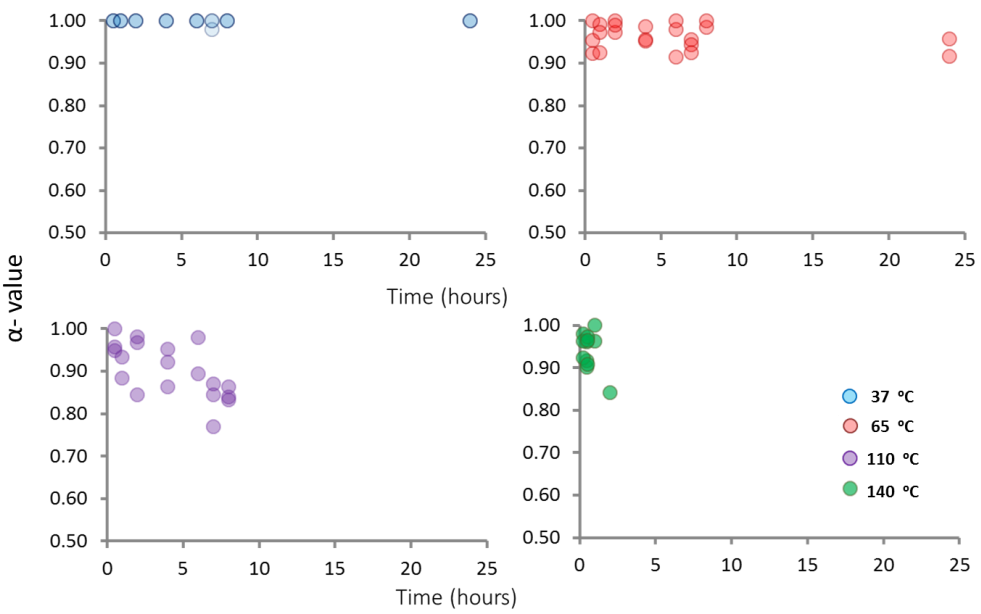


Figure 26: Levels of deamidation observed in protein extracted from powdered bovine metatarsal bone. Deamidation is reported as α-values measured in peptide with *m/z* 1105.57 (GVQGPPGPAGPR). The bone was heated in 100 µL of 50 mM ammonium bicarbonate solution (pH 8 over 24 hours, at either 37 °C (top left), 65 °C (top right), or 110 °C (bottom left). Samples were also heated at 140 °C over three hours

The average α-value at 30 minutes for samples heated at 110 °C was 0.97, after eight hours this decreased to 0.83. Peptide *m/z* 1105.57 was not observed after heating at 110 °C for 24 hours. The range of α-values obtained at this temperature was 0.77 -1.00. In samples heated at 140 °C, peptide *m/z* 1105.57 was not observed after 2 hours of heating. The α-values obtained at this temperature ranged from 0.84 to 1.00. For each temperature a total of 288 observations was possible (8 time points, x 3 replicates per time point, x 12 peptides measured in each sample). 65 °C yielded the highest percentage of observations (51.04 %). At 37 °C the percentage of observations was 40.63 %. This percentage decreased at 110 °C with 31.94 % and 16.67 % for samples heated at 140 °C (Figure 27). Overall spectra obtained from samples heated at 110 °C and 140 °C resulted in poor quality of spectra, with fewer peaks and lower S/N ratios than observed at the lower temperatures of 37 °C and 65 °C.

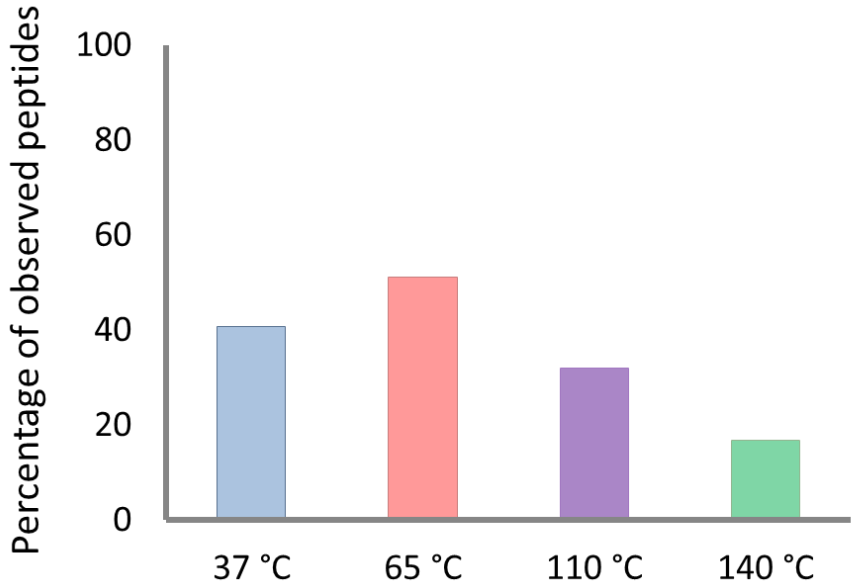


Figure 27: The percentage of the total possible observations in samples heated at four different temperatures. Samples heated at 37 °C, 65 °C and 110 °C were heated over 24 hours while samples at 140 °C were heated for up to three hours

3.1.3 Conclusions

Results from these experiments demonstrate that glutamine deamidation in peptide *m/z* 1105.57 increases with temperature. As well as an increase in deamidation, temperature was also found to decrease the number of peptides observed in spectra. Samples here were heated in solution which may promote hydrolysis of the peptide bonds (Hill, 1965). Cleavage of the peptide bond would result in smaller fragments than those measured in this experiment (Table 2).

In conclusion, it was found that in this sample, deamidation of glutamine in peptide *m/z* 1105.57 increased with temperature when heated in ammonium bicarbonate solution. At the lower temperatures used during sample preparation (37 °C and 65 °C) the level of deamidation observed was very low, even after heating for 24 hours, therefore it was unlikely that these conditions would induce significant levels of deamidation. However it was observed that the quality of spectra obtained from high temperature experiments (110 °C and 140 °C) was poor. At these higher temperatures fewer peaks were observed, and the S/N (signal to noise ratio) of peaks was lower than those observed in samples heated at 37 °C and 65 °C. It was therefore difficult to assess the levels of deamidation in Gln-containing peptides other than *m/z* 1105.57, particularly for time periods longer than 8 hours. Due to this, heating samples in solution at 110 °C and 140 °C would not be favourable for the extraction and measurement of collagen peptides. In order to measure this reaction over longer time periods, lower temperatures and conditions that minimise peptide hydrolysis would be required. One way of reducing the amount of peptide hydrolysis may be to heat the samples in sand rather than solution. Heating bone samples in sand would also be more representative of the burial environment from which bones are most commonly excavated. The use of sand as a medium to heat samples is investigated further in Chapter 5.

It may be difficult to model the way in which glutamine undergoes deamidation in the burial environment, if the rate observed in higher temperature experiments does not reflect just an acceleration of the pattern of deamidation observed at lower temperatures. This is particularly important when looking at archaeological samples, as material will have been buried at temperatures below 65 °C (Goodfrien*d et al*., 1991).

3.2 Trypsin digestion of peptides

Porcine trypsin is widely used in proteomic studies. The enzyme cleaves on the C-terminal side of arginine and lysine residues unless followed by proline (Olsen *et al.,* 2004). Arginine and lysine make up approximately 10 % of the residues in type one collagen resulting in the average length of tryptic peptides being ~ 10 amino acids long. Breaking down the extracted collagen into shorter peptides enables the analysis of specific sites within the protein. Previous studies have used mass spectrometry to measure deamidation in peptides containing asparagine. In these studies samples were labelled with ^{18}O in order to measure induced deamidation and were subjected to four hour (Li et *al.,* 2008), and overnight tryptic digestion (Araki and Moini 2011), with both digestion times showing no deamidation. As asparagine is known to be less stable than glutamine, it has been predicted that the overnight tryptic digestion will not induce deamidation in glutamine (Wilson *et al.,* 2012).

The effect of trypsin digestion (at 37 °C), on Gln deamidation, as well as the S/N of the Gln-containing peaks was investigated. Aliquots of the same protein extract (extracted at 65 °C for one hour), were subjected to digestion for differing lengths of time, incubated with trypsin for either 2, 4, 6, or 16 hours. Samples were prepared in duplicate for each time point, with each sample analysed in triplicate. An average of each triplicate was used to calculate an α-value.

3.2.1 Methods

~ 80 mg of bone powder was transferred to a 2 mL plastic polypropylene (PP) microfuge tube. The bone powered was prepared as described in Chapter 2. 800 µL of 50 mM ammonium bicarbonate pH 8.0 was added to the bone powder; this was vortexed for 60 seconds and heated at 65 °C for one hour. The supernatant was then transferred to a new 1 mL plastic polypropylene (PP) microfuge tube and eight aliquots of 50 µL were transferred to 0.5 mL plastic polypropylene (PP) microfuge tubes. To each sample 1 µL of porcine trypsin (0.4 µg/µL) was added. Samples were heated at 37 °C for 2, 4, 6 and 16 hours with two samples used at each time point. The digests were purified using 100 µL C18 solid-phase tips (Millipore ZipTips) and desalted by washing with 0.1% trifluoroacetic acid (TFA) solution. Peptide mixtures were eluted in 50 µL of 50:50 (v/v) acetonitrile:MilliQ water (0.1% TFA). The resulting peptide mixtures, consisting predominantly of tryptic peptides, were analysed using (MALDI-TOF-MS). α-values were measured in 12 Gln-containing peptides (Table 2).

3.2.2 Results

Out of the 12 Gln-containing peptides, four (*m/z* 2705.26, *m/z* 2089.01, *m/z* 2073.01 and *m/z* 2056.98) were found to show increased deamidation with the length of time they were incubated with trypsin (Figure 28). Peptide with *m/z* 2705.26 shows little deamidation in the first six hours of incubation, with α-values for the first 3 time points all having α-values of 1.00, at 16 hours however, an increase in deamidation was measured with an average α-value of 0.84. Peptides *m/z* 2089.01 and *m/z* 2073.01 show similar patterns of deamidation to each other, with a slight increase in deamidation over time. Peptides *m/z* 2073.01 is the hydroxylated version of *m/z* 2089.01, with the difference in mass of ~ 16 Da due to an additional oxygen. The reproducibility of α-values at the two hour time point is not consistent, and so it is difficult to determine the level of deamidation at this time point. The peptide which showed the strongest correlation of deamidation with incubation time was *m/z* 2056.98, again with the highest level of variability between replicate samples observed at two hours. The average α-values for peptide *m/z* 2056.98 range from 0.79 after two hours, to 0.43 at 16 hours.

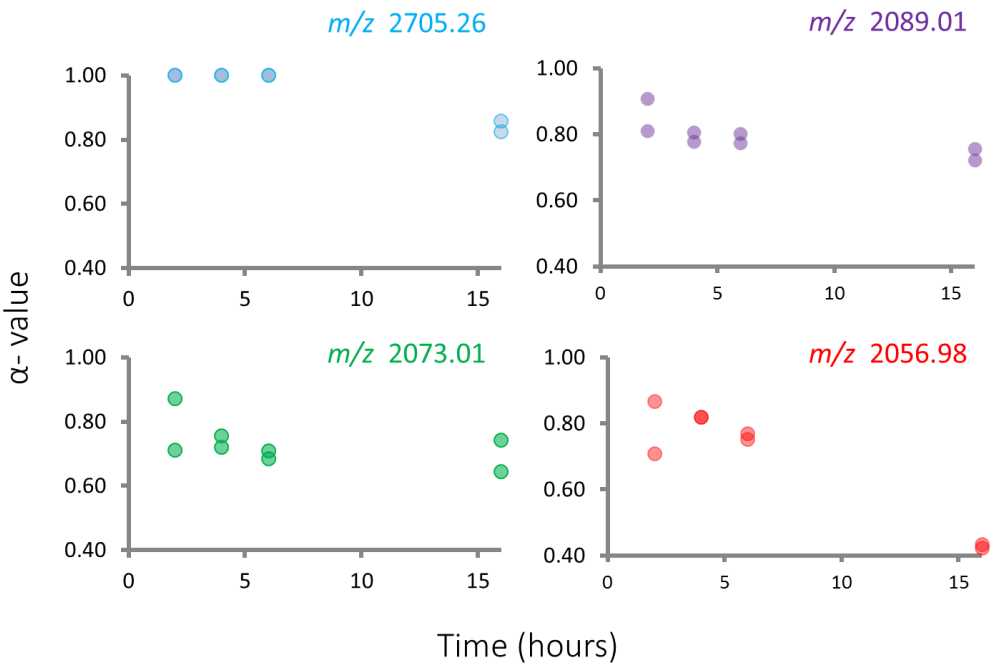


Figure 28: Protein extracts were digested for different lengths of time, in trypsin, at 37 °C. α-values are shown for peptides; *m/z* 2705.26, *m/z* 2089.01, *m/z* 2073.01 and *m/z* 2056.98

Four peptides showed no correlation between deamidation and the length of digestion time (*m/z* 2689.26, *m/z* 3001.50, *m/z* 3100.41 and *m/z* 3665.54). Peptides which did not show an increase with deamidation across the four digestions times were *m/z* 836.44, *m/z* 1105.57, *m/z* 1690.77 and *m/z* 1706.77 (Figure 29). As well as measuring the levels of deamidation in each peptide (Table 2), the S/N (signal to noise ratio) of the mono isotopic signal was also recorded for each of the 12 observed Gln-containing peptides. The results can be split into four groups:

Group one contains two peptides with S/N that did not increase after two hours of incubation with trypsin, *m/z* 1706.77 and *m/z* 1105.57 (Figure 30).

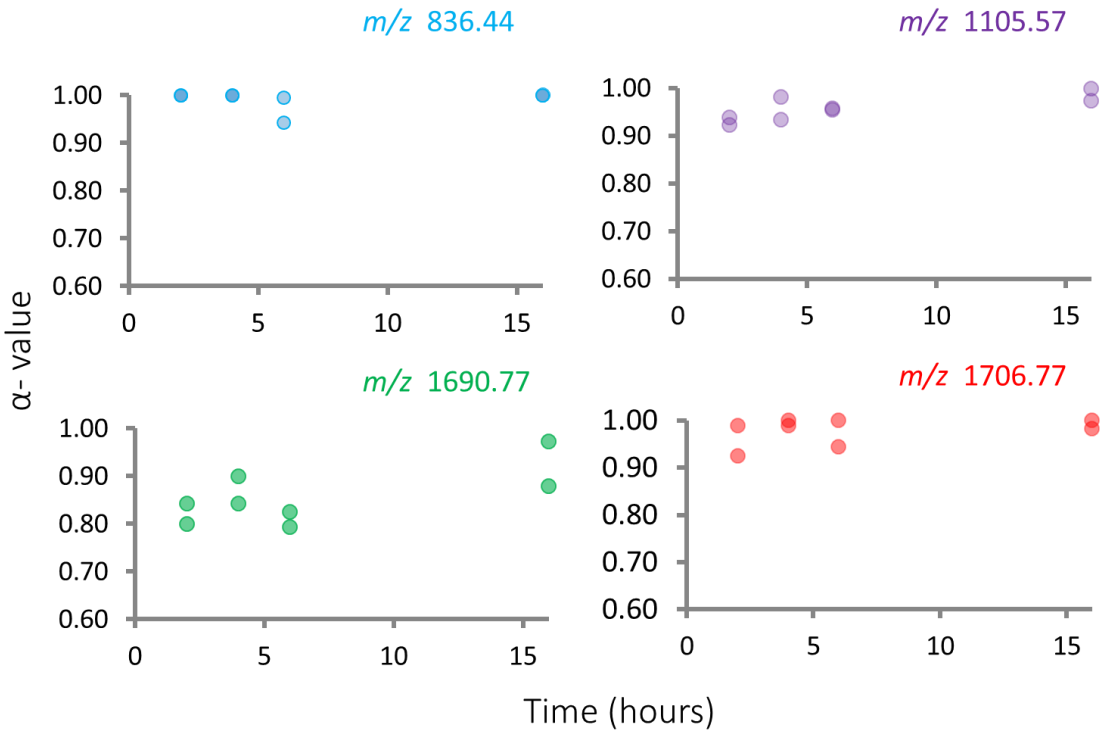


Figure 29: Protein extracts were digested for different lengths of time, in trypsin, at 37 °C. α-values are shown for peptides; *m/z* 836.44, *m/z* 1105.57, *m/z* 1690.77 and *m/z* 1706.77

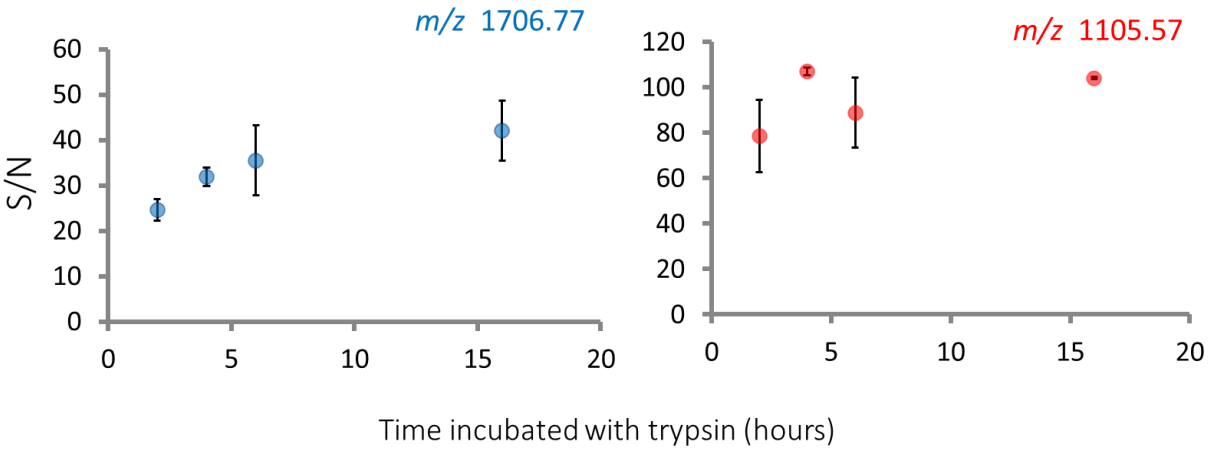


Figure 30: The S/N ratio of two peptides *m/z* 1706.77 and *m/z* 1105.57 in bovine protein extracts, after being incubated with trypsin for either: 2, 4, 6 or 16 hours. The S/N of these peptides did not increase significantly after a digestion time of two hours. The error bars show the range in S/N values obtained between the replicate samples.

Group two consists of three peptides with a S/N ratio that did not increase after four hours of incubation with trypsin, *m/z* 836.44, *m/z* 3001.50 and *m/z* 3100.41 (Figure 31). The difference in S/N observed for peptide *m/z* 3100.41 was particularly high with replicates of samples incubated at 16 hours giving S/N of 56 and 262.

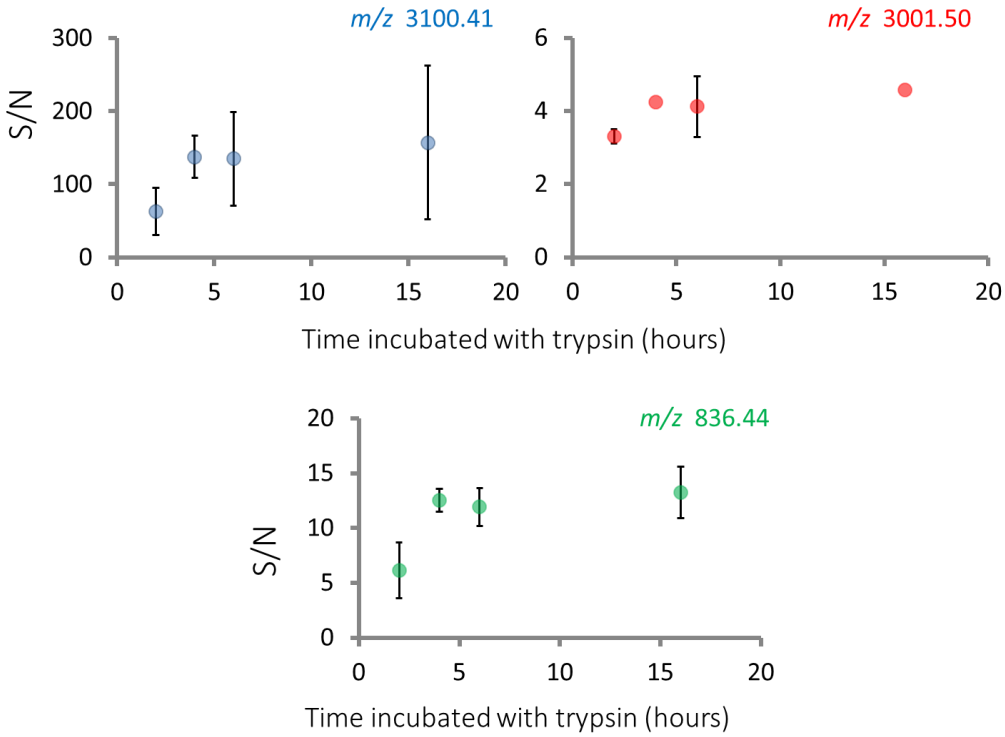


Figure 31: The S/N ratio of three peptides *m/z* 3100.41, *m/z* 3100.41 *and m/z* 836.44, in bovine protein extracts, after being incubated with trypsin for either: 2, 4, 6 or 16 hours. The S/N of these peptides did not increase significantly after a digestion time of four hours. The error bars show the range in S/N values obtained between the replicate samples.

Group three includes two peptides with a S/N that did not increase after six hours (*m/z* 1690.77 and *m/z* 2056.98)(Figure32).

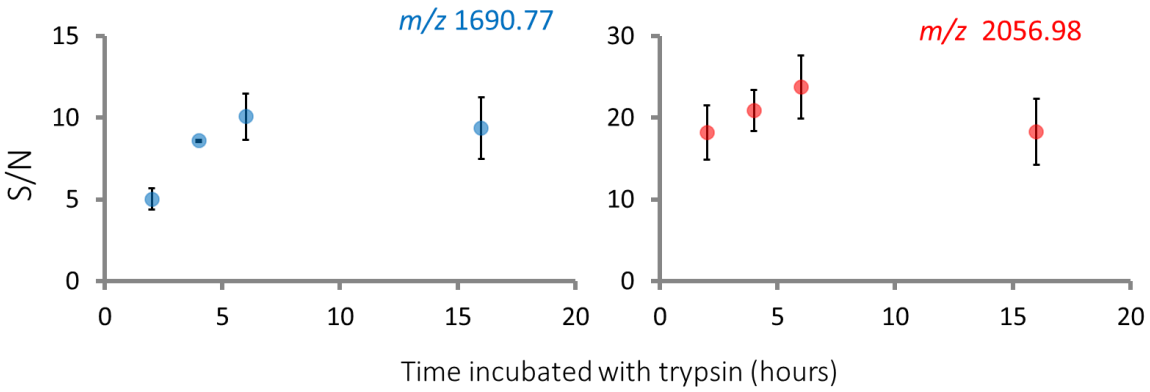


Figure 32: The S/N ratio of two peptides *m/z* 1690.77 and *m/z* 2056.98 in bovine protein extracts, after being incubated with trypsin for either: 2, 4, 6 or 16 hours. The S/N of these peptides did not increase significantly after a digestion time of six hours. The error bars show the range in S/N values obtained between the replicate samples.

Group four includes four peptides which showed an increase in S/N after 16 hours (*m/z* 2073.01, *m/z* 2089.01, *m/z* 2689.26 and *m/z* 2705.26), this is particularly evident in peptide *m/z* 2705.26 (Figure 33).

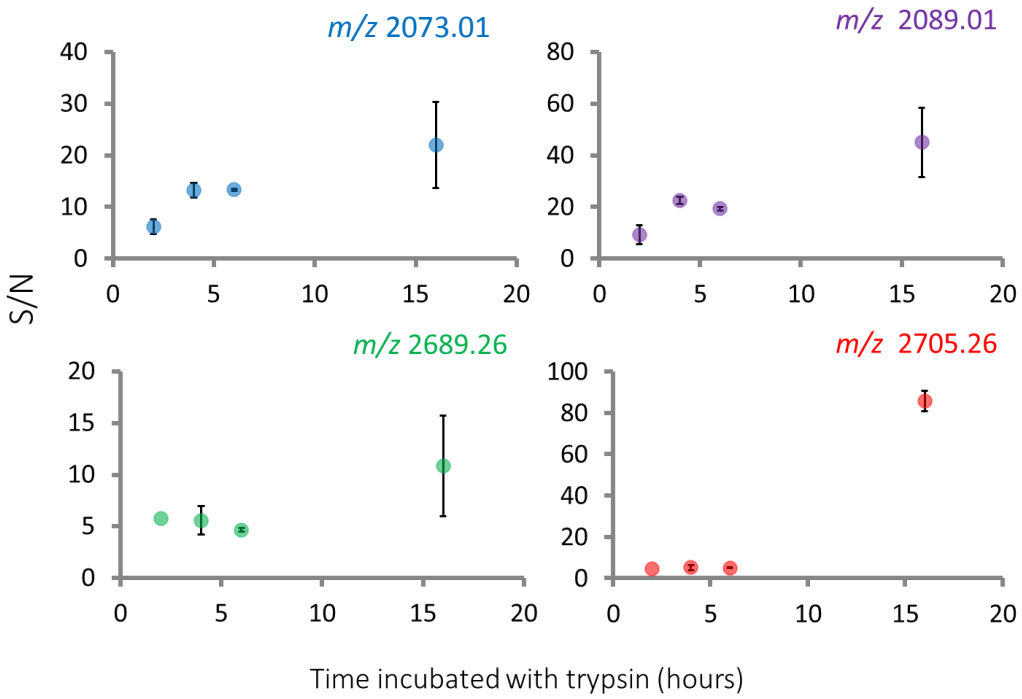


Figure 33: The S/N ratio of four peptides *m/z* 2073.01, *m/z* 2089.01, *m/z* 2689.26 and *m/z* 2705.26 in bovine protein extracts, after being incubated with trypsin for either: 2, 4, 6 or 16 hours. The S/N of these peptides was increased significantly after a digestion time of 16 hour. The error bars show the range in S/N values obtained between the replicate samples.

3.2.3 Conclusions

In was observed that although most of the Gln-containing peptides in Table 2 do not deamidate during overnight tryptic digestion, some do appear to show increased levels of deamidation over time. This experiment was carried out on one bone and therefore may not be more broadly representative of bone collagen. The preservation of the collagen may also play a role in how susceptible it is to deamidation. Further investigations of protein extracted from a range of different bones, of differing ages would help to assess this, as well as triplicate analysis as some of the duplicate values did not replicate well. It is thus important that samples are digested under the same conditions in order to make the results comparable. In order to make the data obtained in this study comparable to previous studies of Gln deamidation in collagen, 16 hour tryptic digestion will be used. It is also possible that this length of time is needed in order to increase the S/N level at which peptides with *m/z* 2073.01, m/z 2089.01, *m/z* 2689.26 and in particular peptide *m/z* 2705.26 are observed. There were no consistent differences between the sequences of the peptides which were easily cleaved after 2 or 16 hours of digestion. It may be that some peptides only undergo deamidation once they are free and therefore observed levels of deamidation might depend their release rates.

3.3 Comparison of Gln-containing peptides observed when analysed using standard or AnchorChip MALDI target plates.

Previous studies by van Doorn *et al.* (2012) and Buckley *et al.* (2008) have used standard MALDI target plates to analyse collagen peptides. One of the main advantages of this plate type is that they are easy to clean. In the protocols used here, these plates are cleaned using a 10 % surfactant solution (RBS 35, Detergent Concentrate, Thermo Scientific), rinsed with purified water and then polished with a brass polish. Once the polish is removed the plate is rinsed with isopropanol (IPA) and left to air dry. An alternative plate type which is used in proteomic analysis is the AnchorChip target plate. In AnchorChip plates, each spot position has a hydrophilic area, surrounded by a small hydrophobic ring. This hydrophilic anchor draws the sample into a small area, resulting in a more concentrated spot. A spot on an AnchorChip plate holds half the amount of sample (0.5 µL) and matrix as a standard MALDI plate (1 µL) and is advertised by Bruker as being able to increase the sensitivity by a factor of 10 to 100. It is therefore thought to be a preferable plate format for low concentration samples, where only a small sample volume is available. Another benefit of AnchorChip plates over standard MALDI plates, is that due to the smaller area over which the sample is dried, there is less chance that the laser will miss sample crystals as it scans across the spot; reducing the need to look for 'sweet spots' which can sometimes happen with standard MALDI plates. One of the main disadvantages of AnchorChip plates is that caution must be taken when cleaning as the hydrophilic and hydrophobic coatings can be damaged. AnchorChip plates are thus cleaned using a series of sonication steps in various solvents, depending on what the sample is soluble in, this can be tricky if you have a number of different sample types on one plate, or if a sample is particularly concentrated. Manufacturers recommend that you do not analyse samples where calibrants have been as high concentration samples can be difficult to remove. AnchorChip plates also have a limited life span as they can only be cleaned a number of times before the AnchorChip is worn away. This section will compare these two plate types, with the aim of determining which format type would be most suitable for the determination of deamidation in the 12 Gln-containing peptides measured in this thesis. It should be noted that in this investigation the standard recommended volumes for samples and matrix, for each plate type were used (1 µL for standard MALDI, and 0.5 µL for AnchorChip). Therefore differences in the spectra observed may be due to concentration, as well as the plate type. It was decided that 1 µL would be used on the MALDI plate as this is more representative of real-life use of the plate. Further investigations to determine only plate differences could be carried out using both 0.5 µL volumes on each plate.

A range of protein extracts from bison bones of different ^{14}C ages were used: EL1050 (50,500 yrs BP), EL1017 (12,365 yrs BP), QZ10346 (34,700 yrs BP) and QZ10220 (52,4000 yrs BP) (It should be noted that although these samples have been radiocarbon dated, some of the dates obtained are at the limits of radiocarbon dating and therefore it is possible that they are older and belong to the

associated dated tephra layer of around 80,000 years BP)The protein extracts from each sample were analysed using the two plate formats and the Gln- Containing peptides and α-values were compared.

3.3.1 Methods

From each of the four bone pieces (EL1050, EL1017, QZ10346 and QZ10220) ten chips were sampled and labelled A-J respectively. Collagen was extracted by warming in 50 mM ammonium bicarbonate pH 8.0 for one hour. 50 µL of each extract was then digested, adding 1 µL of porcine trypsin (0.4 µg/µL), and incubating for 16 hours at 37 °C. The digests were then purified using 100 µL C18 solid-phase tips (Millipore ZipTips) and desalted by washing with 0.1% trifluoroacetic acid (TFA) solution. Peptide mixtures were eluted in 50 µL of 50:50 (v/v) acetonitrile:MilliQ water (0.1% TFA). 1 µL of the each sample, followed by 1 µL of CHCA matrix was spotted on to the standard MALDI target plate. 0.5 µL of each sample followed by 0.5 µL of CHCA matrix was then spotted onto the Anchor Chip plate. All samples were spotted in triplicate. Both plates were left to air dry and analysed by MALDI-TOF-MS. α-values for the observed Gln-Containing peptides shown in Table 2 were measured.

3.3.2 Results

For each bone a maximum of 120 observations was possible (10 replicates x 12 peptides). Three out of the four samples, EL1017, QZ10346 and QZ10220, all had a higher percentage of observed Gln- containing peptides when analysed using a standard MALDI plate over the AnchorChip plate, the percentages weren't vastly different; however for sample EL1050 only 57.50 % of the Gln-containing peptides were observed using a standard MALDI plate, where as 89.17 % were observed in samples analysed using the AnchorChip plate (Figure 34).

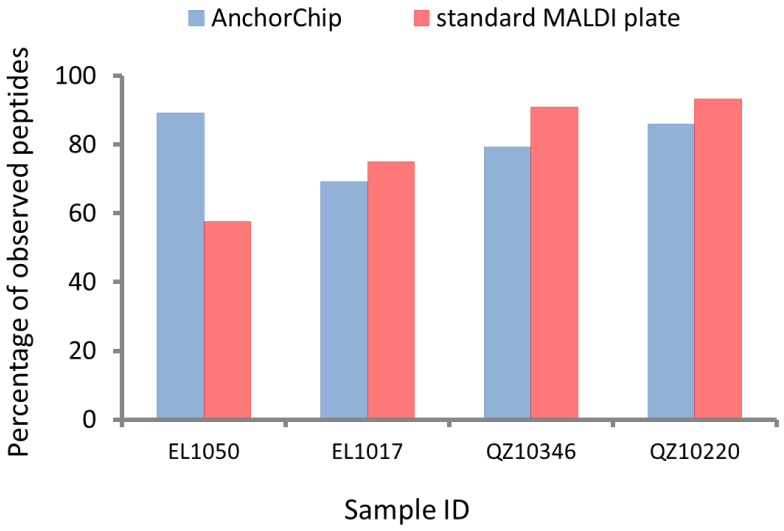


Figure 34: Percentage of peptides observed in four samples (EL1050, EL1017, QZ10346 and QZ10220) when analysed using either a standard or AnchorChip MALDI target plate

Out of the 12-Gln containing peptides measured, four (*m/z* 836.44, *m/z* 1105.57, *m/z* 2056.98, *m/z* 2073.01) gave similar α-values regardless of the type of MALDI plate used, however, three peptides (*m/z* 1690.77, *m/z* 1706.77, *m/z* 3100.41) produced consistently slightly higher average α-values when analysed using the standard MALDI plate, than those measured using the AnchorChip plate (Figure 35).

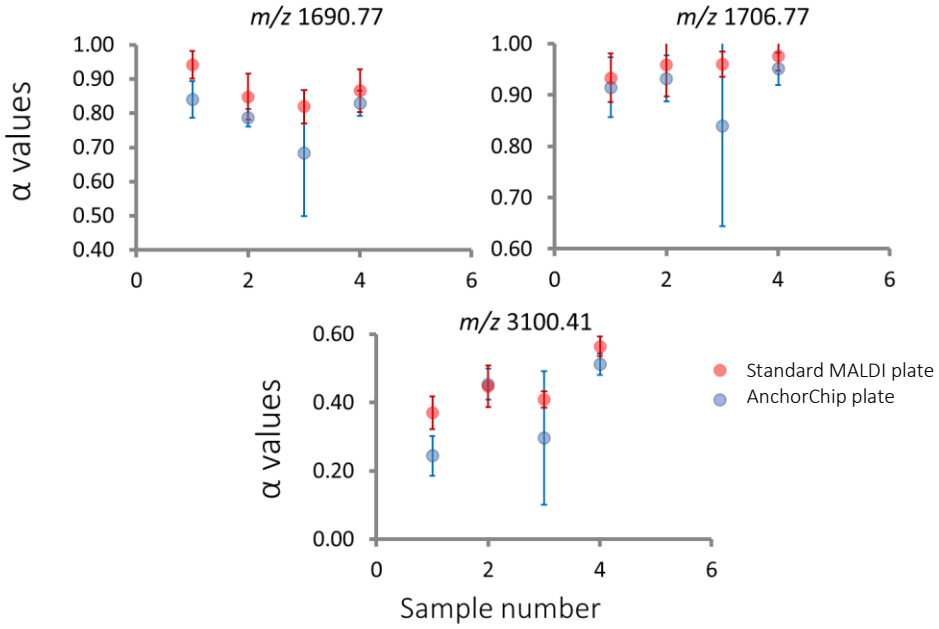


Figure 35: α-values measured in three peptides (m/z 1690.77, m/z 1706.77, m/z 3100.41) across four bison bone samples; 1 = QZ10220, 2 = QZ10346, 3 = EL1017 and 4 = EL1050. α-values from standard MALDI plate are shown in red, and values from AnchorChip plates are shown in blue

The peptide with *m/z* 3001.50 showed different α-values when analysed on the two plate formats, but did not show a correlation with the type of plate used (Figure 36). This peptide was not observed in sample EL1017 when analysed on the standard MALDI plate. The remaining four peptides (*m/z* 2089.01, *m/z* 2689.26 and *m/z* 2705.26) produced similar α-values for both plate types, in at least three of the four samples analysed.

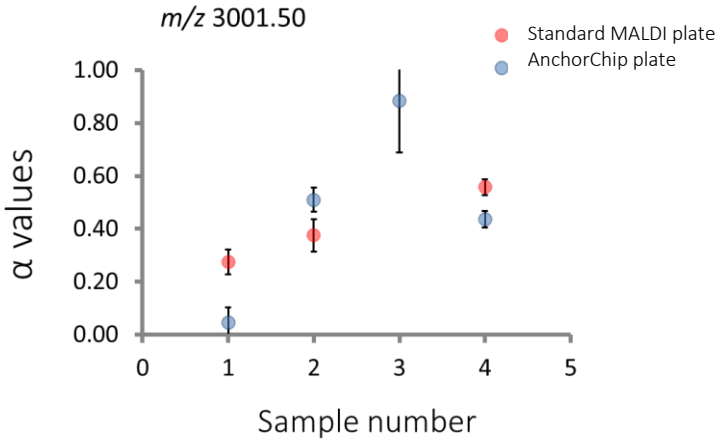


Figure 36: α-values from peptide *m/z* 30001.50 measured in four different samples: 1 = QZ10220, 2= QZ10346, 3 = EL1017 and 4 = EL1050

3.3.3 Conclusions

There are advantages and disadvantages of both plate formats (Table 3). On average, the standard MALDI plates gave better overall coverage of the Gln-containing peptides investigated in this thesis. Although the AnchorChip may be more suitable for the analysis of low concentration samples, especially those looking at measuring hydrophobic peptides, this was not investigated. From this data there is insufficient difference to draw a clear distinction. Because the AnchorChip did not significantly increase the sensitivity, and carries a risk of cross contamination it was not used further. A standard MALDI plate was used throughout the rest of the thesis.

Table 4: Advantages and disadvantages of the two plate formats investigated, Standard MALDI and AnchorChip plates

Advantages of MALDI plates	Disadvantages of MALDI plates	Advantages of AnchorChip plates	Disadvantages of AnchorChip plates
Standard MALDI plates are cheaper than AnchorChip plates as they can be re-used indefinitely.	Due to large spot size sometimes there are areas of the target circle which do not contain sample crystals, this can be missed during automated analysis.	Requires only 0.5 µL of sample per spot and therefore uses less sample, calibration solution and matrix.	Cleaning the plate can be difficult depending on the chemistry and nature of the samples.
Less chance of damaging the plate during cleaning.	Not good for samples of low concentration, multiple spots may be required for low quality samples.	Good for analysing hydrophobic compounds.	Possibility of cross contamination if not cleaned correctly.
Because they can be cleaned as often as you like, they can be cleaned between analysis sets regardless of whether the whole plate has been used or not.	Require 1 µL of sample or calibrant per spot plus 1 µL of matrix solution.	Resulting spot is concentrated and therefore more likely to produce a uniformed, densely packed circle of sample crystals.	AnchorCips are more expensive than standard MALDI plates and can only be reused a certain number of times due to the chemistry of the chemical anchors.
		Good for analysing samples of very low concentration.	

3.4 Reproducibility of α-values

In order to test the reproducibility of α-values obtained using the finalised standard preparation protocol (see Chapter 2); collagen was extracted from some chips of Roman bovine bone. This sample was then spotted across 30 spots across a MALDI plate, generating ten α-values. The same spots were then analysed a second time and α-values across 12 Gln-containing peptides compared.

3.4.1 Method

Collagen was extracted from a bone chip ~ 30 mg , and 100 µL of 50 mM ammonium bicarbonate pH 8.0 was added. The sample was warmed at 65 °C for one hour. The 50 µL of the supernatant was removed and digested over night with 1 µL of porcine trypsin (0.4 µg/µL), and incubated for 16 hours at 37 °C. The digests were then purified using 100 µL C18 solid-phase tips (Millipore ZipTips)

and desalted by washing with 0.1% trifluoroacetic acid (TFA) solution. Peptide mixtures were eluted in 50 µL of 50:50 (v/v) acetonitrile:MilliQ water (0.1% TFA). 1 µL of the each sample, followed by 1 µL of CHCA matrix was spotted on to the standard MALDI target plate; this was repeated across 30 spots. The plate was left to air dry and analysed twice by MALDI-TOF-MS.

3.4.2 Results

An average and the standard deviation of the ten α-values obtained from each of the two sets analysis were compared. All but one of the 12 peptides (*m/z* 3001.50) produced similar average α-values in both sets of analysis (Figure 37). Some peptides produced peptides with a large standard deviation (*m/z* 3001.50 and *m/z* 2689.26) with others showing little or no standard deviation (*m/z* 836.44, *m/z* 1105.57 and 1706.77).

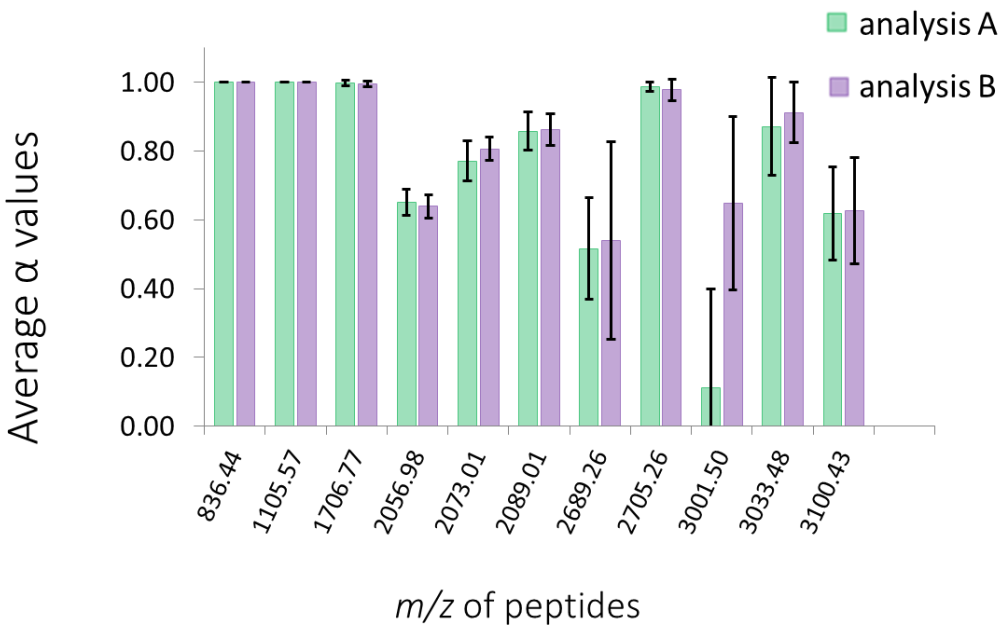


Figure 37: Average α-values taken from ten replicates analysed twice. The standard deviation of α-values is shown for each peptide analysis.

Peptide with an *m/z* 1105.57 (Figure 38, green dots) showed little deamidation and no variability across the 20 spots with consistent α-values of 1.00 ; peptides which showed similar or overlapping levels of deamidation were *m/z* 836.44, *m/z* 1706.77 and *m/z* 2075.26, with average α-values of 1.00, 100 and 0.98 respectively. *m/z* 2089.01 (Figure 38, blue dots) is the hydroxylated version of peptide *m/z* 2073 (Figure 38, red dots); both of these peptides showing similar average α-values of 0.86 and 0.79 respectively. *m/z* 2056.98 produced the lowest of average α-value of out of the four peptides shown in Figure 38, with an average of 0.64. All four of these peptides showed low standard deviations ranging from 0.00 – 0.05.

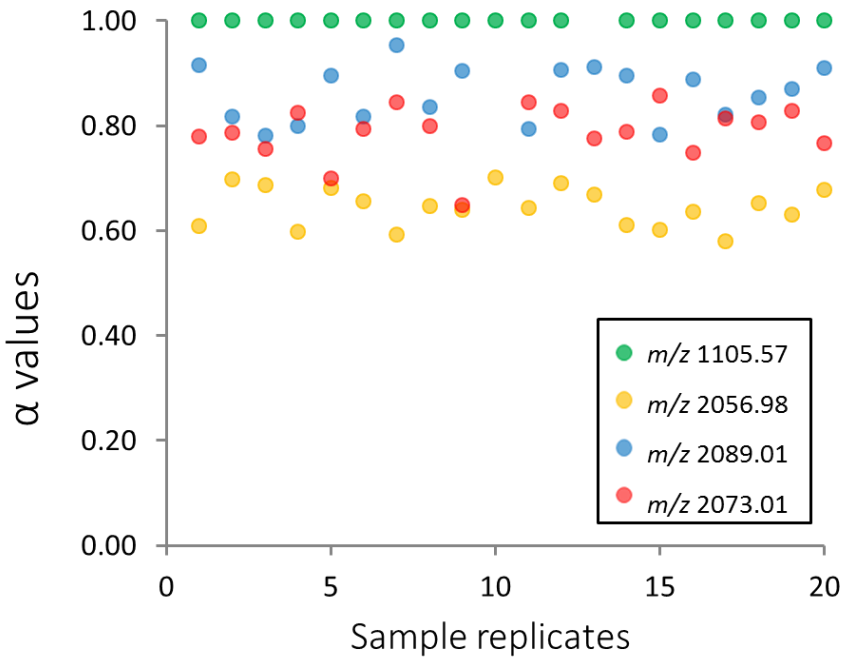


Figure 38: α-values from 20 spots of the same sample for peptides *m/z* 1105.57 *m/z* 836.44, *m/z* 1706.77 and *m/z* 2075.26

Peptides which did not appear to give reproducible α-values were those with *m/z* 3033.48, *m/z* 3100.43 and *m/z* 3001.50 (Figure 39), with average alpha values of 0.89, 0.62 0.57 respectively. It should be noted that peptide *m/z* 3001.50 was observed in only five out of the 20 samples. It does appear that the three heaviest peptides (*m/z* 3033.48, *m/z* 3100.43 and *m/z* 3001.50) (Figure 39) not only appear more deamidated than the three smallest (*m/z* 1105.57 *m/z* 836.44 and *m/z* 1706.77), but also show higher levels of variability.

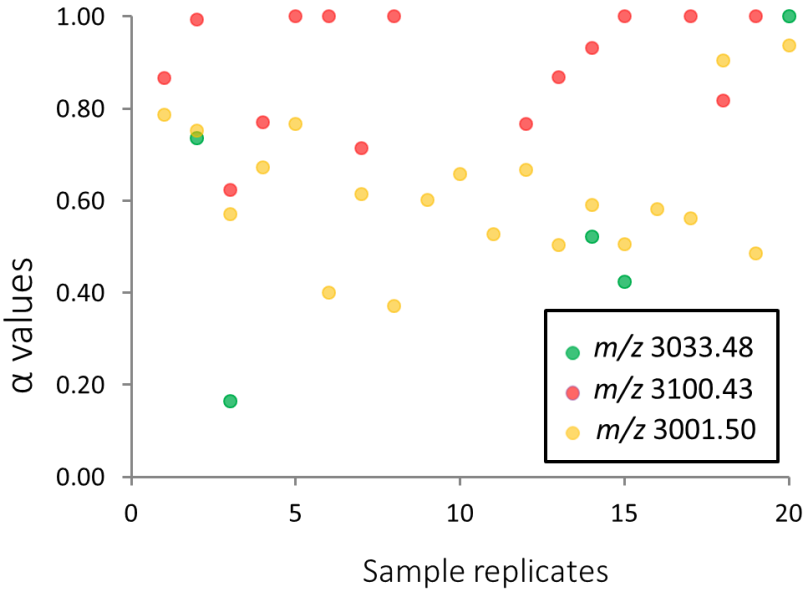


Figure 39: α-values from 20 spots of the same sample for peptides *m/z* 3033.48 *m/z* 3100.43, and *m/z* 3001.50

3.4.3 Overall conclusions

It appears that within the sample analysed here, some peptides (*m/z* 1105.57 *m/z* 836.44, *m/z* 1706.77 and *m/z* 2075.26) show very little variation in α-values; however these peptides have undergone no to very little levels of deamidation, with most having average α-values of 1.00. It is therefore not possible to say how variably these peptides might deamidate without looking at samples in which a larger extent of deamidation has occurred. These three peptides appear to be the most stable of the 12 Gln-containing peptides observed in this sample, and may be useful when further evaluating levels of collagen preservation in older bone. Three peptides which had undergone deamidation in this sample were *m/z* 2089.01 *m/z* 2073 and *m/z* 2056.98. These peptides showed good levels of reproducibility with standard deviations for the three peptides being 0.05, 0.05 and 0.04 respectively. Peptides *m/z* 3033.48, *m/z* 3100.43 and *m/z* 3001.50 did not show good levels of reproducibility and therefore may not be suitable markers of collagen preservation.

As the actual level of deamidation in this sample is not known, it is not possible to determine how accurate the measurements of deamidation in these peptides are; however it does give some idea of how precise and reproducible each Gln-containing peptide might be. Without knowing the actual level of deamidation in each sample it is difficult to determine whether the variation observed in peptides in *m/z* 3033.48, *m/z* 3100.43 and *m/z* 3001.50 is due to analytical error, or a mixture of the same peptides with varying levels of deamidation.

Chapter 4 - Investigating measurements of glutamine/glutamic acid ratios in synthetic peptides using mass spectrometry

Introduction

In order to determine the ratios of glutamine to glutamic acid using mass spectrometry, it has generally been assumed that each of the two forms of the peptide (deamidated and undeamidated) have equal ionisation efficiencies. If this is correct, the ratios of the peak areas of the glutamine/glutamic acid (Gln/Glu; Q/E)-containing species would be directly proportional to the concentration ratios of the two peptide forms.

Matrix-assisted laser desorption/ionisation (MALDI) coupled to a time of flight (TOF) mass spectrometer has been routinely used for the analysis of peptide mixtures, and over recent years for the estimation of levels of deamidation in various sample types. One disadvantage of this instrumentation is that due to the insufficient resolving power of the mass analyser, it was not possible to resolve the two forms, as the n^{th} peak (typically the mono-isotopic signal) of the deamidated peptide signal overlaps the $(n+1)^{th}$ peak (typically the signal for the species containing one ^{13}C atom) of the undeamidated form. In order to calculate the Q/E ratio, a genetic algorithm has been used to deconvolute the overlapping isotopic distributions of the two peptide forms (van Doorn *et al*., 2012; Wilson *et al*., 2012).

In order to investigate how accurate this approach is for the estimation of glutamine deamidation in biological samples such as protein extracted from archaeological bone, a reference standard is needed. Consequently, two peptides have been synthesised with the target amino acid sequences TyrAlaTyrGlyHypGlyGlnValGly and TyrAlaTyrGlyHypGlyGluValGly. The synthesised peptides were analysed using LC-MS to assess their purity. To investigate the applicability of MALDI for estimating the ratio of glutamine/glutamic acid-containing forms of a peptide, and the accuracy of the published deconvolution method (van Doorn *et al*., 2012; Wilson *et al*., 2012) the two peptide forms were mixed in known ratios and analysed using MALDI-TOF-MS and MALDI and ESI-FT-ICR-MS. The data were used to examine the ionisation responses of the two peptide forms using ESI and MALDI, and to assess the data from the two mass spectrometer types for determining the relative amounts of Q and E-containing peptides.

4.1 Methods

4.1.1 Solid phase peptide synthesis

When designing the peptides, several factors were taken into consideration. The peptides needed to be larger than 800 Da, so as to not fall in the same region as the CHCA MALDI matrix peaks. The

length of the peptide is also important, as longer peptides have more chance of secondary structures developing during synthesis, as well as generally decreasing peptide purity due to an increase in the number of coupling steps. It was therefore decided that, in order to maximize peptide purity, the chosen peptide should be no longer than nine amino acids in length. In collagen type I, a third of the protein is made up of glycine, so in order to make the peptide similar to the structure of a collagen peptide, a third of the chosen amino acids were glycine. As glycine (G) is the smallest amino acid with a residue mass of 57 Da, two tyrosine residues (Y, residue mass of 163 Da) were included in the peptide to offset the small size of the glycines. The resulting target amino acid sequences were TyrAlaTyrGlyHypGlyGlnValGly (referred to as product peptide Q) and TyrAlaTyrGlyHypGlyGluValGly (product peptide E), with masses 926 Da and 927 Da respectively.

4.1.1.1 Reagents

The water used in the following protocols was purified using a MilliQ system with the resulting purified water having a resistivity of 18 MΩ. This is referred to throughout as purified water. Amino acids and reagents used were purchased from Sigma Aldrich. These included: pre-loaded glycine resin (H-Gly-2-ClTrt resin), Fmoc-Val-OH, Fmoc-Glu (tBu)-OH, Fmoc-Gly-OH, Fmoc-Hyp (tBu)-OH, Fmoc-Try (tBu)-OH, Fmoc-Ala-OH. Fritted polypropylene (PP) tubes (10 mL) were used as vessels for the solid phase reactions. During the reaction, samples were agitated using a Stuart blood rotator. Reagents used included: trifluoroacetic acid (TFA), dimethyl formamide (DMF), dichloromethane (DCM), methanol (MeOH), piperidine (PIP), triisopropylsilane (TIS), (6-chloro-1H-benzotriazol-1-yl)oxy](dimethylamino)-N,N-dimethylmethaniminium hexafluorophosphate (HCTU), diisopropylethylamine (DIPEA), purified water, cold diethyl ether, liquid nitrogen.

4.1.1.2 Coupling reaction

Both the product Q and E-containing peptides were prepared at the same time in two separate PP tubes. To each tube a 100 mg aliquot of H-Gly-2-ClTrt resin plus 10 mL of DMF was added, in order to cause the resin to swell. The tubes were rotated on a Stuart rotator for 30 minutes at room temperature. The DMF was then removed and this process was repeated three more times. Fmoc-Val-OH (187 mg) and HCTU (223 mg) were dissolved in 10 mL of DMF, and 203 µL DIPEA was added. This solution was prepared in duplicate with 10 mL of the solution added to each of the PP tubes. The tubes were then transferred to the Stuart rotator and left to rotate for one hour. The solution was removed from both tubes and Fmoc deprotection was carried out by adding 10 mL of piperidine solution (20 % in DMF) to each tube and rotating for two minutes. The piperidine solution was then removed and this process was repeated a further four times. Once the final piperidine aliquot had been removed, 10 mL of DMF was added to each tube and the sample was left to rotate for 2 minutes. The DMF was then removed and this process was repeated a further four times.

The coupling reaction was repeated with the same conditions, substituting valine for the following amino acid derivatives and amounts: Fmoc-Glu (tBu)-OH (234 mg)/ Fmoc-Gln (tBu)-OH 234 mg), Fmoc-Gly-OH (163 mg), Fmoc-Hyp (tBu)-OH (225 mg), Fmoc-Gly-OH (163 mg), Fmoc-Tyr-(tBu)-OH (253 mg), Fmoc-Ala-OH (171 mg), Fmoc-Tyr-(tBu)-OH (152 mg). Each amino acid was coupled once, except for the Gly preceding Hyp in the coupling sequence, for which the coupling step was performed twice.

4.1.1.3 Cleavage and isolation of the peptides

The resin was isolated by filtration and dried under high vacuum for three hours. The cleave solution consisted of H^2O:TIS:TFA (2.5:2.5:95; v:v:v). 10 mL of cleave solution was added to each tube containing the resin. The tubes were left to rotate for two hours. The sample solution was then separated from the resin by filtration. The resin was then washed twice with 3 mL of TFA. The two washes were then combined with each of the corresponding sample solutions. 40 mL of dimethyl ether were transferred into PP falcon tubes and left to chill at approx. -20 °C for 2 hours. The sample solutions were each transferred into separate falcon tubes of ice cold diethyl ether, in which the product precipitated. Each sample was centrifuged for ten minutes (6,000 x g) and the supernatant removed and discarded. The pellet was then re suspended in ice cold diethyl ether and the process repeated a further three times. The remaining pellet was dissolved in 2 mL of water and lyophilised. The resulting powdered samples were stored in the freezer (approx. -20 °C).

4.1.2 Measuring levels of deamidation using mass spectrometry

This study used two mass spectrometers to generate data from which to calculate the ratios of the two product peptides. The first, a Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) is a high resolution instrument, with both MALDI and ESI ionisation sources. The second instrument was a lower resolution time of flight (TOF) mass spectrometer, with a fixed MALDI source. When using the FT-ICR-MS to analyse a mixture of peptides Q and E (1:1 ratio), it was possible to resolve the two peptide forms, as the nth peak of the deamidated peptide signal is resolved from the $(n+1)^{th}$ peak of the undeamidated signal (Figure 40); this was not possible using the TOF instrument. A genetic algorithm was thus used to deconvolute the overlapping isotopic distributions in the data from the MALDI-TOF-MS (Wilson *et al.*, 2012).

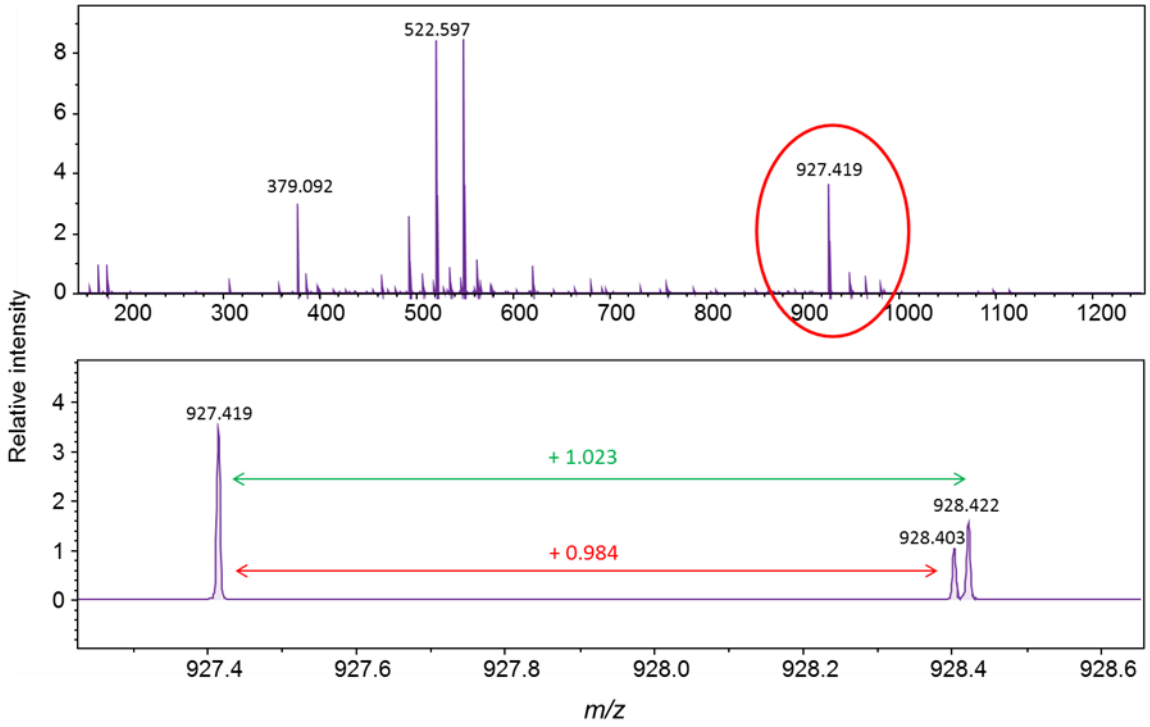


Figure 40: ESI spectrum obtained from the analysis of a mixture of product peptides Q and E (1:1 ratio), using FT-ICR-MS. The peak at *m/z* 927.419 corresponds to $M+H^+$ for product peptide Q. The two peaks at *m/z* 928.403 ($M+H^+$ for peptide E) and *m/z* 928.422 (one ^{13}C atom in $M+H^+$ for peptide Q) are fully resolved from one another. The ratio of the peak areas of the peaks at *m/z* 927.419 and *m/z* 928.403 was used to estimate the ratio of Q to E-containing peptides

4.2 Peptide purity analysis by HPLC-MS and MALDI-TOF-MS

The dried peptides were suspended in water at a concentration of 100 ppm. Each peptide was analysed using an HPLC-HCTultra PTM Discovery System, fitted with a symmetry C18 3.5 µm (4.6 x 7.5 mm) column, using mobile phases of acetonitrile (A) and water (B). The elution was isocratic with a flow rate of 1 mL/min and a total run time of 9 minutes. A multi-step gradient of A and B was used to elute the peptide products (Table 4; Figure 41).

Table 5: percentage of mobile phase A (acetonitrile) used over a 9 minute gradient

Time (minutes)	% A
0	95
1	95
6	70
6.1	5
7	5
7.2	95
9	95

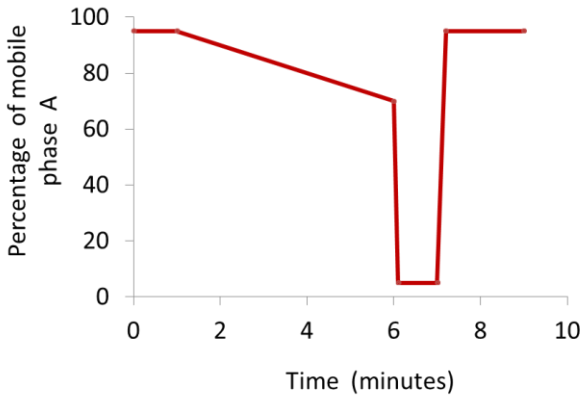


Figure 41: Isocratic elution gradient showing the percentage of mobile phase A (acetonitrile) over the total run time of 9 minutes

4.2.1 Measuring the ratios of the product Q and E peptides using MALDI-TOF, MALDI-FT-ICR and ESI-FT-ICR

Once the purity of each peptide (in solution at 100 ppm in 50:50 ACN:H2O) was determined the solutions were mixed in the following Q/E ratios : 100/0, 90/10, 70/30, 50/50, 30/70, 10/90 and 0/100. This set of seven point calibration standards was then analysed using two different mass spectrometers and two ionisation sources. The two instruments used were the Bruker ultraflex III and the Bruker solariX XR. The ultraflex was used with a MALDI source and the solariX with either an ESI or MALDI source. Both instruments were used in positive ion mode.

The resulting peptide mixtures were analysed using matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF-MS/FT-ICR-MS). A volume of 1 µL of sample solution was spotted on a ground steel MALDI target plate, followed by 1 µL of α-cyano-4-hydroxycinnamic

acid matrix solution (1 % in 50 % ACN/0.1 % TFA (w/v/v)). The sample and the matrix solutions were mixed together on the plate and allowed to air-dry. Each sample was spotted on to the MALDI target plate in triplicate. When using the Bruker ultraflex instrument positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in reflectron mode, equipped with a Nd:YAG smartbeam laser. MS spectra were acquired over a range of 800-3000 *m/z*. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). Bruker flexAnalysis software (version 3.3) was used to perform spectral processing and peak list generation. The same sample MALDI plate was analysed using the Bruker solariX FT-ICR-MS instrument. It was operated in positive mode, with a smartbeam nitrogen laser (337 nm) with attenuation setting of 40 %. Each spectrum was acquired using 200 laser shots, and an average spectra setting of 10. A transient length of 2.5166 seconds was used. The same calibrant spots were used for both instruments (ultraflex and solariX). The spectra were obtained using the solariXcontrol software and processed with DataAnalysis 4.2 software.

4.3 Results

4.3.1 Purity analysis of peptides by LC-MS and MALDI-TOF-MS

4.3.1.1 Determination of peptide sequences

MS/MS of the MALDI sample spots prepared with either 100 % product peptide E or product peptide Q was carried out using MALDI-TOF-MS. The assignment of the sequence of each of the peptides was determined using de novo interpretation of the product ion spectrum (Figures42 and 43).

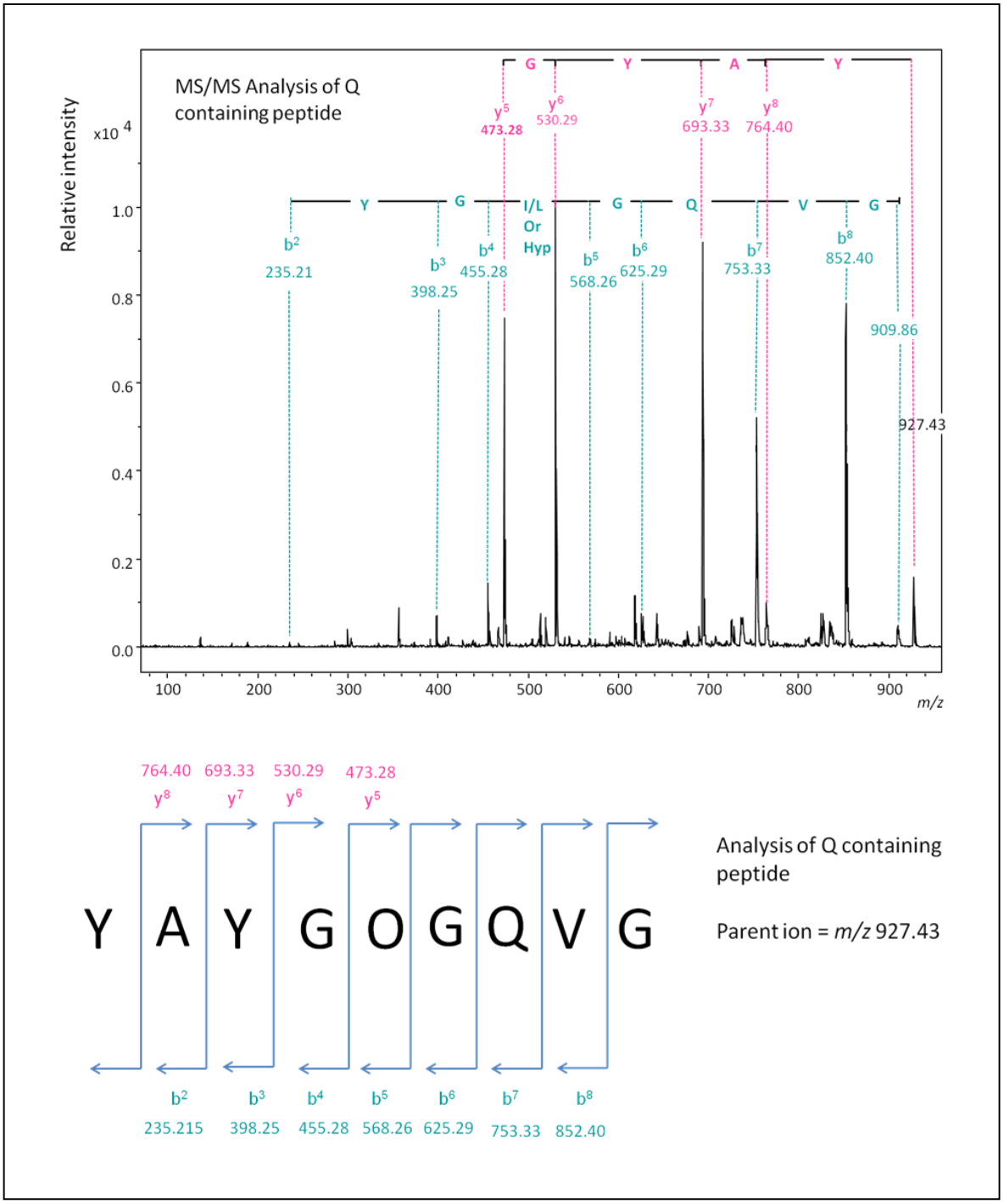


Figure 42: Product ion spectrum obtained from the MS/MS analysis of product peptide Q showing the fragmentation of the y and b ion series

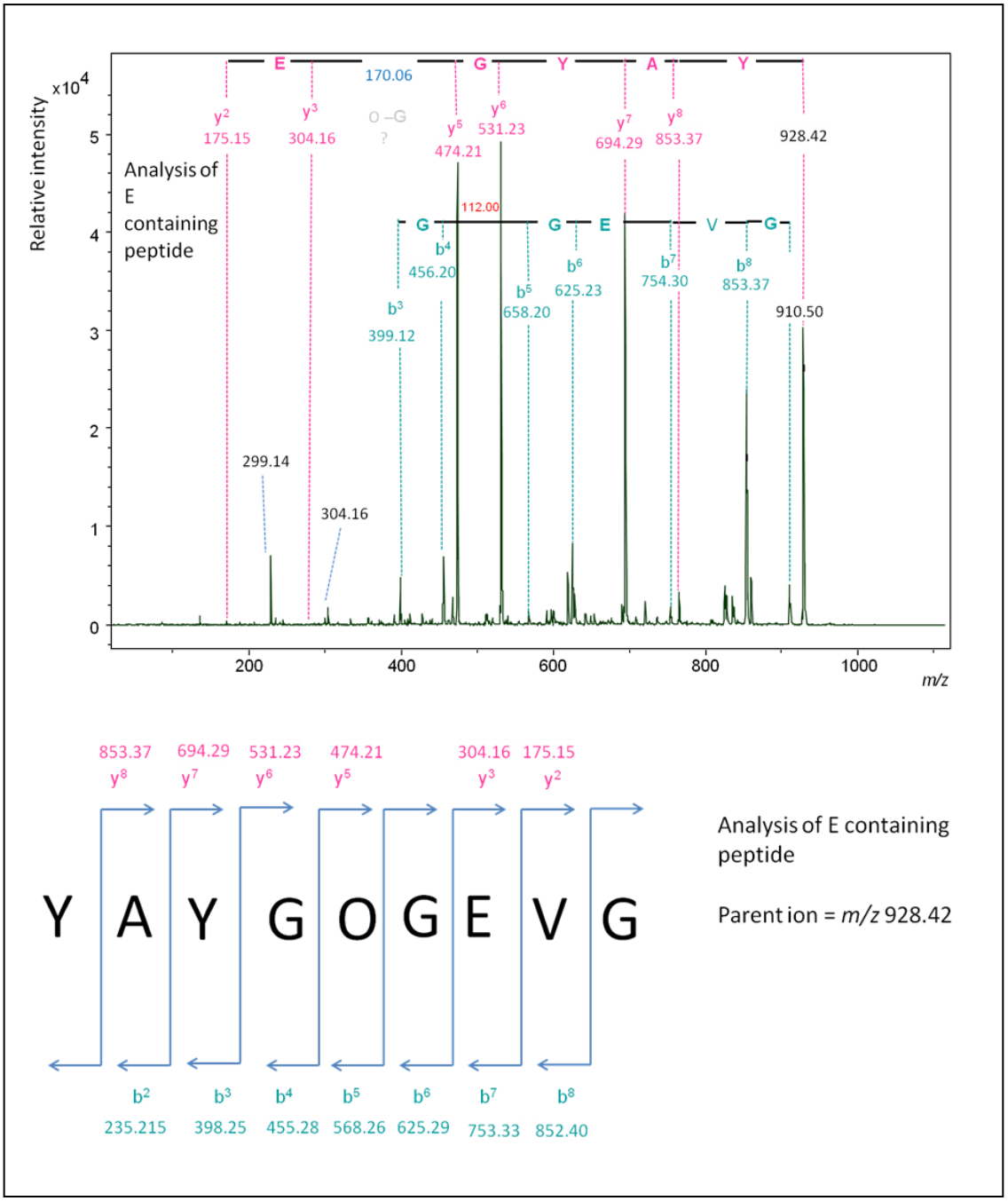


Figure 43: Product ion spectrum obtained from the MS/MS analysis of product peptide E showing the fragmentation of the y and b ion series

The product ion spectrum for product peptide E showed a mixture of both the Q and E-containing peptides. Harsh acidic treatments were used to cleave the synthetic peptides from the resin. It is therefore possible that some of the glutamine in the product peptide Q, may have undergone some minor deamidation during synthesis (Chapter 6).

4.3.1.2 Analysis of the purity of product peptide Q

In order to assess the purity of the two synthesised peptides, each peptide was first analysed separately by reversed phase LC-ESI-MS. UV absorbance (210 – 380 nm) data were also collected. As expected, similar absorbance patterns were observed for each of the two peptides. The UV chromatogram for product peptide Q contained three peaks. The main peak (peak A) was at *t*R 3.9

minutes and had a short tail. Two smaller peaks at *t*R 4.1 (peak B) and *t*R 4.4 (peak C) were also detected (Figure 44).

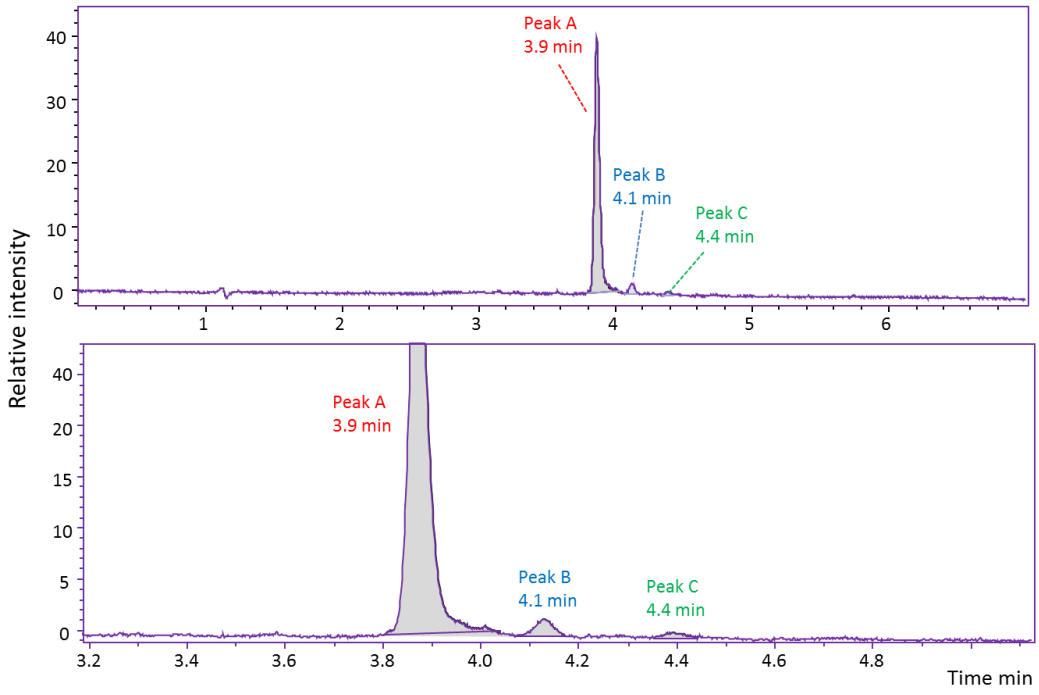


Figure 44: Chromatogram obtained from LC analysis of product peptide Q. The main peak is observed at *t*R 3.9, with two smaller peaks at *t*R 41 and *t*R 4.4 minutes. Panel two (bottom) is a focused in on the peaks indicated in panel one (top)

The spectrum generated from the main peak at *t*R 3.9 min contained a number of singly and doubly charged signals (Figure 45, peak A). The two most intense signal clusters in the spectrum corresponded to the mass expected for peptide Q, $M+H^+$ at *m/z* 927.4 and $[M+2H]^{2+}$ at *m/z* 464.3. The signals at *m/z* 949.3 and 971.4 correspond to $M+Na^+$ and $[M+2Na-H]^+$ respectively. In addition to the main peak observed in the chromatogram, there was a second peak observed at *t*R 4.1 min (Figure 45, peak B). ESI-MS analysis of this peak yielded signals corresponding to peptide E, with $M+H^+$ at *m/z* 928.3 and $[M+2H]^{2+}$ at *m/z* 464.7, along with some low *m/z* signals that have not been assigned. A third peak was observed at *t*R 4.4 minutes. ESI-MS analysis of this peak yielded a spectrum containing one doubly charged signal at *m/z* 422.6 and singly-charged signals at *m/z* 448.8 and *m/z* 806.3 (Figure 45, Peak C).

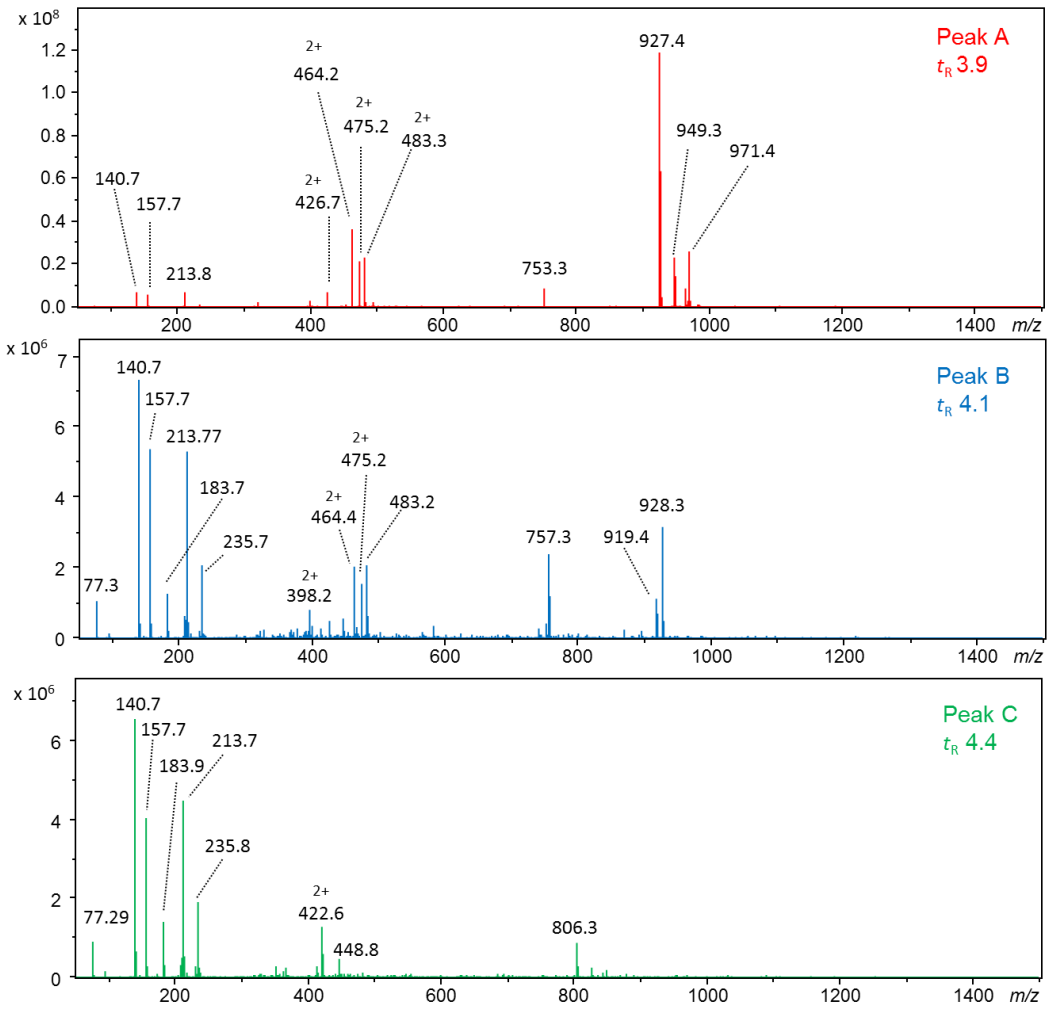


Figure 45: MS data obtained for three peaks present in the UV chromatogram of product peptide Q. The three spectra shown correspond to the peaks at *t*R 3.9 minutes (peak A), *t*R 4.1 min (Peak B) and *t*R 4.4 minutes (Peak C)

On the basis of the peak areas in the UV chromatogram, product peptide Q is estimated to be ~ 94.81 % pure and to contain 3.76 % of peptide E. This assumes that peak A contains solely the Q-containing product, and that all components give a similar UV response. From this analysis it is not possible to distinguish whether the small amount of peptide E in product peptide Q arises from deamidation during peptide synthesis, or from the presence of glutamic acid with the glutamine precursor used to make the peptide.

4.3.1.3 Analysis of the purity of product peptide E

The UV chromatogram obtained from product peptide E contained one main peak with a *t*R of 4.2 minutes (peak B) and two smaller peaks at *t*R 3.9 (peak A)and 4.8 (peak C)(Figure 46).

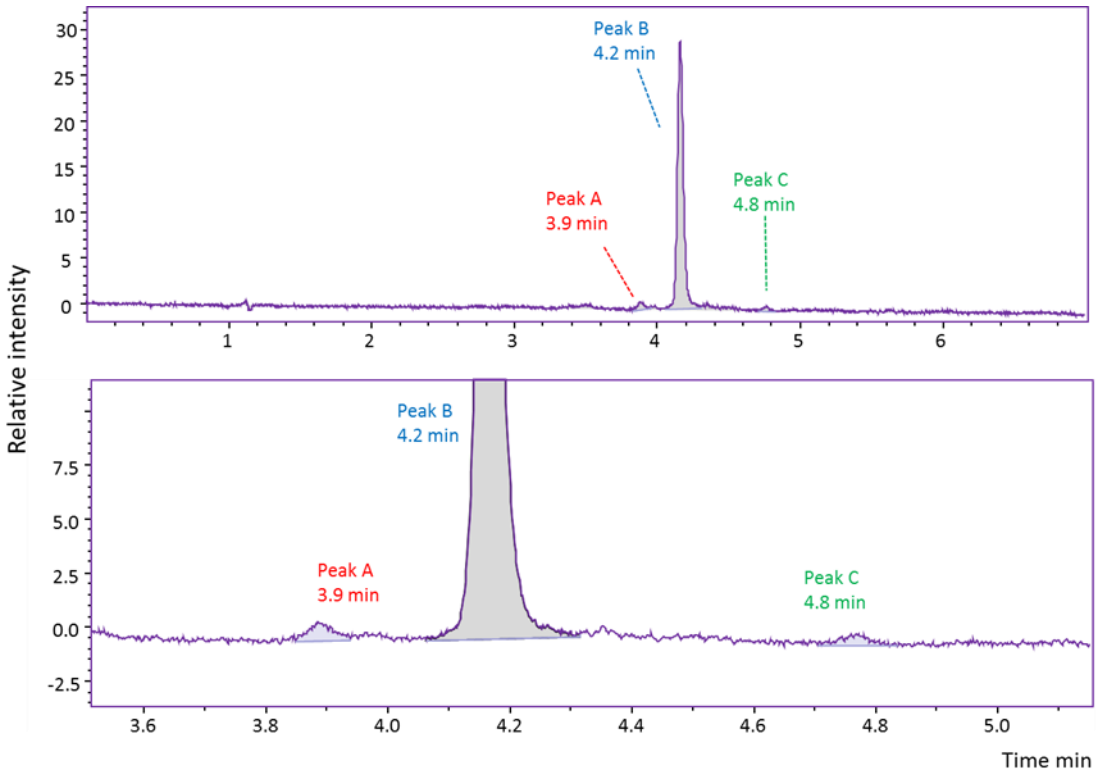


Figure 46: UV chromatogram obtained on LC analysis of product peptide E. The main peak is observed at *t*R 4.2 minutes (peak B), with two smaller peaks at *t*R 3.9 minutes (peak A) and 4.8 minutes (peak C)

ESI-MS analysis of the peak at *t*R 4.2 resulted in signals at *m/z* 464.7 ($[M+2H]^{2+}$) and *m/z* 928.4 ($M+H^{+}$) (Figure 47, peak B) for the expected E-containing peptide. ESI-MS analysis of the peak with a *t*R of 3.9 minutes resulted in signals at *m/z* 927.4 and 464.2, corresponding to $M+H^{+}$ and $[M+2H]^{2+}$ for product peptide Q. This conclusion is consistent with the results of MS/MS analysis of the peptide E product, which was demonstrated to be a mixture of the E and the Q containing peptides. The peak in the chromatogram at *t*R 4.2 minutes did not contain peaks which correspond to either of the masses expected for the Q and E containing peptides (Figure 47), although the main signals are 1 amu higher than those in the equivalent peaks observed in product peptide Q and are therefore consistent with the use of E rather than Q in the synthesis. Taking into account the two observed impurities product peptide E was estimated to be ~ 95.40 % pure and contain 2.68 % of product peptide Q.

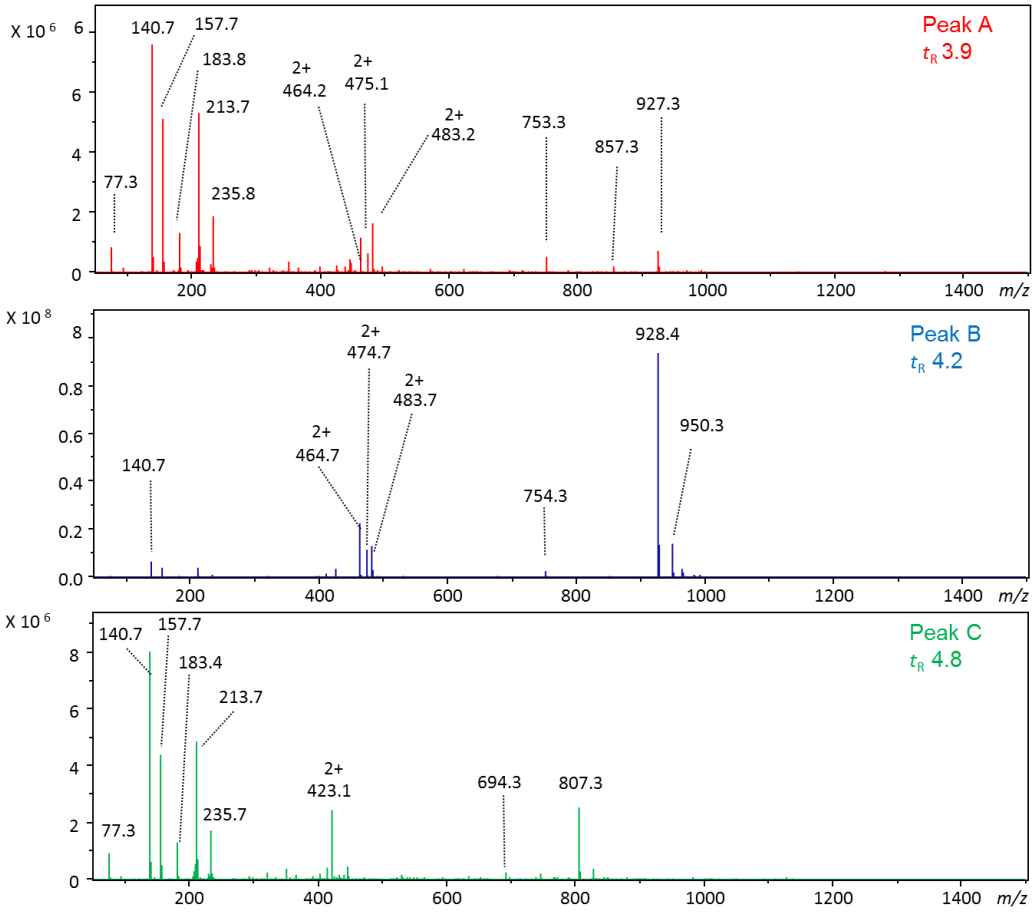


Figure 47 : MS data obtained for three peaks present in the UV chromatogram of product peptide E. The three spectra shown correspond to *t*R 3.9 minutes (peak A), *t*R 4.2 (Peak B) and *t*R 4.8 (Peak C).

The expected structures of product peptides Q and E are shown with the corresponding amino acid sequences (Figure 48).

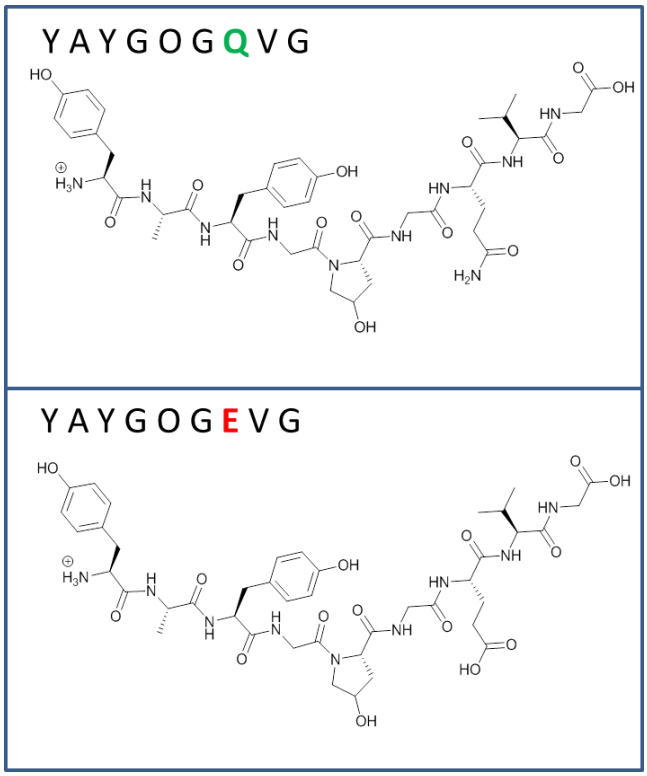


Figure 48: The expected structures of product peptides Q and E, product peptide Q (top) has the amino acid sequence YAYGOGQVG and a mass of 926 Da. Product peptide E (bottom) has the amino acid sequence YAYGOGEVG with a mass of 927 Da

4.4 Comparison of the ratio of product peptides Q to product peptide E, measured using different mass spectrometers and ionisation sources.

2.02 mg of product peptide Q and 1.96 mg of product peptide E were transferred into separate plastic PP microfuge tubes and dissolved in 2 mL of 50:50 (v: v) ACN: purified water. The resulting 1000 ppm stock solutions were diluted to 100 ppm, and mixed in the ratios shown in Table 5.

Table 6: Theoretical percentage of product peptides containing Q or E and corrected percentages taking into account the assigned peptide purities

Theoretical percentage of Q	Theoretical percentage of E	Corrected percentage of Q taking into account peptide purity	Corrected percentage of E taking into account peptide purity
0	100	2.51	97.49
10	90	12.20	87.80
30	70	31.35	68.65
50	50	50.22	49.78
70	30	68.81	31.19
90	10	87.13	12.87
100	0	96.19	3.81

In order to investigate possible differences in ionisation efficiency of product peptides, the seven samples were analysed directly, without chromatographic separation, using both ESI-FT-ICR-MS and MALDI-FT-ICR-MS, on which instrument the mass resolution is sufficient to resolve the deamidated from the ^{13}C isotopic signals. In addition, the seven samples were also analysed using MALDI-TOF-MS in order to assess the results of estimating the extent of glutamine deamidation obtained using this instrument; these data required use of the deconvolution method described in Wilson *et al*. (2012), since the mass resolution is insufficient to resolve the deamidated and the ^{13}C isotopic signals. Use of these two instruments made it possible to compare the quality of the data from a top of the line instrument (FT-ICR) with a much more routinely available instrument (MALDI-TOF).

4.4.1 FT-ICR-MS Results

Using a Bruker solariX XR FT-ICR-MS, two different ionisation sources (MALDI and ESI) were compared. The seven different mixtures of the two synthetic peptides described in Table 5 were analysed. The ESI and MALDI peak areas for the first glutamine- and glutamic acid-containing peptide peaks in the isotopic distributions were measured for each mixture. Each peptide mixture was analysed six times. For each peptide mixture, the average peak areas from each of the product peptides Q and E were plotted against their corresponding theoretical percentages (Figure 49).

Using ESI (Fig 49, left panel), average peak areas for mixtures of product peptides Q and E, ranging from 0 - 50 % Q produced peak areas for peptide Q and peptide E that are reasonably similar to each other, in spite of the excess of peptide E in these mixtures. In contrast, in samples ranging from 50 - 100 % product peptide Q, peptide Q gave higher peak areas than peptide E, as might perhaps be expected, since it is present in higher abundance. When the same mixtures were analysed using MALDI (Fig 49, right panel), peptide Q consistently produced larger peak areas than peptide E, regardless of the percentage concentration of the peptide. The errors observed in peak areas when samples were analysed by MALDI were consistently greater than those observed in samples analysed using ESI.

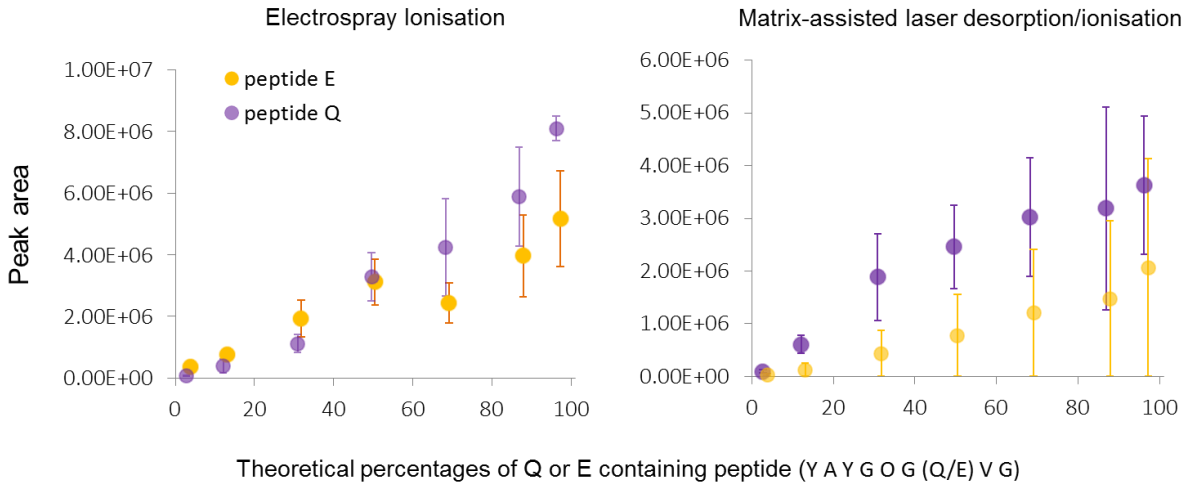


Figure 49: Peak areas from Q and E-containing peptides present in different mixtures of the two peptides, measured using two different ionisation sources, ESI (left) and MALDI (right). The error bars show the standard deviation of the measurements of six analytical replicates

The results shown in Figure 49 were used to determine the ratios of product peptide Q to product peptide E in each of the mixtures. The measured peak area ratios were expressed as percentages and compared to the theoretical percentages (Figure 50). When using the ESI source, the percentage of the Q containing peptide was predicted well across the mixture range, with the average difference between theoretical and measured values being only 1.04 %. The percentage of peptide Q in the same mixtures determined using MALDI resulted in a greater difference between theoretical and measured values, with an average of 14.94 % difference, ranging from 2.81 - 29.65%. This offset was greatest in the range 10-70 % Q, with calculated values for percentage Q overestimated by 16.80 to 29.65 %. It is evident that during MALDI ionisation in positive ion mode that ionisation of the less acidic Q-containing peptide is favoured.

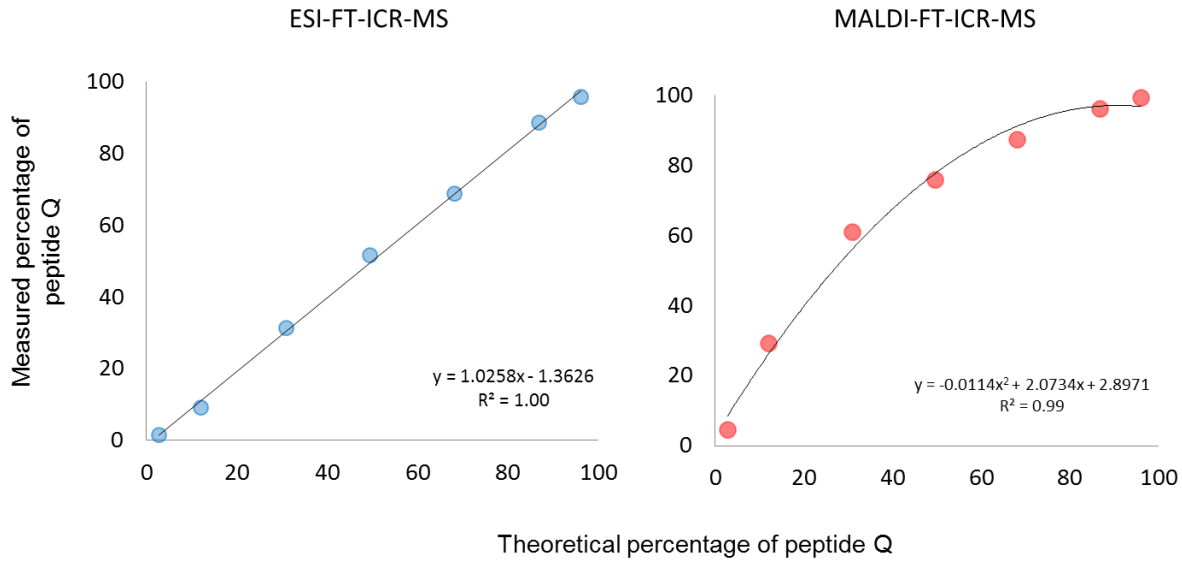


Figure 50: Plots of measured product peptide Q percentage against the theoretical ratios, determined using ESI (left) and MALDI (right). When using MALDI, there is an underestimation of product peptide E

4.4.2 TOF-MS Results

The peptide mixtures analysed in the previous sections were also analysed using MALDI-TOF-MS. The percentage of glutamine-containing peptide in the seven peptide mixtures was estimated using the genetic algorithm published by Wilson *et al.* (2012) (Figure 51). The differences between predicted and measured values obtained for the percentage of Q-containing peptide ranged from 2.51-34.78 % with the level of the product peptide E again, as with the FT-ICR measurements, consistently underestimated.

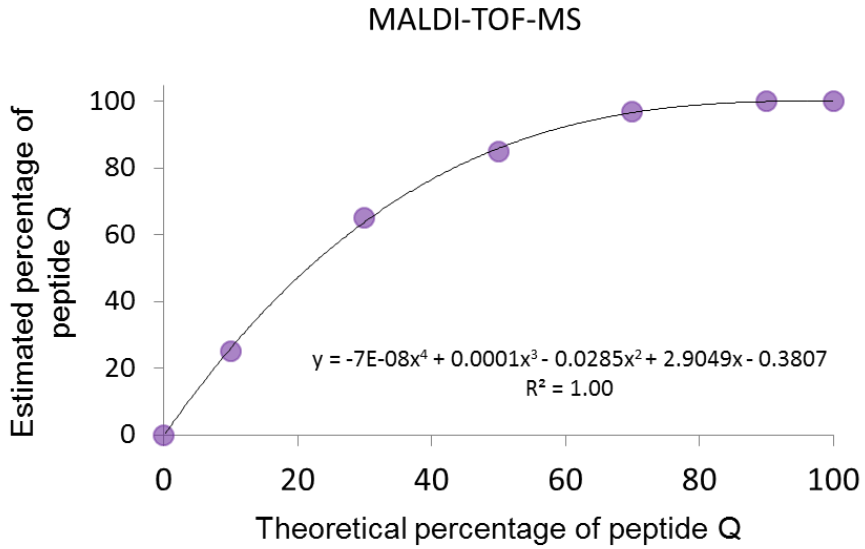


Figure 51: Ratios of Q to E containing peptides, measured using MALDI on a TOF mass spectrometer. Due to an over estimation of Q-containing peptide, a curvilinear relationship is observed

The percentages of measured product peptide Q using all three techniques are summarised in Table 6.

Table 7: The percentage of measured product peptide Q using three techniques: MALDI-TOF-MS, MALDI-FT-ICR-MS and ESI-FT-ICR-MS, as well as the theoretical percentages.

Theoretical percentage of Q	Percentage Q measured using MALDI-TOF-MS	Percentage difference between theoretical and measured value using MALDI-TOF-MS	Percentage Q mesaured using MALDI-FT-ICR-MS	Percentage difference between theoretical and measured value using MALDI - FT-ICR-MS	Percentage Q measured using ESI-FT-ICR-MS	Percentage difference between theoretical and measured value using ESI-FT-ICR-MS
2.51	0	2.51	5	2.49	1.44	1.08
12.20	25	12.80	29	16.80	9.11	3.09
31.35	65	33.65	61	29.65	31.32	0.03
50.22	85	34.78	76	25.78	51.39	1.16
68.81	97	28.19	87	18.19	68.78	0.03
87.13	100	12.87	96	8.87	88.44	1.31
96.19	100	3.81	99	2.81	95.58	0.60

4.5 Discussion

When deciding which ionisation source is most suitable for the samples in question, it is important to investigate relative mass spectrometric responses. This chapter investigates how well two product peptides, (one containing Q and one containing E) ionise in unseparated mixtures, in various ratios. When looking at ratios of deamidated to undeamidated forms of the same peptide, the deamidated peptide is 0.98 Da heavier in mass with respect to the deamidated form. There are several factors which have been shown to affect the ionisation of peptides such as the ratio of K:R (Lys:Arg) residues and the peptide's length. In the two product peptides investigated here the K/R ratios are the same; however factors such as the aliphatic index (Stapels and Barofsky, 2004) or hydropathy index (Kyte and Doolittle, 1982) may be altered by the presence of Q or E. In the study reported by Kyte and Doolittle (1982) glutamine and glutamic acid have the same hydropathy index of - 3.5. Stapels and Barofsky (2004) analysed a number of peptides using both ESI and MALDI. It was found that there were groups of peptides which were only observed when analysed using either ESI or MALDI. From their data set they identified that properties such as the isoelectric point of the peptide made little difference in the peptides' behaviour using the two ionisation methods. However, factors such as the K/R ratio of the peptides were found to be potentially important, with average values for peptides only seen by MALDI having a K/R ratio of 0.61 and peptides ionised by ESI having an average K/R ratio of 1.84. That study also showed that ESI preferentially ionises peptides containing amino acids which have hydroxyl or aliphatic side chains, with MALDI favouring

amino acids with basic or aromatic side chains. Among the peptides analysed by Stapels and Barofsky (2004), peptides containing seven different amino acids were found to have statistically significant differences in ionisation behaviour between ionisation sources and the number of times they were observed. These amino acids included glutamic acid and tryptophan, with peptides containing glutamic acid being preferentially ionised by ESI, and those containing tryptophan by MALDI. The product peptides analysed in this chapter also showed ionisation bias when analysed using MALDI, with product peptide E consistently underestimated. Analysis of the same mixtures however using ESI showed no ionisation bias between the two peptide products. When using ESI the full range of product peptide mixtures was measured accurately. This is consistent with the observations of Stapels and Barofsky (2004).

4.6 Conclusions

Two peptides, differing only in that one contains an internal Q residue, while the other had an E at this position, were successfully and cleanly synthesised using solid phase synthesis. The synthesis process did not appear to induce significant levels of deamidation in the Q-containing peptide, consistent with its reported relative resistance to deamidation (Geiger and Clarke, 1987). Analysis using LC-ESI-MS and ESI-FT-ICR-MS estimated that the product peptide Q contained 94.81% of peptide Q and 4.43% peptide E. From this analysis it is not possible to say whether the presence of small amounts of peptide E in the peptide Q product is due to deamidation or a consequence of impurity in the amino acid precursors used to synthesis the peptide.

It is clear that, when choosing an ionisation source, ESI may be more suitable than MALDI for measuring ratios of Q and E containing peptides. It has been suggested in the literature that both asparagine and aspartic acid containing peptides ionise similarly under the conditions of ESI, (Stroop, 2007; Hurtado and O'Connor, 2012); however (Stroop, 2007; Hurtado and O'Connor, 2012) did not test these assertions. The work carried out in this chapter has shown that when using MALDI, under-representation of product peptide E was observed across the full range of different percentage ratios. The level of underestimation of peptide E varied across the different percentages, so a constant correction factor is inappropriate. It should be noted that the peptide mixtures used in these experiments were relatively pure compared to those typical of biological extracts, with peptide Q estimated to be ~ 94.81 % and peptide E ~ 95.40 % pure. It has been shown here that even under these experimentally near optimal conditions, using reference material containing very few peptides, that preferential ionisation of product peptide Q occurs when using MALDI. Given this, measurement of the ratios of Q and E containing peptides using MALDI as an ionisation source, in complex mixtures such as those found tryptic digests of biological samples should be undertaken with caution. These findings are somewhat inconvenient as MALDI is a high throughput ionisation method, allowing the analysis of ~ 300 samples in ~ two hours,

resulting in shorter analysis time and lower analysis costs. This type of ionisation source coupled with TOF is also widely available and user friendly. As the data obtained from MALDI-TOF did show good non-linear correlation with the expected theoretical percentage, it may be possible to use this technique if reference material were available in which to create a calibration curve. Unfortunately when analysing complex biological mixtures, this is usually not the case. The data presented in this chapter suggests that ESI is a preferable ionisation source for the measurement of mixtures of Q and E-containing peptides, although further research is required to confirm this (see suggestions in Chapter 8). From the analysis carried out in this chapter it would appear that the TOF analyser maybe be a suitable, cheaper alternative to the high resolution FT-ICR mass analyser, however from the data presented in this chapter it is not possible to validate this as a ESI-TOF-MS was not used. The advantages and disadvantages of these two instruments are detailed below in Table 7.

Table 8: A comparison of the advantages and disadvantages of two mass spectrometry instrument systems (ESI-TOF-MS and ESI-FT-ICR-MS).

Advantages of MALDI-TOF-MS	Disadvantages of MALDI-TOF-MS	Advantages of FT-ICR-MS	Disadvantages of ESI-FT-ICR-MS
User friendly, especially for those not familiar with mass spectrometry.	TOF-MS is not able to resolve deamidated from undeamidated peaks and therefore a de-convolution algorithm is required.	High resolution instrument and can therefore separate out deamidated from undeamidated peptides. This allows the calculation of without a genetic algorithm and therefore less data processing required per sample.	Increased cost in analysis compared to TOF instruments, not only is the instrument more expensive, but the sample analysis time is longer (around 60 seconds per sample).
Quick analysis time (a few seconds per sample) with space for around 300 samples on one plate.	Multiple stages of data processing is required in order to calculate the level of deamidation in a peptide.	No observable ionisation bias between in Q and E containing peptide mixtures.	Less user friendly than the TOF-MS and requires more calibration steps to ensure the accuracy of the high resolution measurements.
Not much optimisation required to gain good quality spectra.	Ionisation basis when trying to measure mixtures of Q and E-containing samples.	When using ESI it is possible to get doubly charged ions which can sometimes increase the coverage of peptides, depending on the structure of peptides being analysed.	Larger sample amount required than MADI with around 50 µL of sample needed.
Small sample size required 1 µL per sample.			Once the sample has been injected and analysed it cannot be analysed again.
Once spotted on the MALDI plate, the dried spot can be reanalysed a number of times.			
Low cost per sample compared to higher resolution instruments.			

Chapter 6 - Kinetic experiments carried out on bone investigating the relationship between temperature and glutamine deamidation

Introduction

Numerous studies in experimentally degraded biological materials have shown that both glutamine and asparagine deamidation are temperature dependent (Sinex 1957; Scotchler and Robinson 1974; Stratton *et al.*, 2001; Robinson 2004) however few studies have investigated the rates of glutamine deamidation (Robinson *et al.,* 1970; Scotchler and Robinson 1974; Stratton *et al.,* 2001: Joshi and Kirsch 2004; Joshi *et al.,* 2005). In order to test whether glutamine deamidation is a proxy for age of bone material, it is important to understand the effect of temperature on glutamine deamidation.

This chapter details a number of experiments that were carried out, initially intended to investigate the relationship between glutamine deamidation and temperature. It became clear that optimising the experimental parameters was essential in order to properly test this relationship; therefore the optimisation of the method parameters was investigated. The first three experiments were carried out to determine the extent of deamidation in samples using different: temperatures, heating mediums, as well as the use of different extraction methods to isolate specific protein fractions (Table 8).

Table 9: The sample type, heading step and extraction methods used for the three heating experiments carried out on bone excavated from Tanner Row in this chapter

	Sample type	Heating temp	Heated in	Ambic extraction before demineralisation	Demineralised	Filtered	Ambic extraction after demineralisation
Experiment 1	bone powder	37 °C, 65 °C, 80 °C, 110 °C and 140 °C	ammonium bicarbonate buffer solution	heating solution analysed	no	no	n/a
Experiment 2	bone chip	80 °C,	damp sand	yes	half of chip demineralised using EDTA method and half using HCl method	yes	yes
Experiment 3	bone chips	65 °C	damp sand	yes	EDTA demineralisation method b	no	yes

5.1 Experiment 1 – Heating bone powder in solution

Experiment 1 (Figure 52) investigated the relationship between temperature and levels of glutamine deamidation in bone, when heated in solution. Slices of bone (Roman bovid metatarsal) were powdered and heated at different temperatures in ammonium bicarbonate solution (Table 8). Ammonium bicarbonate is used by van Doorn *et al.,* (2012) to extract soluble protein, (predominantly type I collagen), from mineralised bone. This is done by warming bone chips or powder in ammonium bicarbonate solution at 65 °C for one hour. It was therefore assumed that the experimental heating steps used here would successfully extract the collagen, with no need for a separate extraction step. The protein extracted during these heating experiments was analysed using mass spectrometry, targeting tryptic collagen peptides. The levels of glutamine deamidation in collagen peptides was measured as described in Chapter 2. The experimental steps are shown in Figure 46. Initially heating studies were carried out at five temperatures (37, 65, 80, 110 and 140 °C), over 24 hours. The two lowest temperatures were chosen as these were used during the sample preparation methods adapted from (Buckley and Collins 2011; van Doorn *et al.*, 2012) and are discussed in Chapter 3. The three higher temperatures were chosen to as it was hoped the data could be compared to previous long-term kinetic studies carried out by amino acid racemisation researchers (Penkman *et al.,* 2008; Demarchi *et al.*, 2013; Crisp *et al.,* 2013).

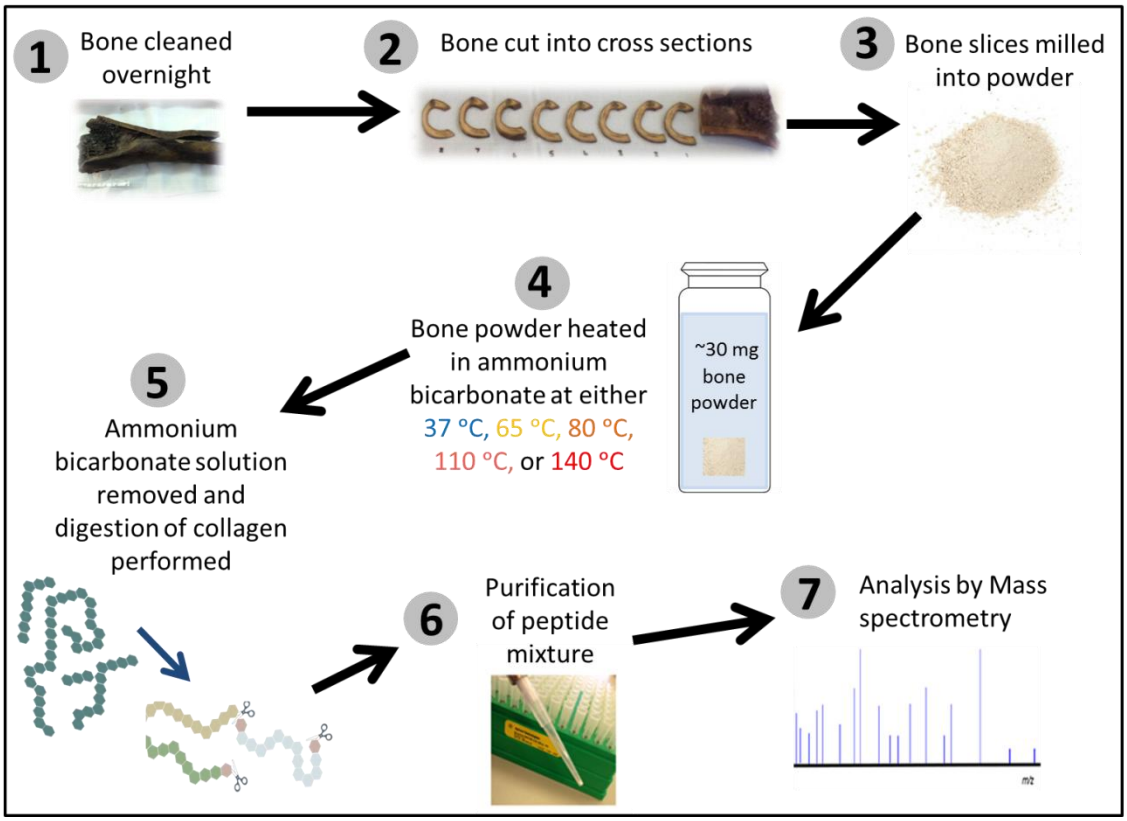


Figure 52: (1) The Roman bovid metatarsal bone was cleaned overnight in ammonium bicarbonate solution (pH 8.0) at room temperature ~20 °C. (2) The bone was then sliced in to thin sections ~ 3 in width using a diamond-edged water-cooled band saw. After slicing the bone, darker sections in the top centre of each of the slices were observed (Figure 47). These darker sections appeared macroscopically more degraded than the surrounding compact bone and were therefore removed using pliers. (3) The remaining macroscopically well-preserved slices were then milled under liquid nitrogen into a fine powder. (4) ~30 mg of powder was transferred to a glass vial, to this 100 µL of ammonium bicarbonate was added. The vials were sealed with screw tops and heated over 24 hours at either 37 °C, 65 °C, 80 °C, 110 °C or 140 °C. (5) After being removed from the ovens after heating, the samples were allowed to cool and vortexed for 60 seconds. 50 µL of the ammonium bicarbonate solution was transferred to a plastic (PP) microfuge tube and the extracted protein was digested overnight at 37 °C with trypsin. (6) The sample digests were then purified using ZipTips. (7) The resulting peptide mixtures were analysed using MALDI-TOF-MS and levels of deamidation for observed peptides containing glutamine were calculated

5.2 Methods

5.2.1 Sample preparation and heating

Six pieces of macroscopically well preserved bone slices prepared and described in Chapter 6 were milled under liquid nitrogen using a SPEX mill. After milling, the bone powder was dried in a fume hood for one week at room temperature. Each sample was prepared by transferring ~ 30 mg of bone powder to a sterilised 2 mL glass screw top vial (VWR International, WHEATON INDUSTRIES INC). To each vial 100 µL of 50 mM ammonium bicarbonate (pH 8.0) was added. The vials were then sealed with screw top lids. Samples were prepared in triplicate for each time point, resulting in a total of 120 samples. After heating the samples were allowed to cool and vortexed for 60

seconds. 50 µL of solution was removed and transferred to a 0.5 mL plastic (PP) microfuge tube. The resulting protein extract, consisting predominantly of type I collagen, was digested using trypsin overnight at 37 °C, filtered using ZipTips and analysed using mass spectrometry as described in Chapter 2. α- values for the observed peptides containing glutamine were then calculated as described in Chapter 2.

5.2.2 Results

This thesis has focused on 12 commonly observed glutamine-containing peptides, extracted from type I collagen. However in this experiment only one peptide (*m/z* 1105) was observed across all heated samples; therefore α-values for this peptide only are reported (Figure 53). At 37 °C, this peptide was detected in all 24 samples, and little to no deamidation observed over 24 hours, with an average α-value of 1.00. Only one of the 24 samples showed any deamidation (α = 0.98). At 65 °C peptide *m/z* 1105 was observed in 22 out of 24 samples, with a slight increase in deamidation to that observed at 37 °C. The average α-value over 24 hours was 0.96 with a range of 0.91-1.00 across the 24 samples. At 80 °C *m/z* 1105 was observed in 22 out of 24 samples heated over 24 hours, with an average α-value of 0.96 and a range of 0.87-1.00. At 110 °C the extent of deamidation significantly increased compared to the lower temperatures, with an average α-value at 0.5 hours of 0.97, and 0.85 after 8 hours. At 110 °C peptide *m/z* 1105 was not visible in 19 out of 24 samples, and undetected after 8 hours of heating. The average α- value in samples heated at 110 °C over 24 hours was 0.90, with a range of 0.91-1.00. At 140 °C peptide *m/z* 1105 was only observed in 12 out of the 24 samples, with no peptides observed after 2 hours of heating. High levels of deamidation were observed, with an average alpha value of 0.84 after 2 hours.

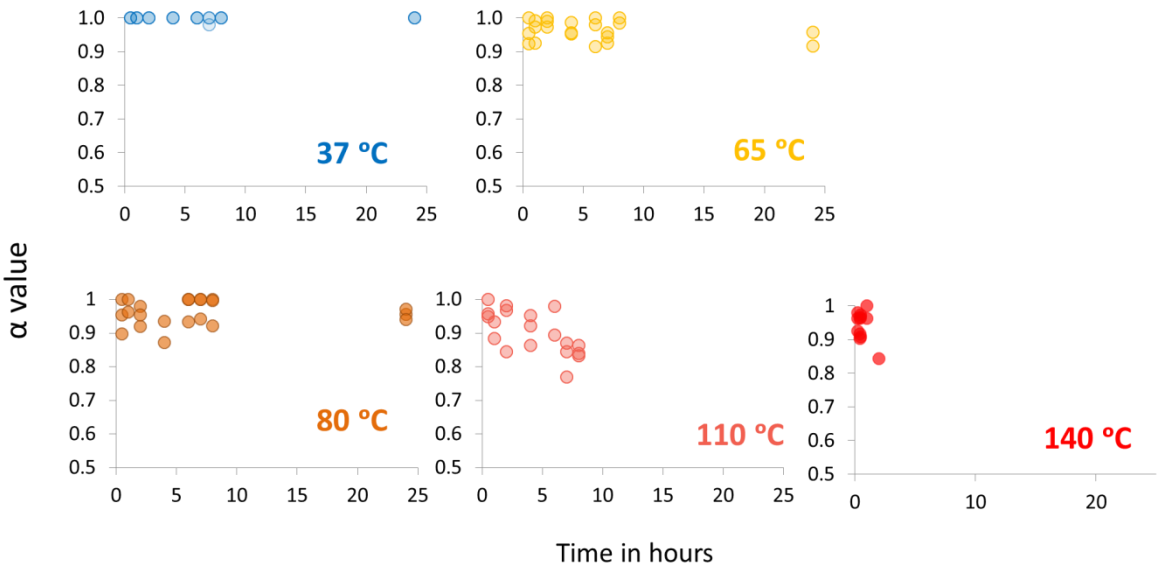


Figure 53: α-values measured in peptide *m/z* 1105 in protein extracts from Roman *bovid* bone powder. The solutions were heated at different temperatures over 24 hours. Samples heated at 37 °C showed little deamidation over this time period, with similar levels of deamidation observed at 65 °C and 80 °C. Greater levels of deamidation were observed at 110 °C and 140 °C, with the peptide undetectable after 8 hours at 110 °C and 2 hours at 140 °C

5.2.3 Conclusions

Although the samples heated in ammonium bicarbonate buffer solution did show a decrease in α-values for peptide *m/z* 1105 over time, with increased temperature; it is difficult to assess the variability of the α-values for samples heated above 65 °C, as this peptide was not always observed in all three triplicate samples. Equally for samples heated at lower temperatures such as 37 °C, where α-values were obtained in all samples, very little deamidation occurred, with only one sample showing deamidation. In samples heated at 65 °C, 80 °C, and 110 °C the standard deviation (SD) of α-values obtained for samples from the same time point increase with temperature, with average SD values of 0.021, 0.028 and 0.035 respectively. Where α-values were calculated for triplicate samples the α-values were relatively consistent. It should be noted however that measurements were only obtained for one peptide.

There was a large decrease in the number of peptides observed when samples were heated in solution above 65 °C, and therefore this method of heating is not suitable for long term kinetic studies of glutamine deamidation in collagen peptides. This may be due to peptide hydrolysis of the extracted gelatine whilst it is heated in solution. For collagen to be extracted from mineralised bone using only the gentle buffer extraction method, the collagen is likely to be in gelatine form, and therefore less structurally conserved. It is known that thermal, chemical or physical degredation of collagen causes the disruption of the triple-helical structure into random coils (gelatine) (Bigi *et al.,* 2004). If the soluble buffer soluble collagen is structurally compromised, it may be that this fraction of collagen is more susceptible to heat induced hydrolysis than the mineral-bound protein. The lack of peptides observed in the heated samples may also be due to peptide adsorption to the glass surface of the vials used, Kraut *et al.* (2009) found that poor reproducibility of LC-MS analysis of peptide mixtures was largely due to poor recovery of peptides from the glass sample vials. Although there has been little investigation into optimal heating conditions for kinetic studies of intact peptides, many studies have investigated the effects of vial type on peptide recovery, both during sample preparation stages, and for longer term storage of peptide samples. Nonspecific adsorption of peptides on to solid surfaces used during sample preparation is a known problem (Stejskal *et al.*, 2013), these solid surfaces include: pipette tips, plastic microfuge tubes, ZipTips, and glass autosample vials. Stejskal *et al.* (2013) reported that the vial in which the samples were stored did not appear to be as much of a problem as the sample solution in which the peptides were stored. However in the Stejskal *et al.,* study, where glass vials and plastic microfuge tubes were compared, the glass vials were pre-treated with polypropylene glycol. Kraut *et al.* (2009) carried out a study over 28 days measuring the recovery of peptides from three vial types (standard glass, standard plastic PP tubes, and low adsorption plastic tubes). They found that glass tubes preformed better (in terms of peptide recovery) than standard plastic tubes, with hydrophobic peptides being preferentialy 'lost' when using plastic. Kraut *et al.* (2009) also compared the recovery of peptides from glass, lo-adsorbtion plastic tubes and standard plastic

tubes. They found that glass and standard plastic tubes did not perform as well as the low-adsorbtion platic tubes.

Therefore in order to reduce the contact of the sample with glass, and minimise the extraction of soluble collagen during the heating step, further heating experiments were carried out in damp sand (Experiments Sections 2-4). It was hoped that this would help maximise the amount of protein extracted from the bone after prolonged heating.

5.3 Experiment 2 – Heating of bone chips at 80 °C in damp sand, demineralisation treatments and ultrafiltration of heat-treated collagen

It is inevitable that during heating some protein will be lost from the bone due to leaching, but when bone samples are heated in solution (e.g. Experiment 1), the extraction of soluble collagen occurs during the heating step. By heating in damp sand, the amount of soluble protein which is extracted from the bone during heating should be less than when heated in solution. The use of damp sand is also likely to be more representative of the general burial environment, than liquid storage such as vials containing water, or buffer solution. . However, other factors need to be considered:

It may be that after heating for long periods of time the collagen in the buffer soluble fraction is broken down, unlike the more structurally conserved mineral bound collagen; if this is true then in order to extract collagen from heated samples, a demineralisation step will need to be performed. In Chapter 6, HCl and EDTA demineralisation methods were compared, both in terms of their effect on levels of deamidation, and on the structure of collagen fibrils. It was found that in unheated bone, HCl induced more deamidation than EDTA and resulted in more disruption of the collagen fibril structure. It was also found that collagen in older bone was more susceptible to HCl-induced damage. As heating has been used to artificially age bone, it may be that heated bone is also more susceptible to HCl-induced damage, so HCl may not be suitable for the demineralisation and extraction of collagen from heated bone.

After demineralisation it is standard practice to gelatinise and filter samples using an Amicon filter, (for example when measuring isotope ratios of collagen or prior to radiocarbon dating (Higham, Jacobi and Ramsey *et al.* 2008)). Theoretically the collagen which does not pass through a 30 kDa filter should be more structurally intact than that which passes through to the filtrate, enabling isolation of the well preserved collagen fraction from potentially degraded, contaminated collagen or smaller proteins.

Therefore in this experiment, four factors were tested:

- Is damp sand a better medium for heating bone chips, prior to extracting the buffer-soluble collagen fraction using the ammonium bicarbonate extraction method?
- Is the quality of the data improved by demineralising the heated bone chip using HCl?
- Is the quality of the data improved by demineralising the heated bone chip using EDTA?
- Does the level of deamidation differ depending on what MW fraction is collected from the Amicon filter after the demineralisation/gelatinisation process?

The experimental steps for this experiment are shown in Figure 48. In order to test if damp sand is a better medium in which to heat bone chips prior to protein extraction, chips were heated in damp sand at 80 °C for up to 21 days. After heating the chips were removed from the sand and an ammonium bicarbonate extraction (described in Chapter 2) was performed on the bone chips to analyse the buffer-soluble collagen. The chips were then demineralised either using HCl or EDTA. The demineralised samples were gelatinised and the protein filtered and separated into three fractions by molecular weight using two different sized Amicon ultra filters (Figure 54). The resulting protein fractions were digested, purified and analysed by mass spectrometry as described in Chapter 6 (sections 6.1.5.1 and 6.1.5.2).

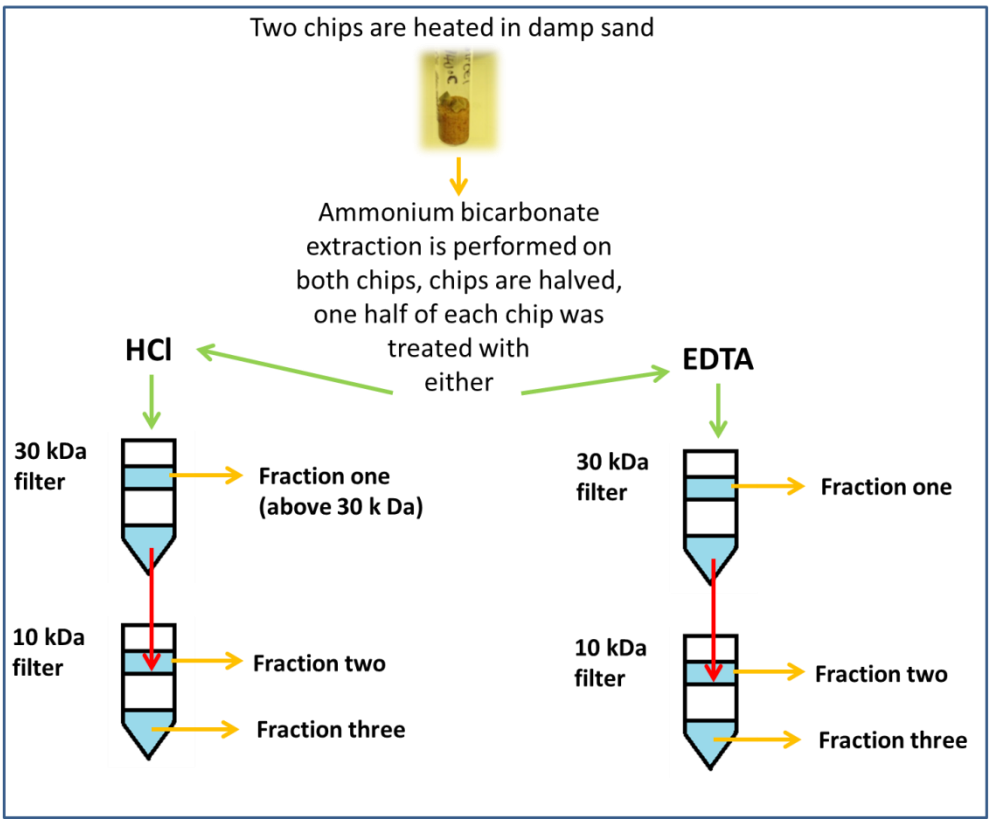


Figure 54: Schematic of steps for Experiment 2. First two chips were heated in sealed glass ampoules containing 2 g of damp sand. The vials were heated at 80 °C for 1, 24, 48, 168 and 504 hours. After heating the chips were removed from the glass ampoules and an ammonium bicarbonate extraction was performed on each chip. Two chips were then taken and broken in half. The first half of each chip was then demineralised using HCl and the second half using EDTA. The demineralised samples were gelatinised overnight at 80 °C in water adjusted to pH 3.0. Once gelatinised, each sample was filtered through a 30 kDa Amicon filter. The retentate was collected and labelled as fraction one (peptides with mass ≥ 30 kDa). The filtrate was then passed through a 10 kDa Amicon filter. The retentate was collected and labelled as fraction two (peptides with mass <30 kDa and > 10 kDa). The filtrate was collected and labelled as fraction three (peptide with mass <10 kDa. The subsequent protein extractions were then transferred to a plastic (PP) microfuge tube and digested overnight at 37 °C with trypsin. The sample digests were then purified using a ZipTip and the resulting peptide mixtures were analysed using MALDI-TOF-MS and levels of deamidation for observed peptides containing glutamine were calculated

5.3.1 Methods

5.3.1.1 Heating and extraction of protein from bone chips

The sand (play sand purchased from B&Q) and glass ampoules (Scientific Laboratory Supplies Ltd) were both sterilised overnight by heating at 400 °C. To each glass ampoule ~ 1 g of sand was added, followed by three bone chips weighing ~30 mg each. The chips were then covered with a further ~ 1 g of sand, and 100 µL of purified water was added. The ampoules were flame sealed with a blow torch (BLOW LAMP, TC2000 WITH CARTRIDGE; Temperature Max: 1750 °C, Onecall, CAMPINGAZ) and transferred to an oven set at 80 °C. The ampoules were heated for 1, 24, 48, 168

and 504 hours. After heating, samples were removed from the oven and allowed to cool. The glass ampoules were then broken open and left to dry in a fume hood at room temperature for 24 hours in order to make recovery of the chips from the glass vials easier. The chips were then removed from the sand and transferred to three 0.5 mL plastic (PP) microfuge tubes. Ammonium bicarbonate extractions were performed on all three chips. Two chips out of the three were then selected at random (due to cost implications). Each of the two chips were broken in half. The first half was demineralised with HCl as described in Chapter 6, section 6.1.5.1. The second half of the chip was demineralised using EDTA as described in Chapter 6 section 6.1.5.2. After gelatinisation, each sample was filtered through a 30 kDa Amicon filter. The retentate was collected (fraction one, peptides with a mass >30 kDa) and the filtrate was then collected, and filtered a second time though a smaller 10 kDa Amicon filter (fraction two). The retentates collected contained peptides with mass of < 30 kDa and > 10 kDa. The filtrate was then collected (fraction three, peptides with a mass of ≥ 10 kDa). All of the fractions were freeze dried overnight and resuspended at a concentration of ~ 2mg/mL in 50 mM ammonium bicarbonate.

The protein extracts were then digested overnight at 37 °C using trypsin, and purified on ZipTips before being analysed by MALDI-TOF-MS. Levels of glutamine deamidation were then measured and reported as α-values, as described in Chapter 2.

5.3.2 Results

Ammonium bicarbonate extractions of mineralised chips heated in sand resulted in higher numbers of detectable Gln-containing peptides than those heated in ammonium bicarbonate solution, with 60 % of the 288 possible Gln-containing peptides observed in samples heated over 21 days at 80 °C. The differences in levels of deamidation observed between sample extraction methods can be split into three categories: (1) there was little difference in measured α-values regardless of the pre-treatment/extraction methods applied (e.g. *m/z* 1105, Figure 55). (2) α-values measured were higher in both the EDTA demineralised and the ammonium bicarbonate extractions of mineralised samples, than those demineralised with HCl (e.g. *m/z* 2705 5, Figure 55). (3) α-values measured were highly variable, with α- values in samples in ammonium bicarbonate extractions of mineralised bone being higher than those observed in both of the demineralised sample sets. (e.g. peptide *m/z* 3100, Figure 55). Out of all of the peptides measured in this experiment, peptide with *m/z* 3100 not only showed the highest level of variability, but is also the only peptide where α-values appeared to be lower in fractions two and three than in fraction one, in both HCl and EDTA treated samples. This may indicate this particular peptide is more susceptible to deamidation once the protein has broken down into smaller pieces.

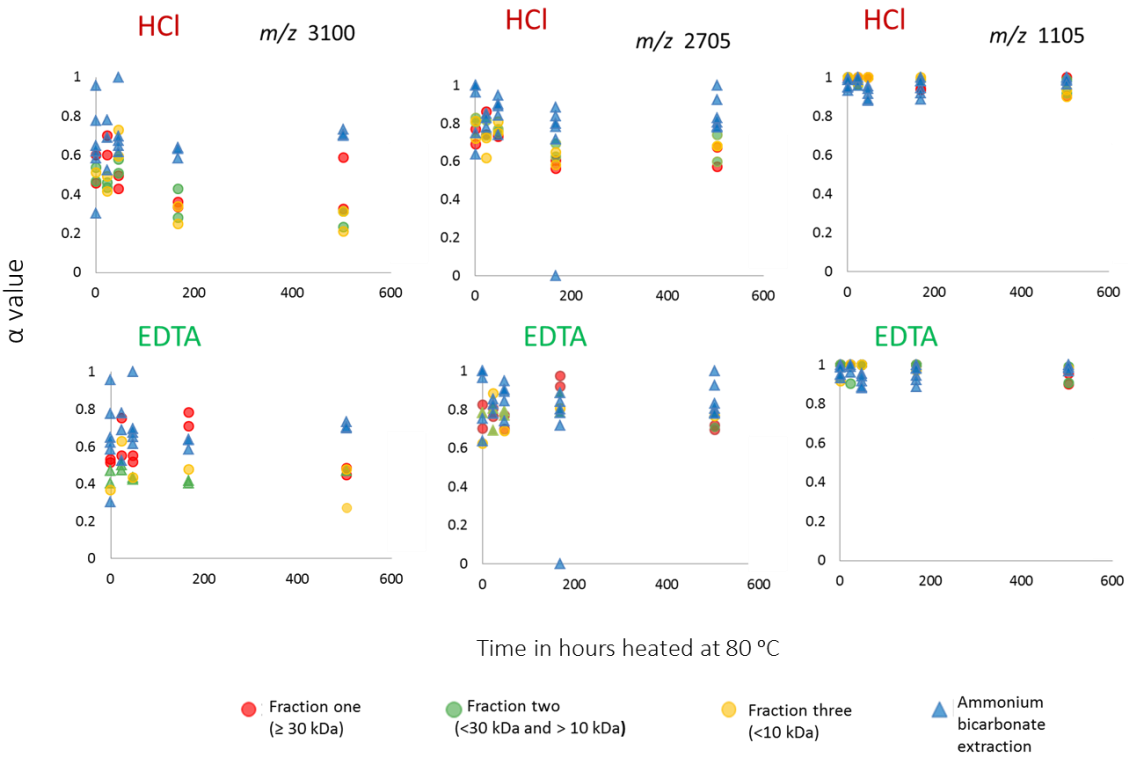


Figure 55: Levels of deamidation measured in collagen tryptic peptides, extracted from bone, heated at 80 °C, across four collagen fractions: ammonium bicarbonate extraction of mineralised bone and > 30 kDa, 10-30 kDa, <10 kDa), fractions of samples demineralised in either HCl (top) or EDTA (bottom)

5.3.3 Conclusions

Samples heated in damp sand and extracted without demineralisation (using the ammonium bicarbonate method) resulted in better quality spectra in terms of the number and S/N of Gln-containing peptides observed, than samples heated over time in ammonium bicarbonate solution. The use of Amicon filters to separate the protein fractions by MW did not appear to have an effect on the alpha value, except in peptide m/z 3100, where fraction one produced generally higher α-values than those in fractions two and three. The method of demineralisation appears to effect levels of deamidation in some peptides: peptides with *m/z* values of 2056, 2072, 2088 and 3100 all showed lower α-values in HCl-treated samples vs samples either demineralised using EDTA or ammonium bicarbonate extracts from mineralised bone. Overall the levels of deamidation measured in samples treated with EDTA appeared similar to those observed in undemineralised samples extracted with ammonium bicarbonate.

This is consistent with the preliminary investigation of extraction methods carried out in Chapter 6; where HCl was found not only to increase deamidation, but also to disrupt the collagen fibril structure. It has been shown here that although demineralisation may be useful when extracting protein from heated samples, the use of HCl is not suitable when extracting protein to calculate

levels of deamidation. When conducting heating experiments on porous mineralised protein such as bone, damp sand appears to be a more suitable heating medium than solutions such as ammonium bicarbonate. Although the use of filtration using Amicon filters may be preferable for some methods of collagen analysis, such as the measurements of isotopes (either for dietary information or radiocarbon dating), it does not appear to make a difference to the level of deamidation. The use of Amicon filters during sample preparation is time consuming, requiring multiple washing, and centrifuging steps, on average adding ~ 2 hours to the sample preparation time per 16 samples. As well as additional time, Amicon filters are also expensive, with a box of 24 filters costing around £70 from Merck Millipore. As filtering samples post-demineralisation is not advantageous when preparing protein for deamidation measurements, future experiments using demineralisation methods will substitute this filtration step with washing steps (See adapted EDTA method Chapter 5).

5.4 Experiment 3 - Levels of glutamine deamidation in samples heated in damp sand, at 65 °C over 75 weeks.

Previous experiments in this chapter showed that as the heating temperature of the experiment increased, fewer peptides were observed. It was also observed that heating in sand was preferable to heating in solution. Although ammonium bicarbonate extraction of mineralised bone does not appear to induce deamidation, it is not always successful when extracting protein from aged or heated bone. It has also been shown in Chapter 6 that EDTA may be a suitable alternative extraction method to HCl. In order to further explore the relationship between temperature and deamidation, a long term (75 week) low temperature (65 °C) experiment was carried out, comparing both ammonium bicarbonate extraction of mineralised bone to protein extraction post EDTA demineralisation. The EDTA demineralisation method was adapted to remove the filtration step; this was replaced by a number of washing steps (Described fully in Section 5.3.1.3). Samples for each time point were prepared in triplicate. After heating, the ammonium bicarbonate extraction was performed on all of the mineralised chips. A second extraction was then performed on all chips using the adapted EDTA demineralisation method (Section 5.3.1.3), and levels of deamidation from these two extractions compared. The experimental steps are shown in Figure 56.

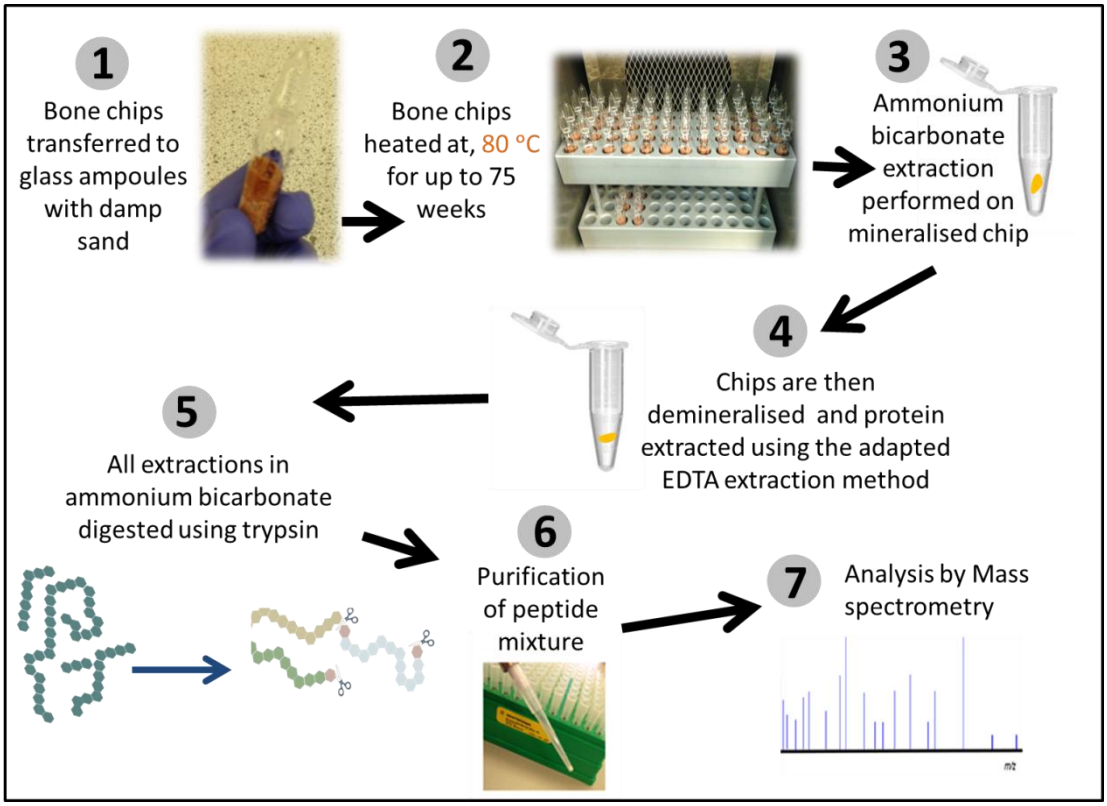


Figure 56: Schematic of experimental steps. (1) Bone chips were heated in sealed glass ampoules containing 2 g of damp sand. (2) The vials were heated at 65 °C over 75 weeks. (3) After heating the chips were removed from the glass ampoules and an ammonium bicarbonate extraction was performed on each chip. (4) The chips were then transferred to separate microfuge tubes and EDTA solution added (pH 7.4), samples were demineralised, rinsed with ammonium bicarbonate solution, and the ammonium bicarbonate extraction preformed. (5) 50 µL of the ammonium bicarbonate solution was transferred to a plastic (PP) microfuge tube and the extracted protein was digested overnight at 37 °C with trypsin. (6) The sample digests were then purified using ZipTips. (7) The resulting peptide mixtures were analysed using MALDI-TOF-MS and levels of deamidation for observed peptides containing glutamine were calculated

5.4.1 Methods

5.4.1.1 Heating of samples

To a sterile glass ampoule 1 g of sand was added, and then three bone chips weighing approximately 30 mg each. The chips were covered with a further 1 g of sand and 100 µL of water was added. The vials were flame sealed using a small blow torch before being transferred to the oven at 65 °C for up to ~75 weeks (time in weeks, 28, 32.9, 37.9, 44, 46, 51, 59, 69.9, 75.6).

5.4.1.2 Protein extraction using ammonium bicarbonate

After heating two ammonium bicarbonate extractions were performed; one pre (as described in Chapter 2) and one post-demineralisation (Section 5.3.1.3). The subsequent Protein extracts were digested and analysed as described in Chapter 2.

5.4.1.3 Adapted EDTA Demineralisation Method

The EDTA demineralisation protocol was adapted from Koon *et al.* (2012). Each chip was placed in a 1 mL polypropylene centrifuge tube and 1 mL of 0.5 M EDTA (pH 7.4) added. The samples were stored at room temperature, and the EDTA solution was replaced every three days. After approx. 14 days the samples did not appear to be fully demineralised. In order to try to optimise the amount of protein extracted, the length of demineralisation was extended. The samples were transferred to the freezer and stored at ~-20 for approximately 6 weeks. After 6 weeks the samples were removed from the freezer and vortexed for 30 seconds. The samples were then spun at x 300 g for 5 minutes and the supernatant removed. The samples were then washed by adding 1 mL of 50 mM ammonium bicarbonate to each tube. The samples were vortexed for 30 seconds and spun at x 300 g for 5 minutes. This washing step was repeated three times. After the last washing step the sample was resuspended in 50 µL of 50 mM ammonium bicarbonate and warmed at 65 °C for one hour. The protein extracts were then digested and analysed as described in Chapter 2.

5.4.2 Results

In protein extracted before the demineralisation step, only 9 % of the possible Gln-containing peptides were observed. In these samples a signal corresponding to peptide with *m/z 1105* was observed in at least half of the samples. In samples extracted post-demineralisation, the percentage of observed peptides increased to 80 %, with four peptides (*m/z* 836.43, *m/z* 1105.57, *m/z* 2705.20 and *m/z* 3665.80) observed in all samples, five peptides (*m/z* 1690.77, *m/z* 1706.77, *m/z* 2056.94, *m/z* 2088.95, and *m/z* 3100.53) observed in at least one of the replicates at each time point. The remaining three peptides (*m/z* 2702.97, *m/z* 2689.20 and *m/z* 3001.47) were observed in at least one of the replicates in eight out of the 9 time points measured. In peptide *m/z* 1105.57 samples extracted using only ammonium bicarbonate (pre EDTA treatment) were found to have increasing levels of deamidation over time (Figure 57, yellow dots), whereas samples extracted post EDTA-treatment showed relatively lower amounts of deamidation (Figure 57, blue dots).

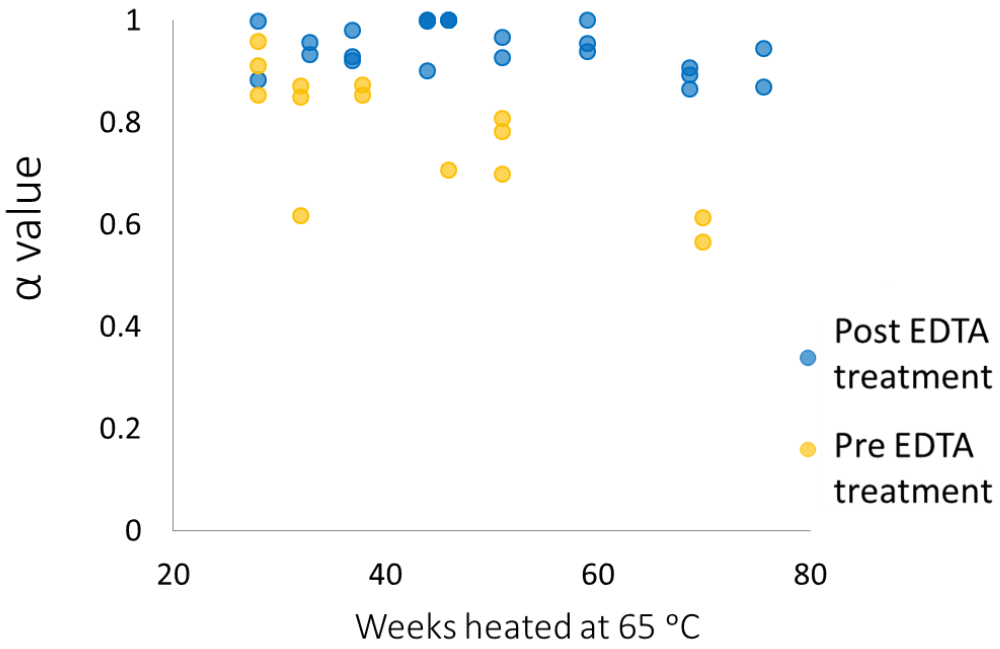


Figure 57: α-values calculated in peptide *m/z* 1105.57, in collagen digests, extracted from bone, heated at 65 °C over 80 weeks, before and after EDTA treatment.

5.5 Discussion

Differences in both peptide coverage and levels of deamidation were observed in samples analysed pre and post EDTA-demineralisation, with peptide coverage and spectra quality greatly improved in samples treated with EDTA. This is most likely due to the highly efficient demineralisation of the samples which was performed slowly at low temperature. Peptide *m/z* 1105 was observed in both sample sets and showed differences in the extent of deamidation between the two sample treatment methods; samples extracted pre-EDTA treatment showed increased levels of deamidation over time. This is most likely due to differences in structure of the protein within the soluble and mineral bound fractions. It is likely that the buffer-soluble collagen fraction exists as random coil or gelatine like structures, whereas the mineral-bound collagen is likely to be still highly ordered and helical. When carrying out heating experiments of intact protein, a balance is needed. The heating experiment needs to be gentle enough so that the protein remains intact, but also there needs to be enough of a temperature increase to accelerate the reaction being measured, in this case glutamine deamidation, over a feasible time scale. Although peptide *m/z* 1105.57 measured in soluble protein extracted pre-demineralisation does appear to undergo deamidation over time at 65 °C, this fraction is not easily extracted after heating. This fraction is also not always successfully extracted from archaeological or sub fossil bone. In addition it should be noted that the composition of this soluble fraction is likely to change over time, as the bone degrades, as well as being susceptible to exchange with the burial environment. For these reasons it may not be directly representative of the preservation state of the bone as a whole.

The mineral bound protein which is extracted post-demineralisation is likely to be more representative of the overall preservation state of the bone; however deamidation of glutamine in this collagen fraction occurs very slowly at 65 °C, with little deamidation observed even after 80 weeks of heating. In order to further investigate the relationship between time and temperature in this collagen fraction, experiments using higher temperatures are required, as well as the measurement of this fraction in dated Pleistocene material. It may be that glutamine in this collagen fraction is too stable to be used as a proxy for thermal age estimations.

5.6 Overall conclusions

Initially the aim of these experiments was to calculate rates of glutamine deamidation in type I protein extracted from bone, at different temperatures. After the preliminary investigations, it became apparent that the effect of heating on the bone and protein was complex, and so the study focused on optimising the experimental design. A series of experiments were required in order to optimise the methods used for both the heating step, and the successful removal and analysis of the subsequent collagen peptides. It was found that heating in solution resulted in the extraction of very few glutamine containing collagen peptides, especially at temperatures above 65 °C. The use of HCl pre-treatment also reduced the number of extracted glutamine containing peptides and increased levels of deamidation. In the experiments conducted in this chapter the use of Amicon filters did not improve the extraction process, or decrease the variability of the calculated α-values and it is therefore suggested that the use of Amicon filters is not required when preparing collagen for deamidation measurements. It is recommended that future heating experiments for this type of analysis are carried out in damp sand, followed by EDTA pre-treatment (Section 5.3.1.3) prior to ammonium bicarbonate extraction.

Chapter 6 - The effects of demineralisation and sampling point variability on the measurement of glutamine deamidation in type I Protein extracted from bone

This chapter is presented as a paper which has been accepted to the Journal of Archaeological Science. The section, figure and table numbers have been renumbered in order maintain the continuity of the thesis. The contributions of the authors are as follows. The Experimental design, sample analysis and data interpretation was carried out by J P Simpson. The statistical analysis was carried out by J Wilson. Samples were provided by B Shapiro. H Koon and Meg stark helped with training on the TEM preparation and analysis. J Thomas-Oates gave guidance on the obtained mass spectra. K E H Penkman and Matthew Collins offered advice and made corrections to the final version of the paper.

J P Simpson[1], K E H Penkman[1], B Demarchi[2], H Koon[6], M J Collins[2], J Thomas-Oates[1,3], B Shapiro[7], M Stark[4], J Wilson[1,5]

[1]Department of Chemistry, University of York, York

[2]Department of Archaeology, University of York, York

[3]Centre of Excellence in Mass Spectrometry, University of York, York

[4]Department of Biology, University of York, York

[5]Department of Mathematics, University of York, York

[6] Department of Archaeological Sciences, University of Bradford, Bradford

[7] Department of Ecology and Evolutionary Biology, University of California Santa Cruz, California

Introduction

Bone can survive in the burial environment for millions of years (Collins *et al.,* 1995) and can provide *direct* information about the individuals during their life and *post mortem*. Bone contains both organic (mainly proteins) and inorganic components, with the most abundant protein being type I collagen. (Rich and Crick, 1961)This fibrous protein consists of three polypeptide chains of similar length (two α-1 chains and one α-2 chain) that form a tightly-wound triple helix (Whitford 2008; Shoulders *et al.*, 2009; Viguet-Carrin *et al,.* 2006). The presence of the hydroxyapatite (mineral) crystals, which embed and protect the protein, contribute to the stability and preservation of bone over geological timescales (Turner-Walker 2008; Covington *et al.*, 2010).

The extraordinary preservation of collagen in bone has been exploited by archaeologists and palaeontologists seeking to address challenges such as species identification (Buckley *et al.,* 2009;

Welker *et. al.,* (2015)), diet (Ambrose and Norr, 1993) and radiocarbon age (Libby 1960; Reimer *et al.,* 2013). Recently, the radiocarbon dating of single amino acids such as hydroxyproline (McCullagh *et al.*, 2010; Marom *et al.,* 2012) and improved pre-treatment methods (Brock *et al.,* 2007; Ramsey and Higham 2007; Brock *et al.,* 2010) have enabled radiocarbon dating to be applied to samples as old as ~ 50 ka BP (van der Plicht and Palstra 2014). However, bones recovered from Middle and Lower Palaeolithic and palaeontological sites must be dated by association with other materials, which can be used as substrates for other absolute dating methods (e.g. luminescence or U-series). Therefore a method that could date bone material directly would be a valuable tool.

Collagen could be an ideal substrate for dating because it has extraordinary potential to be preserved in the fossil record. It was predicted that collagen could survive up to 500,000 years in optimal (i.e. cold) burial conditions (Collins *et al.*, 1995); it has since been found that, even in temperate environments (e.g. in Europe), collagen can survive for much longer than this, up to 1.5 million years (Buckley and Collins, 2011). However, the extent of degradation of collagen increases with thermal age (Dobberstein *et al.*, 2009; Smith *et al.,* 2003), which is defined as an estimate of the equivalent age based upon thermal history, assuming the sample had been held at constant temperature -10 °C (www.thermal-age.eu). A relationship has been suggested between the thermal age and the level of glutamine deamidation (derived from composite estimates of deamidation in several peptides) observed in extracted bone collagen (van Doorn *et al.,* 2012; Wilson *et al.,* 2012). Given the difficulties of using amino acid racemization dating (AAR) to provide robust age information on collagen (Bada and Helfman 1975), such a link could provide the key to age estimation for bone samples beyond the range of ^{14}C dating. However, the data reported by van Doorn *et al.,* (2012) showed high variability (ranging from 40% to 90%) in the levels of glutamine deamidation in peptides extracted and analysed from bones of the same age, obtained from the same site.

Here, we explore the potential causes of this variation, and we test two hypotheses: 1) that variation may occur due to natural variability within the biological tissue; and 2) that variation may be induced in the laboratory, during sample preparation. First, we perform a series of experiments that focus on preservation and decay of a single, well-preserved bovine metatarsus of Roman age. From this bone we determine the variability of glutamine deamidation using mass spectrometry (MS) as a function of:

- the location within the bone from which the sample was taken (section 6.1);
- the visible preservation of the bone - comparing degraded and non-degraded sections (Section 6.2);
- demineralisation method - comparing the effects of two demineralisation methods (using hydrochloric acid (HCl) and ethylenediaminetetraacetic acid (EDTA)) on the levels of deamidation (Sections 6.3 and 6.4).

Second, we explore the preservation of collagen fibrils in samples of different ages, when demineralised using either HCl or EDTA. This was done using TEM to visualise three bones that differ considerably in age: modern, Roman (bone used in previous sections), and Pleistocene (~80,000 years old) (Section 6.5).

Our aim is to improve the understanding of the effects that sample location and pre-treatment methods may have on collagen preservation. This will allow not only more accurate determination of the extent of deamidation in bone collagen, but also may be useful for other analytical methods that require the removal of mineral, such as radiocarbon dating, isotopic analysis or species identification through collagen mass finger printing (ZooMS). Our results provide data that are key to the appropriate interpretation and exploitation of the suggested relationship between deamidation levels and diagenetic history.

6.1 Methods

An overall schematic of the process we have used for the preparation, extraction and analysis of collagen by mass spectrometry is shown in Figure 58.

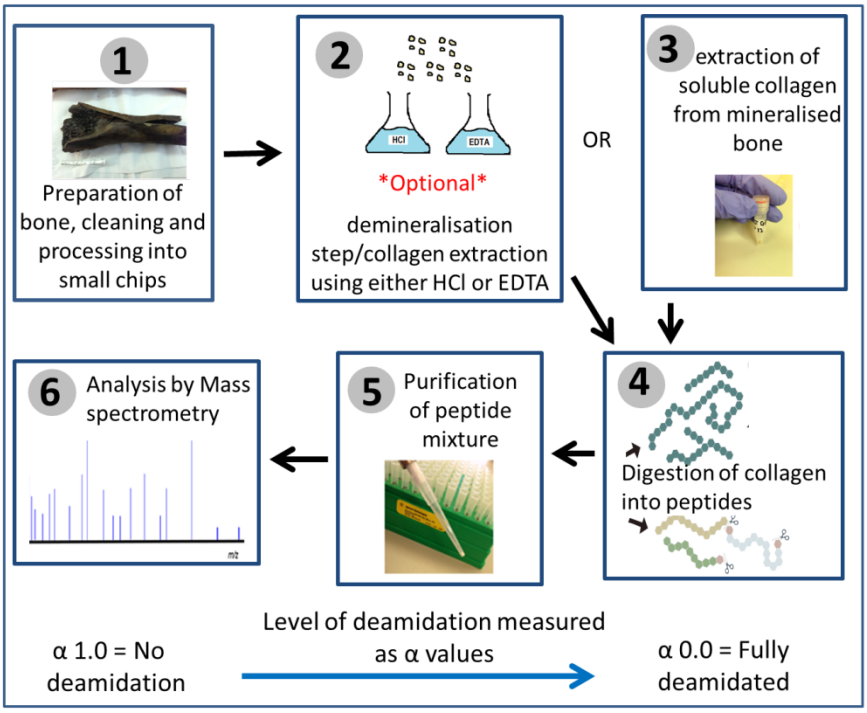


Figure 58: A schematic of sample preparation protocols. (1) Samples are cleaned in 50 mM ammonium bicarbonate at room temperature overnight. The sample is then cut into small pieces as required; (2) For the demineralisation experiments, the bone is demineralised using either HCl or EDTA, gelatinised, ultrafiltered, freeze dried and the resulting lyophilised collagen is re-suspended in ammonium bicarbonate solution (3) If step two has not been performed then collagen is extracted directly from the mineralised bone by warming in ammonium bicarbonate solution (at 65 °C) for one hour; (4) A tryptic digestion of the extracted protein is carried out overnight in ammonium bicarbonate solution at 37 °C ; (5) The resulting peptide mixture is purified using solid phase ZipTips; (6) the peptide mixture is analysed by MALDI-MS (section 2.5); the spectrum is used to estimate the level of deamidation occurring in specific peptides (section 2.6). The calculated glutamine deamidation level is given by the α-value, with a value of 1.0 representing no deamidation and 0.0 indicating complete deamidation of glutamine to glutamic acid.

6.1.1 Preparation and cleaning of bone samples

All three bone sample types (modern, Roman and Pleistocene) were cleaned at room temperature (~22 ºC) by soaking in 50 mM ammonium bicarbonate solution (pH 8.0, prepared in purified water, 18.0 MΩ) overnight. After cleaning, the bones were allowed to dry in a fume hood at room temperature.

6.1.2 Tanner Row bovine bone

The main sample used in this analysis was a bovine bone (Figure 29) from the site of Tanner Row (23/36). Tanner Row is thought to be a Roman site and was excavated in 1984 by York Archaeological Trust. Excavations revealed a number of organic remains including a large number of cattle bones thought to date around the late 2nd century. A piece of bovid bone (Figure 59) from this site has been used throughout this thesis for experimental investigations on the effects of different aspects of the sample preparation process. This bone was sub-sampled first by slicing into 17 cross sections; some of these cross sections were then further sub-sampled by breaking parts of them into small chips. Because deamidation may be induced thermally (van Doorn *et al.*, 2012), after cleaning (see Section 2.1), the bone was cut into 17 slices (~ 3 mm in width) using a diamond-edged water-cooled band saw (Figure 30). The separate slices were then cleaned in 50 mM ammonium bicarbonate solution and left to dry for one week in a fume hood at room temperature.

After slicing the bone, darker sections in the top centre of each of the slices were observed (Figure 30). These sections of bone were darker than the surrounding areas, with a dark brown appearance. In addition, the sections classified as degraded also appeared to be more porous, with small circular holes throughout the thickness of the bone. The remaining areas of bone were a pale brown/yellow colour, with no observable surface erosion. These darker sections have therefore been classified as macroscopically more degraded than the surrounding compact bone, and were removed using pliers before further analysis. The remaining pieces of each slice were immersed in liquid nitrogen for 60 seconds and then removed and broken into small chips using a small impacting hammer; the chips were then sieved though a 2 µm metal sieve and the chips of more than 2 µm were rinsed in purified water and subjected to a range of different protein extraction procedures (Chapter 2, Sections 2.2, 2.3 and 2.4).

Figure 59: Cross-section of cortical bone from the Tanner row site, sliced using a water cooled band saw, showing a darkened porous area of bone

6.1.3 Pleistocene bone

A fragment of bison metapodial bone excavated from a permafrost site in the Klondike region of Canada's Yukon Territory was investigated. This bone was AMS radiocarbon dated at the Center for Accelerator Mass Spectrometry, Lawrence Livermore National Laboratory, California USA, which provided in a non-age estimate (>50,300 ^{14}C years BP; CAMS 157517). This sample was found in association with a volcanic ash (tephra) layer, Sheep Creek-K, that has been dated to ~80,000 years old (Westgate *et al.* 2008). As the exact age of this sample is unknown, we refer to this sample throughout this paper as Pleistocene in age. The bone piece was cleaned prior to all analyses as described in Section 2.1.

6.1.4 Modern bone

A piece of modern bovine tibia obtained from a local butcher (Newcastle) was prepared by Dr C. Smith (Smith *et al.,* 2005): the periosteum and marrow were removed with a scalpel and the bone was then sawn into chunks and defatted for 24 h in acetone. The chunks were freezer-milled under liquid nitrogen.

6.1.5 Extraction of collagen from mineralised bone using ammonium bicarbonate

50 mM ammonium bicarbonate (pH 8) was added to each sample (approximately 100 µL per 30 mg of bone). The sample was then warmed for one hour at 65 °C (adapting extraction procedures described in van Doorn *et al.,* 2011).

6.1.5.1 Hydrochloric acid demineralisation/Protein extraction

For demineralisation in hydrochloric acid (HCl) the standard preparation protocol for stable isotope analyses of Ambrose (1990) was adapted: each chip was placed in a 15 mL polypropylene centrifuge tube and 5 mL of 0.6 M HCl (pH 1) added. The samples were stored at 2 – 8 °C and the HCl replaced every three days. After 10 days the samples appeared to be visually demineralised, and the acid-insoluble fraction of collagen was gelatinised in 5 mL of pH 3.0 HCl (purified water adjusted to pH 3.0 with 0.6 M HCl solution) at 80 °C for 24 hours, filtered through a 30 kDa centrifugal filter (Amicon) and freeze-dried overnight. Prior to MS analysis the lyophilisate was resuspended in 50 mM ammonium bicarbonate (pH 8.0) at a concentration of 2 mg/mL.

6.1.5.2 EDTA demineralisation/Protein extraction

The EDTA demineralisation protocol of Koon *et al.,* (2012) was adapted as follows. 0.5 M EDTA solution was prepared by dissolving 93.06 g of EDTA disodium salt in 500 mL of purified water, and the pH was then adjusted to 7.4 using 0.5 M NaOH. Each bone chip was placed in a 15 mL polypropylene centrifuge tube and 5 mL of 0.5 M EDTA (pH 7.4) added. The samples were stored at room temperature on an electric sample rocker, and the EDTA solution was replaced every three days. After 20 days the samples appeared to be visually demineralised, and the acid-insoluble fraction of collagen was gelatinised in 5 mL of pH 3.0 HCl at 80° C for 24 hours, filtered through a 30 kDa centrifugal filter (Amicon) and freeze-dried. The resulting lyophilised protein was then resuspended in 50 mM ammonium bicarbonate (pH 8.0) at a concentration of 2 mg/mL.

6.1.6 MALDI-MS analysis

The protein extracts suspended in ammonium bicarbonate solution (pH 8.0)) at a concentration of 2mg/mL, were digested with 1 µL of porcine trypsin solution (0.4 µg/µL 50 mM acetic acid) overnight at 37 ^{o}C. Digests were purified using 100 µL C18 solid-phase tips (Millipore ZipTips). After loading, the tips were washed with 0.1% trifluoroacetic acid (TFA) solution. Peptide mixtures were then eluted in 50 µL of 50:50 (v/v) acetonitrile: 0.1% TFA). The resulting peptide mixtures, consisting predominantly of tryptic peptides, were analysed using matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF-MS). A volume of 1 µL of sample solution was spotted on a ground steel MALDI target plate, followed by 1 µL of α-cyano-4-hydroxycinnamic acid matrix solution (1% in 50% ACN/0.1% TFA (w/v/v)). The sample and the matrix solutions were mixed together on the plate and allowed to air-dry. Each sample was spotted on to the MALDI target plate in triplicate. Each spot was analysed in reflectron mode using a calibrated ultraflex III (Bruker Daltonics, Bremen, Germany) MALDI-TOF instrument. Spectra were analysed using flexAnalysis software version 3.0 (Bruker Daltonics).

6.1.7 Determining the level of deamidation in a peptide

The deamidation of glutamine results in an overall mass increase of 0.984 Da. One disadvantage of the TOF instrumentation used in this work is that due to the insufficient resolving power of the mass analyser, it was not possible to resolve the deamidated and undeamidated signals: the nth peak of the deamidated peptide signal (typically the mono-isotopic signal) overlaps the $(n+1)^{th}$ peak of the undeamidated form (typically the signal for the species containing one ^{13}C atom). The extent of deamidation of glutamine (Q), converting it to glutamic acid (E) can be estimated by

deconvolution of the two overlapping distributions as described in Wilson *et al.* (2012). For a peptide containing just one glutamine residue, a value between zero and one (referred to as the α-value) denotes the proportion of glutamine that is deamidated, and is determined by optimizing the fit of overlapping theoretical distributions with the experimental distributions. An α-value of 1 indicates no deamidation, while a value of 0 results from complete deamidation. The method can be extended to peptides with more than one glutamine residue. Each sample was analysed in triplicate by MALDI-MS and the α-value obtained from a weighted average of the three spectra, where the weights reflect the signal to noise ratio (S/N) of each peptide. Full details are given in Wilson *et al.* (2012). The code used to calculate deamidation levels is available as an R package from GitHub (https://github.com/franticspider/q2e.git)

6.1.8 Analysis of collagen fibrils by transmission electron microscopy (TEM)

The modern, Roman and Pleistocene bovid bone samples were prepared for TEM analysis following the protocol of Koon *et al.* (2012). Small bone chips around 60 mg in weight from each sample were treated either with 0.6 M HCl or 0.1 M EDTA. Once demineralisation was complete (approx. 2 weeks) the demineralisation solutions were discarded, each sample was then washed five times, once with 2 mL of phosphate buffer (PBS) at pH 7, and four times with 2 mL of purified water. Before discarding the last water wash, the pH was measured to check that it was neutral. 60 mg of bone chips treated in this way were transferred to a plastic 15 mL polypropylene tube and 3 mL of 1% w/v solution of phosphotungstic acid (PTA) (adjusted to pH 7.0 with 0.5 M NaOH solution) was added. The sample was placed in an ice bath and homogenized for six 30-second intervals, with 30 seconds between pulses using a Yellowline DI25 basic homogeniser. The resulting solution was centrifuged at 3,000 RPM at 4° C for 15 min. The supernatant was discarded and the pellet resuspended in 1 mL of PTA. The sample was vortexed for 30 s. A drop of sample was then placed on a copper grid (with a formvar/carbon support film), 3 mm in diameter, 300 mesh, and left to settle for 5 minutes. Excess liquid was carefully removed from the grid using filter paper and the grid was left to air-dry. In order to stain the collagen fibrils a piece of filter paper was placed into a glass Petri dish and dampened with a 50:50% ethanol: purified water solution. A 2% w/v solution of uranyl acetate was prepared in 50:50 methanol and purified water. A square of dental wax was placed onto the filter paper and 500 µL of uranyl acetate solution spotted onto the wax. The dry sample grid was then floated (sample side down) on top of the uranyl acetate solution. The Petri dish was covered in tin foil and the sample grid left to stain for 30 min. The grids were then washed with 50:50 methanol: purified water and left to air-dry before being analysed by TEM. An FEI Tecnai G2 transmission electron microscope fitted with a CCD camera was used for analysis. The typical optical settings used were as described in Koon *et al.,* (2012). Beam setting was 120 kV.

6.2 Results

The results obtained for the Roman bone are described in terms of the variation in α-values calculated from the MALDI-MS data with respect to: a) the sub-sampling location (and localised areas displaying "macroscopic degradation" on the bone) and b) the protein extraction protocol. These results are then linked to the structural properties observed in protein extracted from modern, Roman and Pleistocene bone, investigated by TEM (Section 3.5).

6.2.1 Variation of Gln deamidation as a function of sampling location

To investigate the variability in levels of glutamine deamidation (α-values) between different sampling locations within a bone, chips were sub-sampled from parts of macroscopically well-preserved sections of slices 1 (~3 mm from the right), 2 (at ~ 15 mm) 3 (at ~27 mm), 4 (at ~39 mm) and 5 (at ~ 117 mm) were sampled (Figure 60). Two chips were taken from each slice, and extracts from each of these two chips were analysed in triplicate by MALDI-MS. Each triplicate analysis generated one α-value; the two α- values generated for each chip were then averaged, and the average α-values for each slice are what is represented in Figure 60. Although, initially, twelve peptides were investigated (Table 9), α- values are only reported here for the ten collagen peptides that were observed in protein extracts from all five slices (Figure 60).

Table 10: 12 peptides that are observed in MALDI mass spectra of tryptic digests of bovine type I collagen and contain at least one glutamine residue. Where possible the theoretical amino acid sequence of the peptides has been demonstrated by product ion analysis. For peptides where this was not possible, due to poor spectral quality, sequences were taken from published data (Wilson *et al.,* 2012) and assigned on the basis of the peptides' accurate *m/z* values.

* Assignment of sequence demonstrated using product ion spectrum.

[M + H]$^+$	Peptide sequence	Collagen chain	Position in collagen chain
836.44	GPAGPQGPR*	COLL 1A1	[1084-1092]
1105.57	GVQGPPGPAGPR*	COLL 1A1	[685-696]
1690.77	DGEAGAQGPPGPAGPAGER	COLL 1A1	[612-630]
1706.77	DGEAGAQGPPGPAGPAGER	COLL 1A1	[612-630]
2056.98	TGPPGPAGQDGRPGPPGPPGAR*	COLL 1A2	[552-573]
2073.01	GAPGADGPAGAPGTPGPQGIAGQR	COLL 1A1	[934-957]
2089.01	GAPGADGPAGAPGTPGPQGIAGQR	COLL 1A1	[934-957]
2689.26	GFSGLQGPPGPPGPSGEQGPSGASGPAGPR	COLL 1A1	[1111-1140]
2705.26	GFSGLQGPPGPPGSPGEQGPSGASGPAGPR*	COLL 1A1	[1111-1140]
3001.50	GPSGEPGTAGPPGTPGPQGLLGAPGFLGLPGSR	COLL 1A2	[845-877]
3100.41	GLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPR*	COLL 1A1	[187-219]
3665.54	GSQGSQGPAGPPGPPGPPGPPGPSGGGYEFGFDGDFYR*	COLL 1A2	[1079-1116]

Figure 60 shows the average α-value for each peptide from the two chips from each slice. Some peptides produce similar α-values regardless of the sampling location (for example peptides with *m/z* values 3100.5, 1105.6, 1706.8, 2705.2), but other peptides (for example, peptides with *m/z* values 2056.9, 2073.0, 2689.1 and 3665.8 in particular) show greater variability with sampling location.

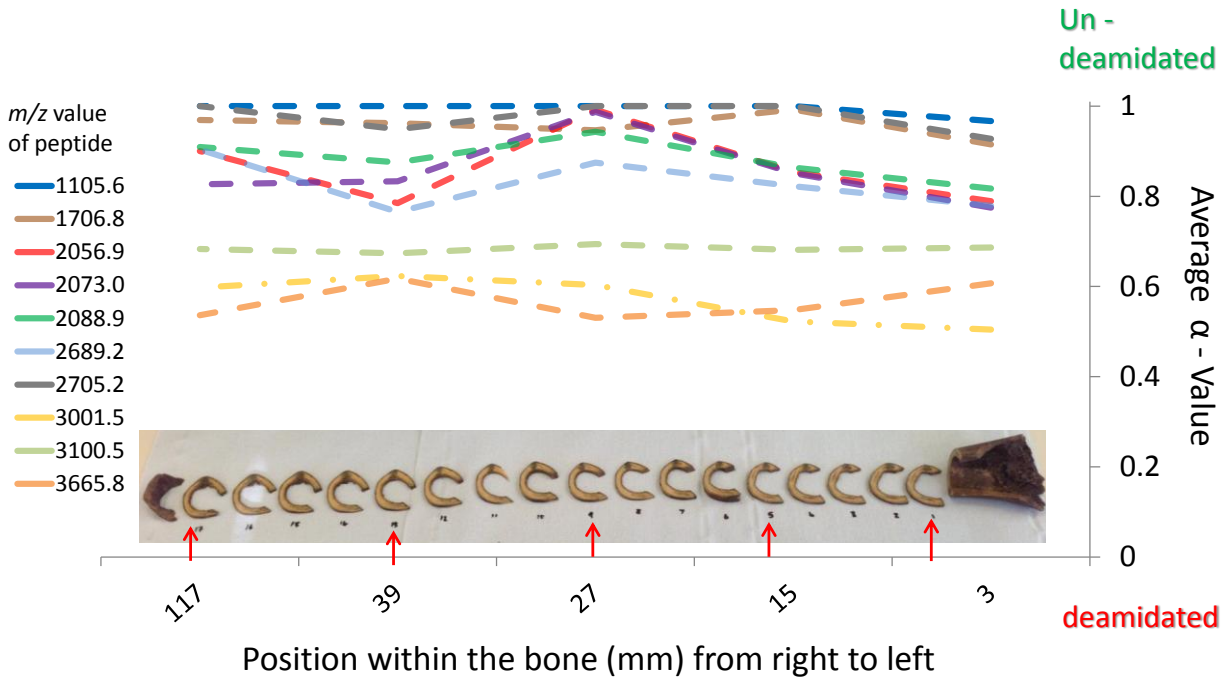


Figure 60: α-values for 10 peptides, in 10 samples, obtained from two chips from each of the five different positions (slices 1, 5, 9, 13 and 17) across the length of a Medieval bovine metatarsal bone. The average value for the two chips from each slice is plotted.

Considering each slice as a group, the usual equations for within-group and between-group variance can be used to calculate the variances within and between slices for each peptide (Snedecor 1934). Thus, the between-slice variance is given by equation 1:

$$V_b = \frac{1}{(S-1)} \sum_{s=1}^{S} n_s (\bar{x}_s - \bar{\bar{x}})^2 \qquad (1)$$

where $S = 5$ is the number of slices, n_s is the number of α-values from each slice (i.e. 2, here), $\bar{x}_s$ is the mean α-value for slice s and $\bar{\bar{x}}$ is the grand mean, taken over all slices and the within-slice variance is given by equation 2:

$$V_w = \frac{1}{(N-S)} \sum_{s=1}^{S} \sum_{i=1}^{n_s} (x_{is} - \bar{x}_s)^2 \qquad (2)$$

where $N = 10$ is the total number of α-values and x_{is} is the ith α-value from slice s. Table 10 and Figure 61 show the within-slice and between-slice variances for the ten peptides, together with the p-values obtained for F-tests comparing the two variances. The variance between slices is shown to be significantly greater than the variance within slices (at the 95% confidence level) for just two peptides, those with *m/z* values 2073.0 and 2689.2, although with a p-value of 0.038 for both the evidence against the null hypothesis is not strong. The peptide with *m/z* value 2056.9 has the

highest level of within-slice variation, but is of a similar level to the variance between slices. The remaining peptides also show similar levels of variation within and between-slices.

Table 11: The variation in α-values obtained from 10 peptides measured in tryptic digests of collagen, extracted from bone chips of different slices compared with the variation obtained from replicate chips of the same slice. The p-values for F-tests show that, in general, the between-slice variance is not significantly greater than the within-slice variance. *denotes statistically significant values (at the 95% confidence level).

m/z of peptide	1105.6	1706.8	2056.9	2073.0	2088.9	2689.2	2705.2	3001.5	3100.5	3665.8
Between-slice variance, V_b	0.001	0.001	0.012	0.012	0.005	0.006	0.002	0.006	0.001	0.004
Within-slice variance, V_w	0.001	0.001	0.011	0.002	0.004	0.001	0.001	0.002	0.001	0.001
p-value for F test	0.486	0.486	0.451	0.038*	0.398	0.038*	0.233	0.13	0.486	0.08

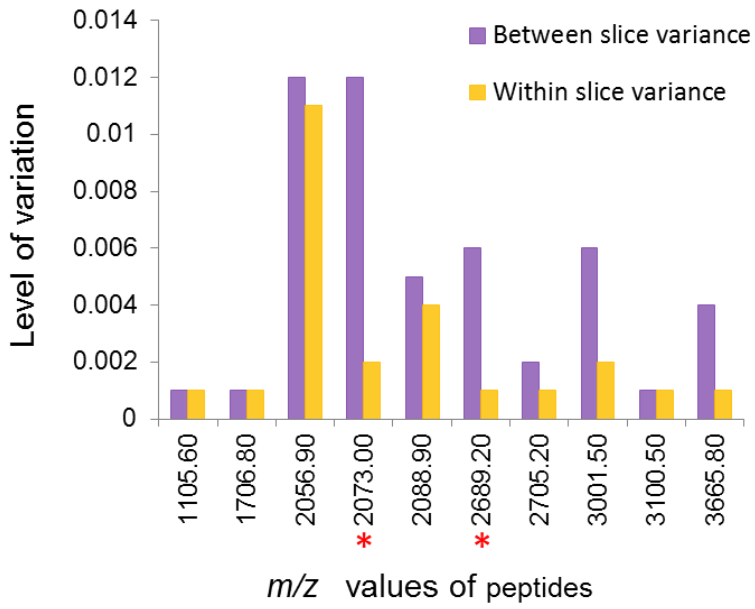


Figure 61: The variation in α-values obtained from 10 peptides measured in trypsin digests of protein extracted from bone chips of different slices compared with the variation obtained from replicate chips of the same slice. The P-values for F-tests show that, in general, the between-slice variance is not significantly greater than the within-slice variance. *denotes statistically significant values (at the 95% confidence level)

6.2.2 Variation due to localised diagenesis

In order to investigate the effect of localised diagenesis on α-values, two chips were taken from the degraded sections of slices 1, 5, 9, 13, and 17 (Figure 62) and the α-values compared with those obtained from chips in macroscopically well-preserved areas of the same bone slice. The spectra obtained from chips from locally degraded regions contained fewer peaks than those from the macroscopically well-preserved chips, with the heavier peptides (*m/z* 3001.5, 3100.5 and 3665.8) absent in spectra of samples from degraded regions. In the spectra from visibly degraded chips, there were a total of 106 observations of these peptides in comparison to 114 observations in the spectra from well-preserved chips (out of a possible 120). In most cases, the average α-values obtained for macroscopically degraded sections were lower (i.e. the peptides were overall more deamidated) than those extracted from macroscopically well-preserved areas. Figure 4 shows the average α-values for the two chips in each case. Interestingly, the four peptides that show least deamidation in well-preserved chips (*m/z* values 1105.6, 1706.7, 2088.9 and 2705.2, with mean α-values of 0.99, 0.96, 0.88 and 0.98 respectively) also show little deamidation in the degraded chips (mean α-values of 0.98, 0.99, 0.86 and 0.92 respectively).

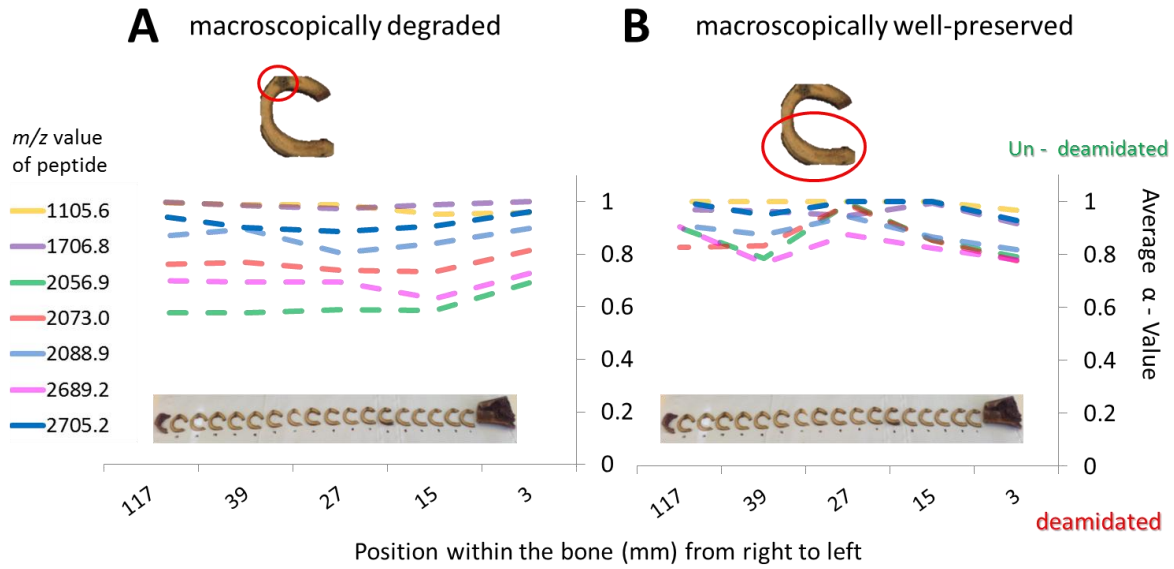


Figure 62: Comparison of α-values obtained from peptides observed in tryptic digests of protein extracted from macroscopically degraded sections of bone (A: left) with those from macroscopically well-preserved areas of the same slice (B: right). Here α-values are only plotted for the seven peptides which were observed in all five slices.

Other peptides (*m/z* values 2056.9, 2073.0 and 2689.2) show greater changes between the visibly well-preserved (mean α-values of 0.86, 0.85 and 0.83 respectively) and degraded areas (mean α-values of 0.59, 0.76 and 0.69 respectively). Figure 63 shows that the variation in deamidation levels

across the bone is slightly less for the degraded samples than the spread for the well preserved region-derived samples.

6.2.3 The effects of acid demineralisation on deamidation

The removal of mineral using HCl is common in most bone preparation techniques such as those for isotope analysis and radiocarbon dating (e.g. Brock *et al.* 2007). An alternative to the use of HCl for the decalcification of bone is the use of EDTA as a chelating agent. EDTA decalcification is often used when trying to minimise damage to the surface histology of bone (Jonsson, Tarkowski, and Klareskog, 1986; Tuross 2012).

HCl demineralisation was compared with the ammonium bicarbonate protein extraction method developed by van Doorn *et al.* (2011), which does not involve the removal of mineral from the bone. We assessed the effects of HCl demineralisation on the overall deamidation using bone chips from macroscopically well-preserved areas of slices 1, 5, 9, 13 and 17. The α-values of 12 peptides produced after HCl treatment were compared with those determined from chips from similarly well-preserved areas of the same slice, in which collagen was extracted using the ammonium bicarbonate extraction method. The 12 peptides were observed less frequently in spectra from samples treated with HCl than from those treated with only ammonium bicarbonate (Figure 63A). In the spectra obtained from the HCl-treated samples, only 74 (of a possible 120) observations of the peptides were recorded, compared with 114 in spectra from mineralised protein extracted with ammonium bicarbonate (Figure 63A). This suggests the HCl treatment affects the peptides detected in the samples. Five of the twelve peptides (*m/z* 1690.8, *m/z* 1706.8, *m/z* 2057.0, *m/z* 2073.0, *m/z* 2089.0) were observed in less than half of the HCl-treated samples. It should be noted that each of these peptides has an aspartic acid on the N-terminal side of glycine. The remaining peptides, observed in at least half of the HCl-treated samples, did not contain aspartic acid.

In observed peptides, the α-values calculated for samples treated with HCl were generally lower than those from samples treated only with ammonium bicarbonate (Figure 63 B), indicating greater levels of deamidation in HCl-treated samples.

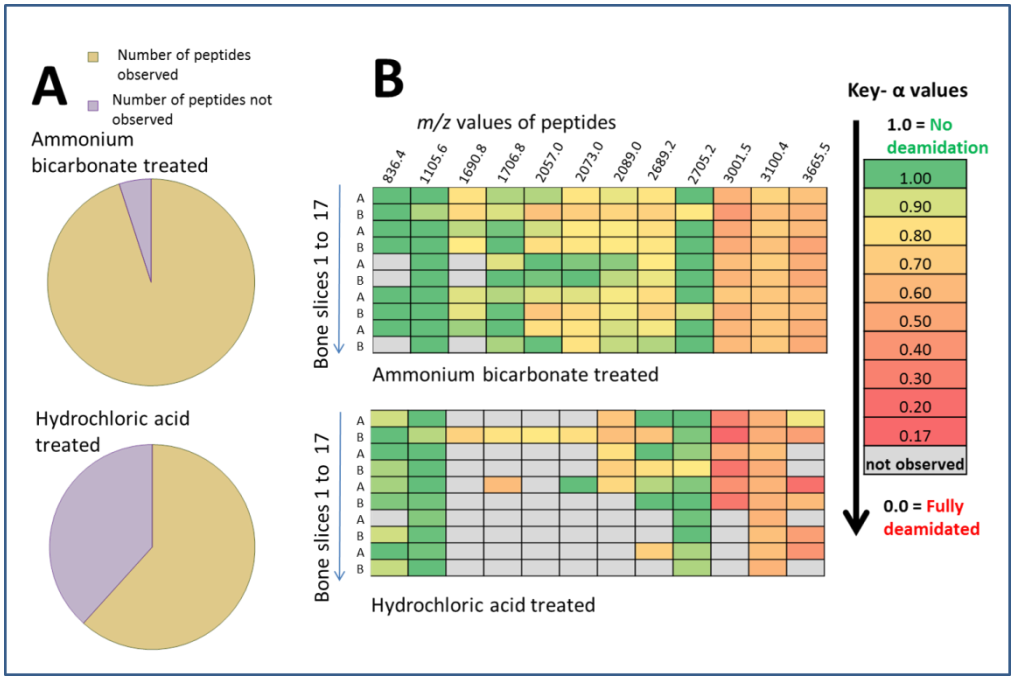


Figure 63: (A) A comparison of the number of times the peptides in Table 1 were observed in spectra obtained from samples treated with HCl or ammonium bicarbonate solutions. (B) Comparison of α-values obtained for these 12 peptides in spectra from macroscopically well-preserved areas of the Medieval bone (2 each from slice: 1, 5, 9, 13, and 17) after treatment with ammonium bicarbonate (top) or HCl (bottom)

6.2.4 Effects of demineralisation time on α-values

In order to compare the effects of HCl (pH 1) and EDTA (pH 7.4) on glutamine deamidation, the remaining unanalysed chips from the macroscopically well preserved sections of the 17 slices of bovine metatarsal were mixed together. A total of 24 chips from this sample set were demineralised for up to four weeks in either HCl or EDTA (see sections 2.3 and 2.4). For each demineralisation method, four chips were removed from the solutions after 2, 3 or 4 weeks. The collagen was extracted as described in sections 2.3 and 2.4. The resulting protein extracts were digested and purified as described in section 2.5 and analysed using mass spectrometry.

For each of the samples, levels of glutamine deamidation were calculated (section 2.6), and the patterns observed can be split into three categories: 1) peptides (i.e. *m/z* 2689.3, 2705.2 and 3100.4) which showed lower α-values (i.e. more deamidation) with increased variability when treated with HCl than EDTA (Figure 34). 2) peptides (i.e. *m/z* 2705.2 and 3100.4) which showed increased levels of deamidation on acid treatment over time, with α-values for 2705.2 ranging from 0.57 – 0.87 in HCl-treated samples; this peptide shows little or no deamidation in samples treated with EDTA over the four week period, with values of EDTA treated samples producing α-values ranging from 0.92 – 1.00. 3) Some of the smaller peptides (*m/z* values 836.4 and 1105.6) showed

little difference in deamidation levels regardless of the demineralisation procedure used, or the length of time they were treated. Examples from the three categories are shown in Figure 64.

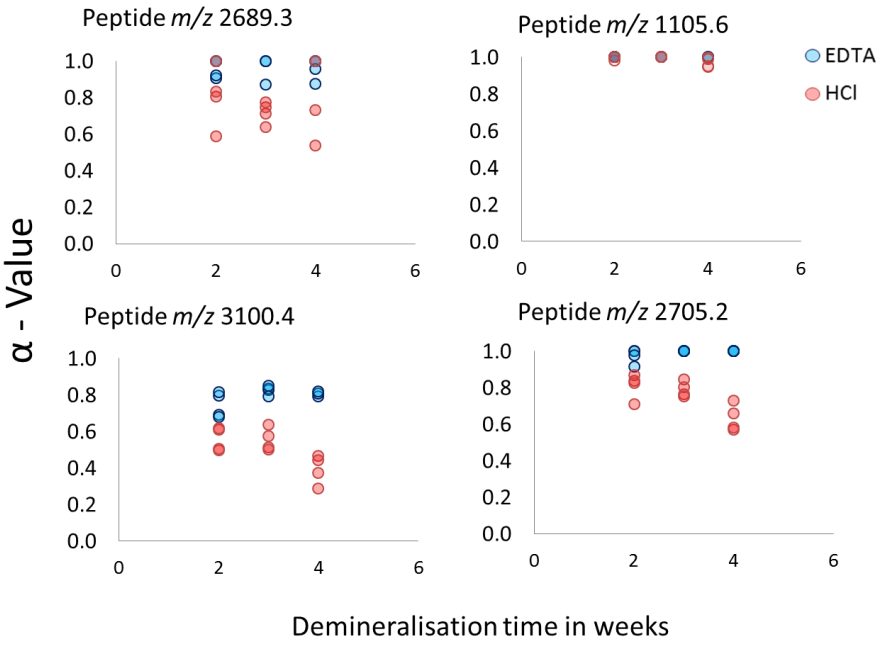


Figure 64: Comparison of α-values obtained for four peptides after demineralisation in HCl or EDTA for 2, 3 or four weeks. Peptides with smaller masses such as 1105 showed little deamidation regardless of the demineralisation method used. In samples pre-treated with HCl three peptides (*m/z* 2689.3, 2705.2 and 3100.4) showed an increase in deamidation over time, in contrast to EDTA pre-treatment which did not appear to induce deamidation over time.

6.2.5 Comparison of collagen fibril structure in modern, Roman and Pleistocene bone demineralised with either EDTA or HCl using transmission electron microscopy (TEM)

To investigate the effect of different demineralisation methods on the structure of collagen fibrils, three bovid bones of different ages, modern, Roman and Pleistocene were used. Bone chips from each sample type were sampled and the mineral from each sample was removed using either HCl or EDTA. The extracted collagen was visualised using TEM and the preservation state and average width of the collagen fibrils was investigated. Measurements of the width were taken at ten points along the length of 20 fibrils, resulting in a total of 200 measurements for each of the six samples. The distribution of measurements was assessed to be plausibly normal for each sample and the statistical significance of the difference in mean fibril width between HCl and EDTA treated samples was determined using a two-tailed, two sample t-test for unequal variances for each of the modern, Roman and Pleistocene samples.

In each case, the average fibril width was found to be significantly larger for HCl-treated samples than in EDTA-treated samples (Table 11).

Table 12: Average fibril width measurements from three samples of bone of different ages (modern, Roman and Pleistocene). Fibril widths measured in all three samples were found to be statistically significantly different at the 95 % confidence level, when prepared using the two pre-treatment methods. In each case, the t-test shows the average fibril width is significantly greater in HCl-treated samples

Sample	Mean fibril width (HCl treated)	SD	Mean fibril width (EDTA treated)	SD	p-value
Modern bone	90.63	14.63	76.88	13.31	1.99 E-20
Mediaeval bone	96.36	29.90	72.18	23.37	1.02 E-17
Pleistocene	96.77	33.68	69.11	15.91	5.2 E-22

In the TEM observations, collagen fibrils are shown by the characteristic dark and light banding along the length (Figure 65). This is due to the highly regulated structure and arrangement of the fibrils within the collagen protein (Orgel *et al,.* 2001). However, the HCl-treated modern collagen resulted in fibrils with less uniform fibril widths than those treated with EDTA, as well as regions of swelling along the fibril length (Figure 65). In contrast, the collagen from the modern bone treated with EDTA resulted in a higher number of fibrils per square on the grid than those treated with HCl. The effect of HCl demineralisation was also evident in the Roman bone. When treated with HCl, the extracted fibrils showed less defined structure with areas of swelling and more disruption to the banding than those treated with EDTA (Figure 65). The detrimental effect of HCl demineralisation on fibril structure was most evident in Pleistocene bone, with very few collagen fibrils displaying the characteristic banding, whereas banding was still evident in the majority of the fibrils in the EDTA-demineralised sample.

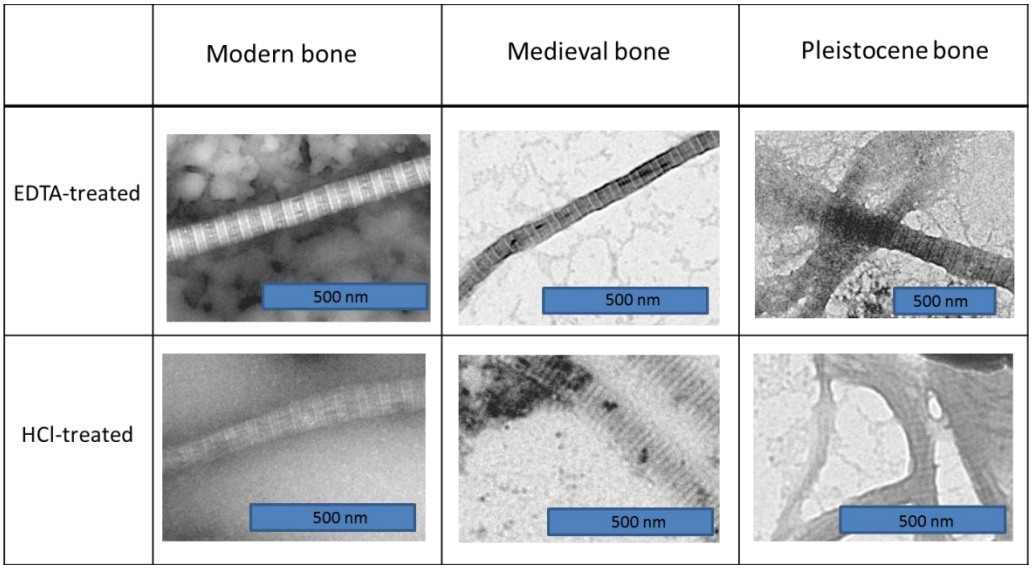


Figure 65: Transmission electron micrographs of protein extracted from modern, Roman and Pleistocene bone treated with either 0.6 M HCl or 0.5 M EDTA.

6.3 Discussion

6.3.1 Spatial variation in deamidation levels within a sample

Our findings show that, in the Roman bovine metatarsal bone investigated here, the sampling location across areas of well-preserved compact bone does not generally contribute significantly to differences in the level of deamidation observed. This may be attributable to highly structured and repetitive nature of the protein and the dense packing of the surrounding mineral. Samples taken from areas of bone that displayed localised macroscopic diagenesis showed elevated levels of deamidation of some peptides. This may be due to localised differences in the bone structure in this “darkened” region. It should be noted that only one bone was used to investigate sampling point variability in this study and although the protein structure is conserved throughout different bone types (e.g. long or flat bones) the level of mineralisation or the effect of structural anatomical differences on levels of glutamine deamidation has not been investigated. The increased deamidation from areas of the bone that display localised, macroscopic diagenesis highlights the importance of sampling from areas that are representative of the overall preservation of the bone, i.e. by avoiding areas that are clearly and visibly compromised.

6.3.2 Effects of sample pre-treatment and extraction methods on glutamine deamidation and the collagen fibril structure

The gentle protein extraction method developed by van Doorn *et al.*, (2011) has the advantages of being fast to perform and minimally destructive to the bone, as it does not require decalcification pre-treatment. However, we have found that this extraction does not always yield sufficient amounts of collagen for successful MS analysis. Extraction using only ammonium bicarbonate solution may result in partial protein extraction for a number of reasons. For example, as the buffer-soluble collagen is easily extracted, it is possible that some of it is lost due to leaching or exchange within the burial environment, especially in sites with fluctuating water tables (High *et al.*, 2015). Also, the buffer-soluble fraction is likely to be gelatinised and therefore may not be truly representative of the general state of preservation of the majority of the bone collagen.

Our results show that demineralisation treatment using HCl influences the extent of deamidation; HCl increases the level of glutamine deamidation and decreases the number of peptides detected in comparison with EDTA treatment. Both asparagine and glutamine deamidation have been studied in a range of sample types, from short synthetic peptides (Li *et al.*, 2010; Robinson *et al.*, 1970; Geiger and Clarke, 1987; Robinson, 2004; Stratton *et al.*, 2001) to proteins such as α-crystallin of the eye lens (Takemoto and Boyle, 1998), collagen (Hurtado and O'Connor, 2012; van Doorn *et al.*, 2012; Wilson *et al*., 2012), keratin (Araki and Moini, 2011) and protein binders in paint (Leo *et al.*, 2011). Asparagine is known to have two deamidation pathways: either via a cyclic succinimidyl (five membered ring) intermediate, or via direct side chain hydrolysis (Capasso *et al.* 1991; Radkiewicz *et al.*, 1996; Xie and Schowen, 1999). The latter reaction has been found to be favoured at low or high pH (Robinson, 2004). Glutamine can also deamidate via two pathways (Robinson 2004; Li *et al.* , 2010), forming a cyclic glutimidyl (six membered ring) intermediate. The two residues have different rates of deamidation via cyclic intermediates, with glutamic acid forming at a slower rate than aspartic acid (Li *et al.*, 2010). The most probable route of deamidation for both residues in a highly structured protein such as collagen is via direct side chain hydrolysis, due to the lack of flexibility necessary for the protein backbone to adopt the appropriate interatomic distance needed for the formation of the cyclic intermediates (Van Duin and Collins, 1998). It is therefore likely that the two residues are equally stable in proteins such as collagen. However once in solution, gelatine (the soluble form of collagen) no longer has the same rigid structural constraints, and exists in the form of random coils.

We observe increased levels of glutamine in HCl (pH 1) treated samples. This is most likely due to an increase in direct side chain hydrolysis, which is less likely to occur during the ammonium bicarbonate or EDTA extractions, both carried out at around pH 8.0 (Van Duin and Collins 1998) . Low pH is known to induce peptide bond hydrolysis (Hill 1965). However, in the experiments presented here the bone was treated in a fairly weak acid solution (0.6 M HCl) under refrigerated

conditions (4-5 °C). It is therefore unlikely that these conditions would result in the activation energy required to significantly hydrolyse the peptide bonds of the protein. Of the 12 Gln-containing peptides studied here, those that were not observed in spectra of HCl-treated samples all contained aspartic acid (peptide sequences in Table 2). The literature has shown that aspartic acid-proline bonds undergo hydrolysis at low pH under conditions where other aspartyl bonds are found to be stable (Pisskiewicz *et al*., 1970). In the peptides measured here, the aspartyl is always to the N-terminal side of Gly; Radkiewicz *et al.* (2001) found that the degradation of aspartyl-glycine bonds can be promoted due to an increased rate of ring formation, with Asp-Gly having a short half-life compared to Asp bound to other amino acids (Ser, Ala, Cys and His). The half-life of Asp-Gly degradation at 37 °C, pH 7.4 was found to be 41-71 days, in comparison with 266 days for Asp-His and Asp-Ala. It is possible that at low pH cyclisation at the aspartyl-glycine occurs, although currently not enough is known about how these bonds in collagen are affected over time, or at different pH. If the aspartyl-glycine bond is more prone to breakage than other Asp-amino acid bonds, this may explain the lack of Asp-Gly containing peptides in the spectra of HCl treated samples. However, from these experiments we have no direct supporting evidence of preferential breakage at the aspartyl-glycine bond.

Low pH has been found to emphasise areas of damage in cooked collagen, as it induces observable swelling at sites of damage (Koon *et al*., 2010). The TEM findings presented in this paper support the theory that HCl treatment of bone causes degradation of the collagen structure and that older bone may be more susceptible to pH-induced damage. Greater knowledge of the contribution of the 3D structure to the stability of residues at specific sites would help further understanding of the breakdown pathways of bone collagen, as well as of the observed differences in deamidation rates for different Gln-containing peptides.

6.4 Conclusions

We have explored two potential causes of variation in Gln deamidation determined in bovid bone. This study found that for some peptides, levels of deamidation were reproducible across the length of areas of macroscopically well-preserved bone. Given that sample point variation was investigated in only one bone, in order to fully understand the possibility of sample point variation, a wider study of multiple bone types would be necessary. Our results suggest that the level of glutamine deamidation is linked to the preservation state of collagen in bone, with macroscopically degraded sections resulting in increased levels of deamidation. Measurement of glutamine deamidation may therefore be a useful screening tool when selecting bone material for collagen-dependent analysis.

When looking to extract collagen, especially from old or poorly preserved bone, it appears that EDTA-treatment is preferable to HCl-treatment. We conclude that, although acid demineralisation has been shown to be suitable for other types of collagen analyses (e.g. for radiocarbon dating, or

dietary studies (*Sealy et al.*, 2014)), this pre-treatment method clearly disrupts the collagen structure and causes some damage to the protein structure. EDTA demineralisation is preferable for mass spectrometric analyses aimed at quantifying the extent of glutamine deamidation in samples where ammonium bicarbonate extraction is unsuccessful, or in particularly degraded or old samples.

In the 12 peptides considered here, some appeared to be more stable than others and underwent deamidation more slowly, similar to the observation of van Doorn *et al*. (2012) and Wilson *et al*. (2012), who calculated different half-lives for glutamine in different peptides. We suggest that these stable peptides may be particularly useful when evaluating the preservation state of Pleistocene bone material. On the other hand, rapidly-deamidating peptides may be most suited to determination of the extent of diagenesis in younger (Holocene and/or Late Pleistocene) bones. In order to further investigate the relationship between thermal age and glutamine deamidation, a number of bones from dated sites are currently being analysed which should help answer this question.

Chapter 7 - Measurement of glutamine deamidation in collagen peptides extracted from bones spanning ~11,000 – 300,000 years BP

Introduction

Despite the robust nature of bone material, in terms of its survival in the burial environment, there have been no systematic studies investigating the degradation of protein in bone older than the radiocarbon limit (~ 50 ka BP). In a study by Buckley *et al.,* (2011) it was shown that collagen could be successfully extracted from bone > 1.5 million years in age and it has been postulated that bone protein could survive over geological time periods (Smith *et al.,* 2005). Unfortunately, even though there is an abundance of bone this material found in sites throughout the Quaternary, it is not currently possible to directly date bone beyond the radiocarbon limit of ~ 50 ka BP (van der Plicht and Palstra 2014). Although attempts have been made to extend this age limit using other methods, such as uranium-thorium dating, this technique has been problematic and associated with large errors since its inception (Grün 2006). This is largely due to the fact that bone is an open system, meaning that early uptake and leaching of uranium can occur (Pike *et al.*, 2002). Levels of deamidation have been measured in protein extracted from a number of biological materials such as: keratin in wool (Araki and Moini, 2011), collagen in bone (van Doorn *et al.*, 2012; Hurtado and O'Connor, 2012; Wilson *et al.,* 2012), α-crystallin eye protein (Takemoto and Boyle, 1998; Flaugh *et al.*, 2006), protein binders in paint (Leo *et al.,* 2011), and keratin in mummified skin (Mikšík *et al.,* 2014). These studies suggest that the measurement of deamidation in protein may provide a versatile tool for the investigation of bone protein diagenetic history.

It has been shown in this thesis (Chapter 5) and in the literature that the rate of deamidation is temperature dependent (Sinex, 1957; Scotchler and Robinson, 1974; Stratton *et al.,* 2001; Robinson, 2004), although there are other factors that may affect the rate at which deamidation occurs, or the observed extent of deamidation. These include: the extraction method, analytical instrumentation used, pH, bone preservation, peptide sequence, as well as the positioning of the residue within the 3D structure of the protein. It has been observed in experiments presented in this thesis (Chapter 3) and in the literature (van Doorn *et al.,* 2012) that some collagen peptides consistently undergo deamidation more slowly than other Gln-containing collagen peptides (e.g. *m/z* 1105 appears to undergo deamidation more slowly than other peptides investigated in this thesis). It may be possible that these slower, more stable peptides may be useful when evaluating the preservation states of older bone material, i.e. those across the Pleistocene.

It has been suggested that there may be a relationship between the rate of glutamine deamidation in bone collagen and thermal age (van Doorn *et al.*, 2012), and levels of deamidation in HCl-treated bone from a range of Pleistocene sites were found to correlate with the predicted thermal ages

(Welker *et.al*, 2015). However, it should be noted that the levels of error observed in the 5 samples in this study ranged from ± 3.2 to ± 24.5 %. Currently no conclusive study has been carried out to measured levels of glutamine deamidation in a wide range of dated bone material in order to test this robustly. In this chapter, mass spectrometry is used to measure levels of glutamine deamidation in bone samples dating from ~11,000 years BP to as old as 300,000 years BP in order to investigate the relationship between glutamine deamidation and thermal age.

7.1 Materials and Methods

7.1.1 Sites selected for the study

Ideally when choosing material it would be best to analyse: multiple samples from areas of compact bone, multiple bones from each site, bone from sites which have been securely dated, and to compare data from identical or similar species across the sites. Unfortunately this is not always possible. The bone available for sampling from palaeontological sites can quite often consist of small fragments, bones from un-stratified sites, or sites for which the exact MIS or interglacial period is unknown. In addition, there is often limited availability of comparable species across all sites. Taking this into account, the samples selected for analysis in this thesis were deemed suitable for preliminary investigation of the applicability of deamidation as a measure for thermal age. Bone samples from nine sites have been analysed (Table 12): Tanner Row, York (UK), Yukon permafrost (Canada), Stump Cross (UK), Scladina (Belgium), Joint Mitnor Cave (UK), Victoria Cave (UK), Neumark Nord (Germany), Ilford (UK), and Schöningen (Germany). These sites were selected in order to investigate levels of glutamine deamidation in samples which cover a range of chronological and thermal ages. The thermal ages for each of the sites were estimated using www.thermal-age.eu. It should be noted that the activation energy for glutamine deamidation in collagen at different temperatures is not known, and therefore the level of thermal degradation of the protein is estimated using data from collagen and DNA degradation. When using www.thermal-age.euthe thermal age is calculated by calculating the probability of damage as a function of temperature (reported as γ). γ is then calculated with a constant temperature of 10 °C and the difference between the two values (relative difference) is then multiplied by the age of the sample, giving the estimated thermal age.

Sample sites, showing the age of the sample in years BP, the MIS, species analysed and estimated thermal age

Table 13: Sample sites, showing the age of the sample in years BP, the MIS, species analysed and estimated thermal age

Site Name	MIS	Age in years BP	Species	Thermal Age (years)	Latitude and longitude co-ordinates
Tanner Row	1	~800 -1000	*Bovid*	731 - 12515	53.958906, -1.081160
Yukon permafrost	2-5a	~12,000-~80,000	Bison and horse	See table 13	64.050395, -139.435922
Stump Cross	5a-c	~82-96 ka	*Bison*	7681 - 7926	54.067658, -1.865090
Scladina Cave	5a-e	~82 -123 ka	*cervid*	15021 - 23940	50.485731, 5.026254
Joint Mitnor Cave	5e	~123 ka	*Bovid*	28383	50.484601, -3.771299
Victoria Cave	5e	~123 ka	*Bovid*	12002	54.294735, -1.983461
Neumark – Nord	5e or 7 or	~123 ka, or ~243 ka	Forest elephant	19368 -23187	51.336785, 11.888964
Ilford	7	~ 243 ka	*Bovid*	51454	51.573410, 0.083750
Schöningen	8	~ 300 ka	*cervid*	35133	52.143458, 10.957362

*MIS ranges taken from Lisiecki and Raymo (2005). Thermal age is reported as a function of degradation in 10 °C thermal years and calculated using www.thermal-age.eu

7.1.1.1 Tanner Row (UK)

Tanner Row (23/36) is thought to be a Roman site and was excavated by York Archaeological Trust in 1984. Excavations revealed a number of organic remains including a large number of cattle bones thought to date around the late 2nd century. A piece of *bovid* bone (Figure 66) from this site has been used throughout this thesis for experimental investigations on the effects of different aspects of the sample preparation process.

Figure 66: The piece of *bovid* bone excavated from Tanner Row, thought to date from around ~ late 2nd century Roman York (UK)

7.1.1.2 Yukon (Canada)

A total of ten0 bison, and three horse bone fragments (Table 13) excavated from a permafrost site in the Klondike region of Canada's Yukon Territory were investigated. These bones were AMS radiocarbon dated at the Center for Accelerator Mass Spectrometry, Lawrence Livermore National Laboratory, California, USA (CAMS 157517). It was initially thought that these samples dated from ~ 80,000 years BP as they were found in association with a volcanic ash (tephra) layer, Sheep Creek-K, that has been dated to ~ 80,000 years (Westgate *et al.,* 2008). Six of these fragments were dated as younger than 40,000 years BP with dates ranging from 12,365 to 34,830 years BP (CAL). Seven of the fragments were dated as older than 40,000 years, with five of these producing dates over 50,000 years BP. These older dates are at the limit of radiocarbon dating (~ 50 ka BP; van der Plicht and Palstra, 2014), so it is therefore likely that results for samples > 50,000 ^{14}C years BP are not accurate age estimates, and that these samples may well be associated with the ~ 80,000 year old volcanic ash layer. This is discussed more detail in Section 6.1.

Table 14: Details of samples obtained from Yukon Canada including: sample ID number, species, bone type, radiocarbon date with associated errors, estimated thermal ages with associated errors and laboratory reference numbers (CAMS). The samples were radiocarbon dated at the Lawrence Livermore Nationtal Laboratory (LLNL) California.

sample ID number	species	bone type	radiocarbon date	variance from radiocarbon date (+/-)	CAMS # (from LLNL)	estimated thermal age	thermal age (min)	thermal age (max)
EL-10-17	horse	Radius	12365	35	157471	1306	1303	1308
EL-10-238	horse	metacarpal	13085	40	157473	1341	1340	1308
QZ-10-11	horse	pelvis fragment	25610	140	157467	1531	1530	1534
QZ-10-81	bison	radius fragment	28970	220	158440	1574	1572	1576
QZ-10-346	bison	scapula fragment	34700	450	158448	1643	1637	1649
QZ-10-348	bison	humerus fragment	34830	460	158446	1644	1640	1652
QZ-10-83	bison	humerus fragment	42700	1200	157509	1753	1740	1768
QZ-10-92	bison	humerus	42800	1200	158442	1754	1742	1769
QZ-10-232	bison	humerus fragment	>52200	27800	158451	1859	1859	2368
QZ-10-360	bison	astralagus	50300	3100	158438	1861	1817	1911
EL-10-50	bison	metapodial	>50300	29700	157517	1864	1864	2368
QZ-10-353	bison	tibia	>50500	29500	158439	1864	1864	2368
QZ-10-220	bison	metatarsal	>52400	27600	157513	1894	1894	2368

* All samples were chips removed from bone fragments using pliers; each chip sampled was ~ 30 mg

7.1.1.3 Stump Cross (UK)

Stump Cross caves are located between Wharfedale and Nidderdale and have been described as key sites for understanding the latter stages of MIS 5 (O'Connor and Lord, 2013). The cave has been extensively dated (the site is reported as having 55 age determinations (O'Connor and Lord, 2013)) mainly by measuring U/Th in speleothems; these dates have been correlated with the biostratigraphy and fluctuating hydrology of the cave. The cave also offers an insight into the fluctuating hydrology, with dry and wet stages identified and evidence that the cave was repeatedly flooded from ~ 225,000 years BP to the mid-Holocene (O'Connor and Lord, 2013). The gravel beds and silts were found to be deposited before MIS 7. Ages for the site have identified a younger deposit dated to around MIS 5. This deposit is described as a classic horizon, and has been named "wolverine cave". This is due to the cave acting as a trap for large herbivores, which in turn lured in carnivores such as wolves and wolverine. In the 1980's U/Th dating of speleothems was carried out, resulting in a correlation with MIS 5b. At the site, bones from bison and reindeer as

well as carnivore species such as: brown bear, wolf and wolverine were found. In this chapter a bison bone found in wolverine cave (Figure 67) was sampled.

Figure 67: Piece of bison bone excavated from the wolverine cave at Stump Cross correlated to MIS 5a-c, ~ 82-96 ka BP

7.1.1.4 Scladina Cave (Belgium)

Scladina Cave is situated in Sclayn (Numur, Belgium), situated on the west side of the small Fond des Baux valley. The Cave was first discovered in 1971 and has been dated to between 80,000 and 127,000 years BP (Smith *et al.,* 2007). It was found that the sediment in the cave contained evidence of a number of climatic fluctuations, correlated to the Upper Pleistocene Period. Although Scaldina Cave has been described as one of the most complete sequences available in Belgium, it should be noted that the site of Scladina is complicated, especially since the reappraisal of the stratigraphy by Stéphane Pirson. The samples analysed were collected from layers 4 and 5b of the sequence (Figure 68).

The top of layer four of the cave was dated measuring $^{230}Th/^{234}U$ in speleothems, with an age of 110,000 years BP (Bocherens *et al.*, 1999). Thermoluminescence dating has also been carried out on a calcite sample from the top of layer four, providing a date of 100,000 years BP (Huxtable and Aitken, 1992). These dates indicated a correlation to around the middle of MIS 5, but palynological correlations suggested that the site belonged to early MIS 5a (Bocherens *et al.*, 1999). In addition to these dates, a mandible was dated using gamma spectroscopy, which produced a date of 127 ka (+46 or -32 ka) also collected from layer 4, this would place the site at late MIS 5e. Due to this range of dates, the exact sub-stage of the *Cervid* bone sampled in this chapter is unknown, and therefore has been reported as 5a-e. Four *Cervid* bones were sampled from this site: two fallow deer, (SC85185, radius) and (SC8687, phalanx), excavated from layer 4, and two red deer bones from layer 5b (SC8293, phalanx) and (SC84 109, lower tooth bud). The samples from these bones were powder and small fragments.

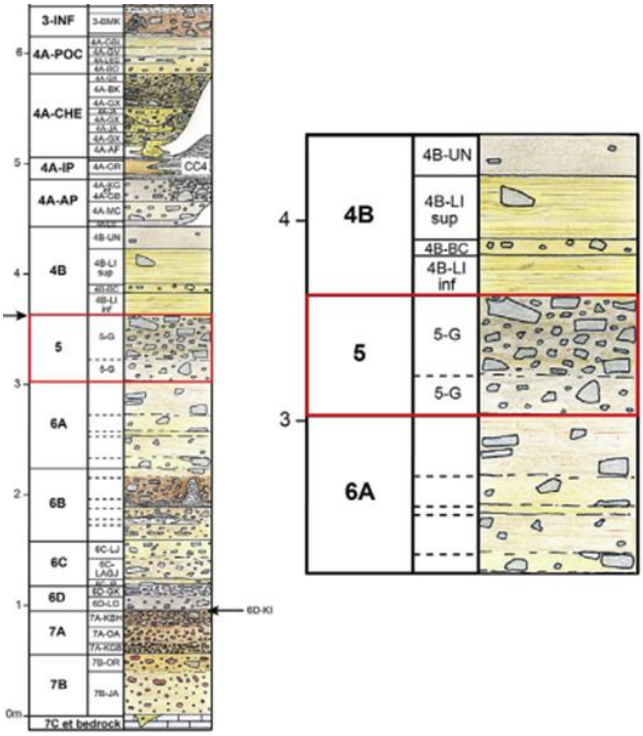


Figure 68: Stratigraphic sequence of Scladina Cave looking in detail at Unit 5 (taken from Abrams *et al.*, 2014), the samples analysed from this site were collected from units 4 and 5b, with Neanderthal artefacts found throughout the sequence, including units 1, 4 and 5.

7.1.1.5 Joint Mitnor Cave (UK)

Joint Mitnor Cave (Figure 69) was discovered in 1939 and excavated from 1939-1941 (Sutcliffe, 1960). Located in Higher Kiln Quarry, Buckfastleigh, the site has given its name to a vertebrate assemblage, with the most distinctive element being the presence of *Hippopotamus*, dating the site to the last Interglacial Period (Currant and Jacobi, 2001). As well as *Hippopotamus*, *Bovid*, fallow and giant deer were also found. It had been noted that the absence of horse in this assemblage is indicative of later stage 5. The site has been dated with consistent radiometric dating and fauna correlating to 120 ka BP (Schreve 2001). This site has now been correlated with MIS 5e (O'Connor and Lord, 2013). A piece of bovid long bone (Figure 66) excavated from the test pit was sampled from this site.

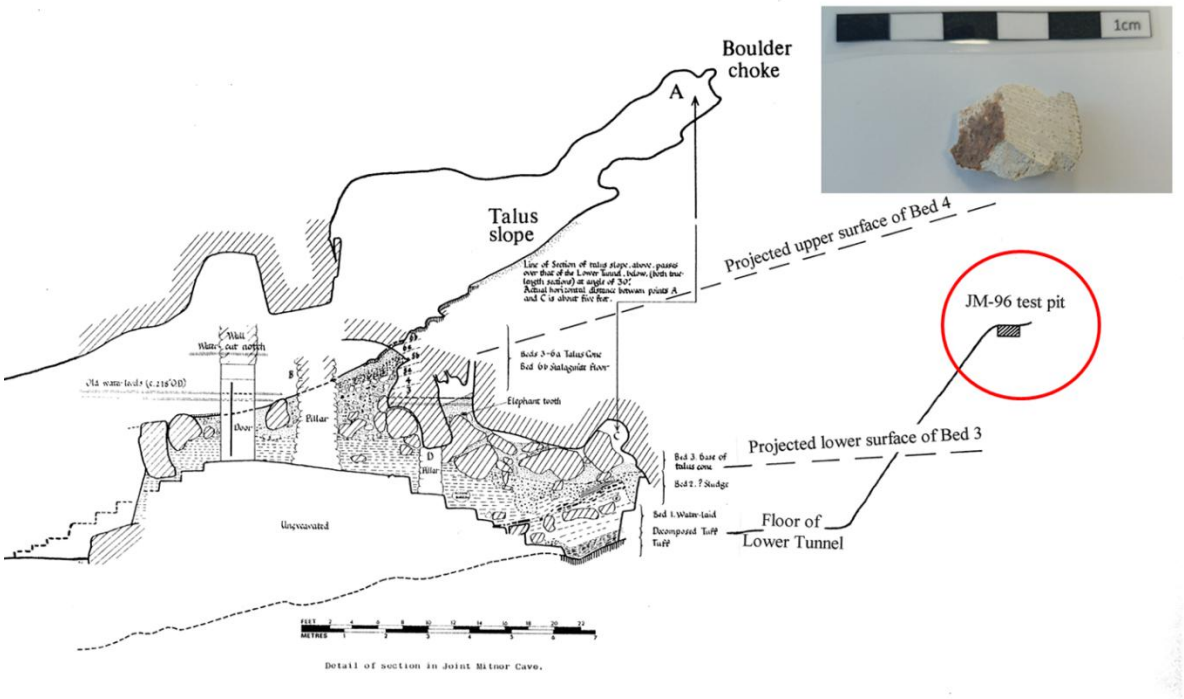


Figure 69: Site section and picture of a piece bovid bone (long bone) which was excavated from the test pit at Joint Mitnor Cave. Site map provided by Andrew Chamberlain. The bison bone shown here has been correlated with MIS 5e and thought to be ~123 ka BP

7.1.1.6 Victoria Cave (UK)

Victoria Cave is located north east of Settle in North Yorkshire. The cave was excavated from 1870-1878 (Gascoyne *et al.,* 1981). Two distinct fossiliferous horizons were found at the site, with the lower cave containing mammalian remains consistent with those present during the last interglacial (Ipswichian) and have been placed at or before ~120,000 years ± 6,000 years BP (Candy and Schreve, 2007). Animals found within the lower cave basin include: hyena, *Hippopotamus* and narrow-nosed rhinoceros. The presence of *Hippopotamus* fauna is particularly useful when dating Pleistocene sites in England. This species is commonly referred to as an 'indicator species' as in Britain it has only been found in post-MIS 12 deposits that date to the last Interglacial (Schreve, 2001; Candy and Schreve, 2007). Speleothems, close to or adjacent to fossil bone and teeth at the floor of the cave were analysed by measuring the ratio of $^{230}Th/^{234}U$. Initially this produced dates between 71,000 – 131,000 years BP (Gascoyne *et al.,* 1981). It was thought that the uranium rich fossil material may be leaching into the speleothems, and further sampling was carried out, between 1.5 and 2 cm from the fossil material (Gascoyne *et al.,* 1981). These samples produced an older, narrower age range of 114,000 – 135,000 years BP. The site was therefore correlated with the Ipswichan interglacial and therefore sub-stage MIS 5e (Gascoyne *et al.,* 1981). From this site a piece of bovid bone was provided for analysis (Figure 70).

Figure 70: The piece of *bovid* bone excuvated from Victoria Cave, Yorkshire, (UK). This sample has been correlated with MIS 5e and thought to be ~123 ka BP

7.1.1.7 Neumark-Nord (Germany)

Neumark-Nord, situated in the Geisel valley close to Merseburg, is famous for its paleontological and archaeological remains (Brühl and Mania, 2003). The Palaeolithic site is located east of Leipzig, and consists of several lake basins, three of which will be discussed here (Neumark-Nord 1, 2 and 3). The first of the basins (Neumark-Nord 1) was discovered in 1985 by geologist Matthias Thomae (Brühl and Mania, 2003), during which time the area was being used for lignite mining. Neumark-Nord 1, excavated between 1985 to 1996 (Gaudzinski-Windheuser *et al.*, 2013), is the largest of the three sites, measuring approx. 600 metres square (Brühl and Mania, 2003). The site was found to contain artefacts from a late middle Pleistocene interglacial lake, the shore areas of the lake containing evidence of human activity. Towards the end of the excavation of Neumark-Nord 1, another series of lake deposits was found approx. 400 meters to the east, this site was named Neumark-Nord 2. In 1998 bone and other artefacts were found (Brühl and Laurat , 2010) with the site correlated with MIS 5e (Gaudzinsk-windheuser and Roebroeks, 2014). Neumark-Nord 3 is the oldest of the lake basins and is thought to be approximately 400,000 years old, containing typical Lower Palaeolithic artefacts. A piece of forest elephant bone (long bone) (Figure 71) was sampled from Neumark-Nord, but unfortunately it is not known whether this bone was excavated from site 1 or 2, therefore the bone could date from either MIS 5 or 7.

Figure 71: A piece of forest elephant bone excavated from one of the Neumark-Nord sites in the Geisel valley (Germany). It is not known whether this bone was excavated from site 1 or 2, therefore the bone could date from either MIS 5e or 7, with dates of ~123 ka, or ~243 ka respectively

7.1.1.8 Ilford (uphall pit assemblage)

The Ilford mammalian assemblage, also described in the literature as 'The uphall pit assemblage', was first discovered in 1824 by workmen during a clay excavation (Schreve, 1998). At the site a number of mammalian remains were found including bones of rhino, horse, red deer, and most famously mammoth. There has been some debate about the age of the site. It was originally thought that the site was Ipswichian in age (West *et al.*, 1964), however it is now largely accepted that the site correlates with MIS 7 (Schreve 1998; Bridgland *et al.*, 2003; Penkman *et al.*, 2013). Ilford is described as having distinctive mammalian remains which have been correlated to MIS 7, such as the primitive 'Ilford type mammoth' along with supporting terrace stratigraphy and molluscan biostratigraphy (Schreve, 1998).

7.1.1.9 Schöningen (Germany)

Schöningen is a well-known site, located between Hanover and Berlin, ~ 10 km from Helmstedt. The site became famous with the discovery of a range of wooden spears and hearths which were estimated to be ~ 400 ka and claimed to be the first evidence for hominins hunting, as well as the first evidence of controlled use of fire (Thieme, 1997). Investigators from the University of Tübingen later found that the site was more likely to be ~ 300 ka (Balter, 2014), with the site sedimentary sequence correlated to the Reinsdorf interglacial of around 300 ka (Kuitems *et al.*, 2015). However it was also suggested by Balter, (2014) that the red debris which Thieme reported as hearths could be reddening due to precipitated iron compounds, caused by the rising and falling of the lake waters (Balter, 2014). The spears and other objects such as bones are thought to be well preserved due to the anaerobic sediment environment, with the artifacts remaining underwater until the present day (Lang *et al.*, 2012). From this site five bovid bones were sampled in total, with B28 identified more specifically as bison. Two of the bones were metatarsus fragments (B04 and B19), two were metacarpus fragments (B28 and B29) and the last sample B30,

was a metapodium fragment. The Schöningen samples analysed in this thesis are from the Schöningen 13 site (Figure 72), in channel II of the sequence. Samples B4 and B28 were excavated from dispositional phase 4, most of the large mammalian remains were collected from channel II of this phase. The remaining three samples (B19, B29 and B3 were collected from channel II of dispositional phase 3.

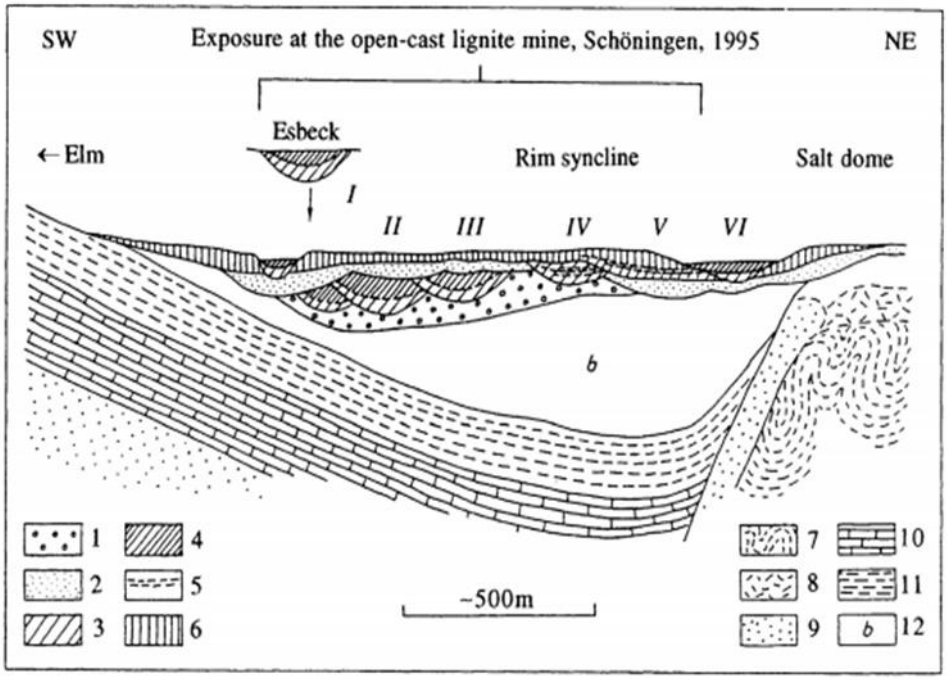


Figure 72: Figure taken from van Kolfschoten (2014). Schematic section through the Quaternary sedimentary cycles Channel I-VI of open-cast lignite mine Schöningen (modified after Mania, 1995). Key: 1) Elsterian glacial deposits; 2) Saalian glacial deposits; 3) Lacustrine deposits; 4) Limnic telmatic sequences; 5) Soil complexes; 6) Loess deposits; 7) Evaporites; 8) Gypsum cap-rock; 9) Buntsandstein; 10) Triassic limestone (Muschelkalk); 11) Triassic deposits (Keuper); 12). Palaeogene deposits. Samples analysed in this thesis were collected from channel II of dispositional phases 3 and 4.

7.2 Methods

The methods described in Chapter 2 for extraction, digestion and mass spectrometric analysis, were performed on all of the sub-fossil samples in this chapter. In brief, collagen was extracted from bone chips/powder by warming in 50 mM ammonium bicarbonate pH 8.0. After the protein extraction the samples were digested into shorter peptides and analysed using matrix-assisted laser desorption/ionisation mass spectrometry. The spectra obtained were processed using Data analysis 4.2, (Bruker) and the α-values calculated as described in Chapter 2. In order to investigate the reproducibility of values obtained, each sample was sub-sampled as many times as the sample size allowed, with three analytical replicates of each sample performed. The number of technical replicates of samples ranged from 1-15 depending on the amount of sample available. On average

a minimum of 30 mg of bone was used for each technical replicate. The samples from each site and their replicates are detailed in Tables 13 (Yukon samples) and 14 (all other samples).

Table 15: Samples collected from the following sites, Joint Mitnor Cave, Stump Cross, Victoria Cave, Nuemark-Nord, Scladina, Schöningen, and Ilford

Site Name	Sample ID	Sample replicates	Sample type and size of each sub-sample	Tools sampled with
Joint Mitnor Cave	Joint Mitnor Cave	1 bone sub-sampled 12 times (labelled A-L)	~ 30 mg of bone powder	Electric drill
Stump Cross	Stump Cross	1 bone sub-sampled 12 times (labelled A-L)	~ 30 mg of bone powder	Electric drill
Victoria Cave	Victoria Cave	1 bone sub-sampled 12 times (labelled A-L)	~ 30 mg of bone powder	Electric drill
Tanner Row	Tanner Row	1 bone sub-sampled 3 times (labelled A-C)	bone chips ~30 mg	Water cooled and saw, pliers
Neumark – Nord	Neumark – Nord	1 bone sub-sampled 2 times	bone chips ~30 mg	pliers
Scladina	SC15000	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Scladina	SC15100	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Scladina	SC18900	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Scladina	SC19200	1 bone, whole sample used	Bone powder ~ 10 mg	Metal spatula
Schöningen	B13	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Schöningen	B19	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Schöningen	B28	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Schöningen	B29	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Schöningen	B30	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Schöningen	B4	1 bone, whole sample used	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford01	1 bone, sub-sampled 3 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford02	1 bone, sub-sampled 4 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford03	1 bone, sub-sampled 4 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford04	1 bone, sub-sampled 4 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford05	1 bone, sub-sampled 2 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford06	1 bone, sub-sampled 2 times	Bone powder ~ 10-30 mg	Metal spatula
Ilford	Ilford07	1 bone, sub-sampled 2 times	Bone powder ~ 10-30 mg	Metal spatula

7.3 Results

7.3.1 Yukon samples

13 bone samples in total were analysed from a permafrost site in the Yukon region of Canada. In each sample the α-values were calculated for 12 glutamine-containing peptides (Figure 73). The α-values in this sample set ranged from 1.00 to 0.03, with the three heaviest peptides (*m/z* 3001, 3100 and 3665) showing the lowest α-values, and the lightest peptides (*m/z* 836 and 1105) showing the highest α-values. Peptides with *m/z* values of 1105, 2689, 2705 and 3100 were the most commonly observed peptides across all 13 samples.

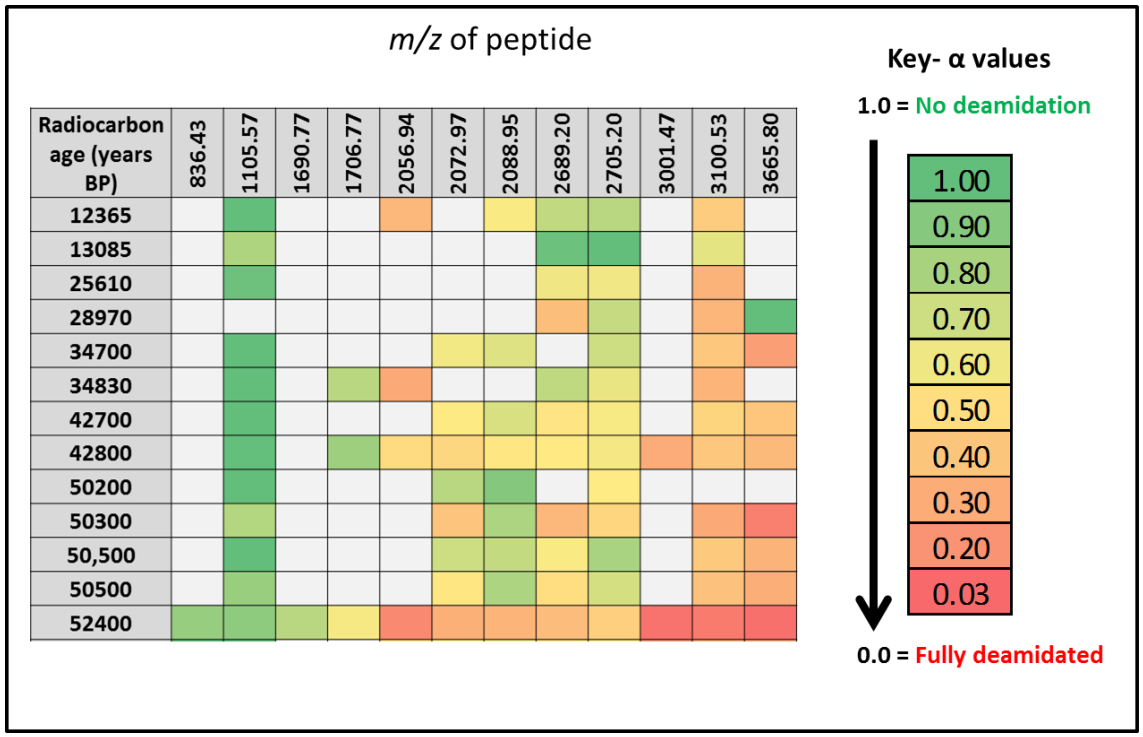


Figure 73: α-values in collagen, extracted from bone chips from Yukon with radiocarbon ages (Cal years BP) ranging from ~12,000 - 52,000 years BP. The errors obtained for these dates are reported in table 2 (It should be noted that some of the radiocarbon dates obtained are at the limits of radiocarbon dating, around 50,000 years BP, and threfore it is possible these samples are older, belonging to the surrounding dated ash layers (80,000 years BP)

Although the peptide with *m/z* of 1105 was observed in 12 out of the 13 samples (Figure 74, little deamidation was observed, with the lowest α-value of 0.87, obtained from two samples with radiocarbon dates of 13,085 and 50,300 years (CAL) BP. Peptides with *m/z* values of 3100 and 2705 showed a gradual decrease in α-values with radiocarbon age (Figure 74).The α-values in *m/z* 3100 were lower than those measured in *m/z* 2705 with ranges in α-values of 0.78- 0.13 and 1.0 – 0.56 respectively. α-values obtained for *m/z* 3665 ranged from 1.0-0.07. It should be noted that the α-value of 1.0 in peptide *m/z* 3665 obtained for a sample radiocarbon dated to 28,970 years BP does not fit with the rest of the values obtained, including the value for this particular peptide obtained

from the relatively young Tanner Row sample (~800 -1 000 years BP). The reason for this unusually high α- value is unknown, as even contamination with modern material is unlikely to give such a high α- value, as this peptide has been shown to deamidate quickly (see Chapter 6, analysis of Tanner Row bone).

From the peptides investigated here (Figure 73), peptide *m/z* value 2705 appeared to produce α-values that correlated most with the age of the sample; however a range of α-values (0.59-0.88) (Figure 74, right) were obtained for samples which were radiocarbon dated to around 50,000 years BP. As the accurate age of samples with a ^{14}C date of > 50,000 years BP are in question, the data was plotted using the radiocarbon dates of all samples (Figure 74 left), and also using the assumed date of 80,000 years BP for the samples radiocarbon dated as >50,000 ^{14}C years BP (Figure 74 right).

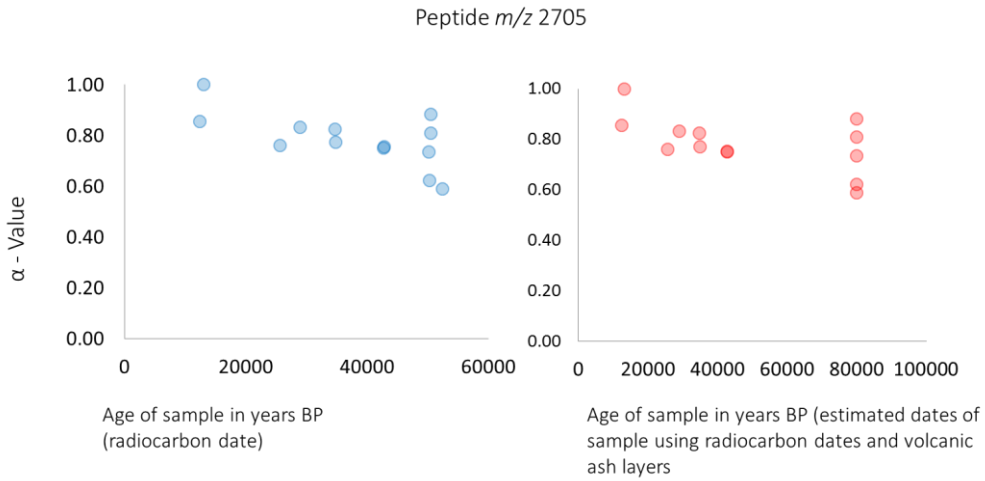


Figure 74: α- values obtained from peptide *m/z* 2705 observed in tryptic digests, of protein extracted from 13 bone samples, from the Yukon region Canada. These values are plotted against the radiocarbon ages of the samples (left), and against radiocarbon age (samples > 50,000 years BP) and dates obtained from the surrounding volcanic ash layers (samples found to be >50,000 ^{14}C years BP) (right)

Before obtaining the radiocarbon dates for these samples, it was thought that they were all ~ 80,000 years BP, due to the dates obtained from the surrounding volcanic ash layer (Westgate *et al.,* 2008). This highlights the difficulty in dating samples thought to be older than or close to the radiocarbon limit (~ 50,000 years BP). Radiocarbon dating is clearly useful for identifying younger samples in mixed deposits, as shown in this sample set. However, if the older samples here are ~ 80,000 years BP, then they should be radiocarbon dead. This demonstrates the importance of using multiple dating techniques for samples thought to be ~ 50,000 ^{14}C years BP. However, throughout the rest of this section the samples reported will be plotted against their radiocarbon age.

In order to further investigate the reproducibility of the α-values obtained, four of the samples were sub-sampled and analysed a further ten times (QZ-10-200, QZ-10-346, EL-10-50, EL-10-17)

(Figure 75). These samples were chosen for further sub-sampling in order to cover the range of ages, as well as investigating two of the older samples dated to ~ 50 ka BP which gave different α-values of 0.59 and 0.88. Here peptide *m/z* 2705 is investigated as it was the most frequently observed peptide that appeared to correlate with chronological age. Replicate analysis of the four samples resulted in variation which didn't appear to increase with the age of the sample, with three of the four samples chosen for sub-sampling producing similar standard deviations (0.4, 0.4 and 0.5 for samples EL-10-17, QZ-10-220, and QZ-10-346). Sample EL-10-50 showed the highest error in α-values, with a standard deviation of 0.32 (Figure 76). In one of the replicates of EL-10-50, peptide *m/z* 2705 was not observed with an α-value of 0.0 was measured in two of the nine replicates.

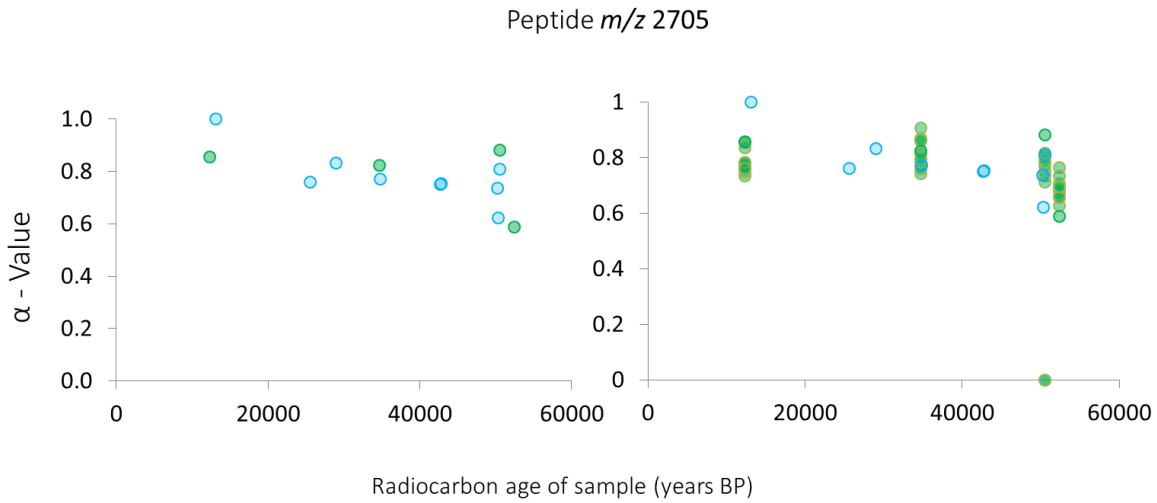


Figure 75: α- values obtained from peptide *m/z* 2705 observed in tryptic digests, of Protein extracted from 13 bone samples, from the Yukon region, Canada. Four of these samples (QZ-10-200, QZ-10-346, EL-10-50, EL-10-17) (left, green points) were sub-sampled and analysed a further ten times each, in order to investigate the reproducibility of α- values (right, green points)

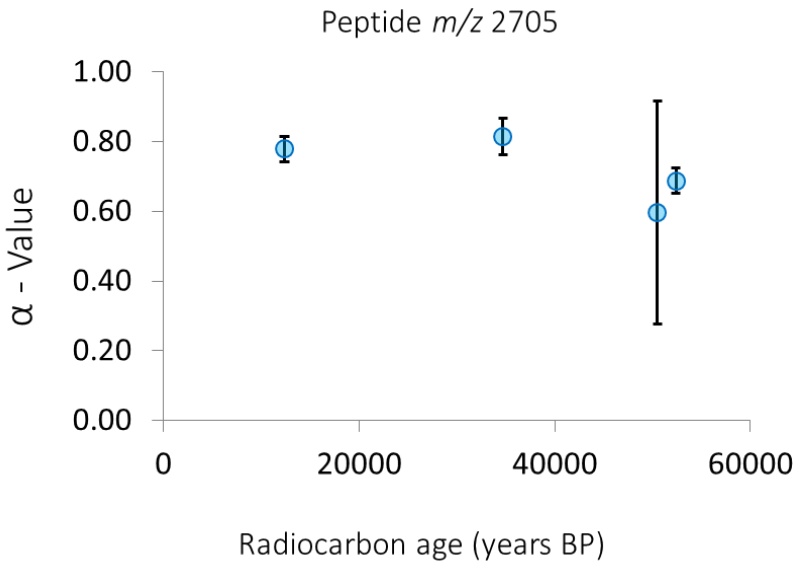


Figure 76: Average α-values from the ten replicates of four Yukon samples (QZ-10-200, QZ-10-346, EL-10-50, EL-10-17). Error bars show the standard deviation for each sample

Van Doorn *et al.* (2012) found that the deamidation in their samples correlated more with thermal age than chronological age, which is consistent with the recognised effect of temperature on deamidation (Sinex, 1957; Robinson *et al.,* 1970; Scotchler and Robinson, 1974; Stratton *et al.,* 2001; Joshi and Kirsch, 2004; Robinson, 2004; Joshi *et al.,* 2005). In order to investigate this, the thermal age of the Yukon samples were calculated based on the probability of damage as a function of temperature (www.thermal-age.eu). The level of deamidation measured in each sample was plotted against the chronological age (Figure 77) and the thermal age (Figure 78). In these particular samples thermal age did not appear to correlate more with deamidation than the radiocarbon age, with peptides *m/z* 3665 and 2705 having the same R^2 values when calculated using chronological age or thermal age (0.56 and 0.43 respectively). Peptide *m/z* 3100 had slightly higher R^2 values in samples plotted against thermal age than those plotted against chronological age (0.28 and 0.27 respectively). This is probably because the samples are all from the same site, and therefore have similar thermal histories. The application of thermal age is likely to be more useful when comparing samples from sites of different locations which may differ in terms of their thermal history.

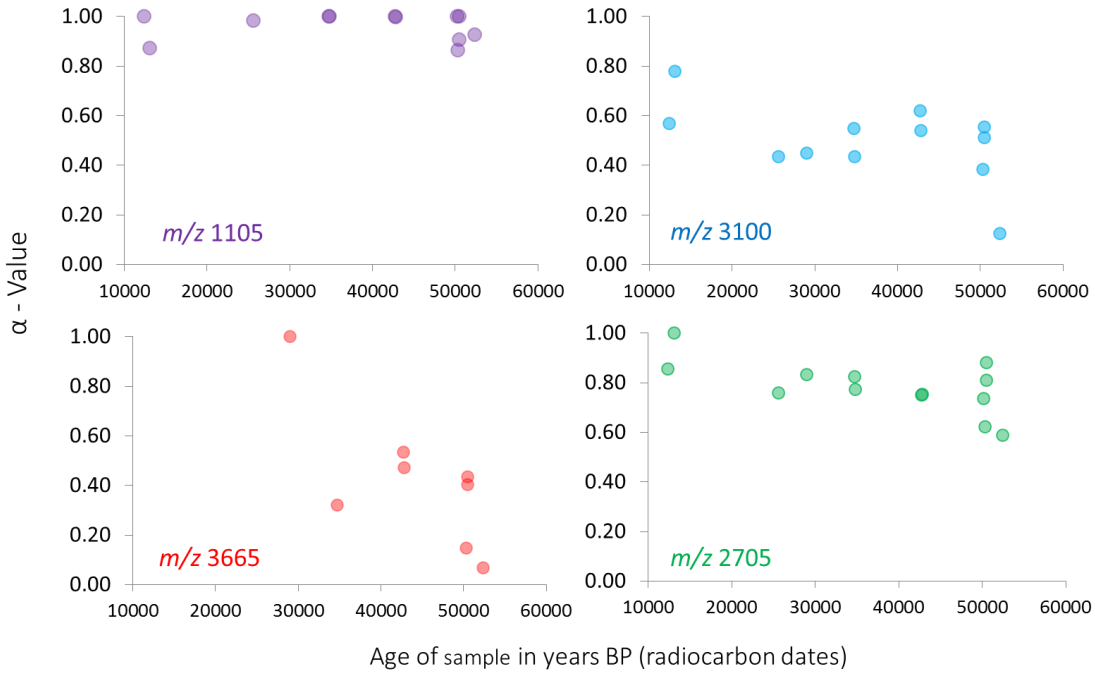


Figure 77: α- values obtained from peptides observed in tryptic digests, of protein extracted from 13 Yukon bone samples. Deamidation is shown for four peptides: *m/z* 1105 (top left), *m/z* 3100 (top right), *m/z* 3665 (bottom left), *m/z 2705* (bottom right), and plotted against the chronological age obtained from radiocarbon dates

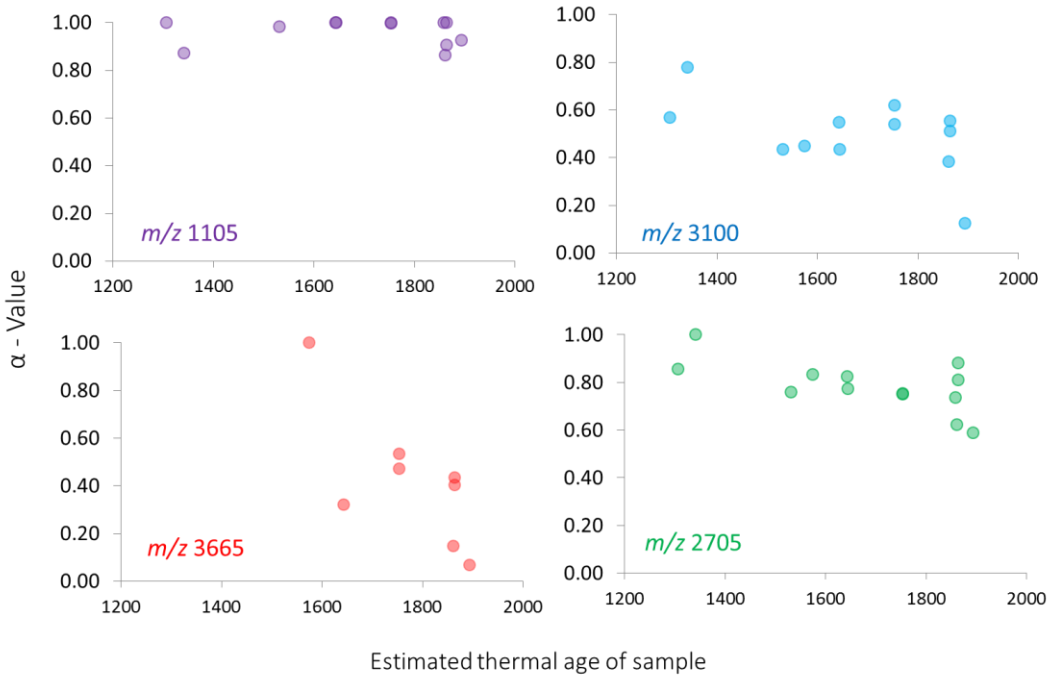


Figure 78: α- values obtained from peptides *m/z* 1105, 3100, 3665 and 2705 observed in tryptic digests, of protein extracted from 13 bone samples, from the Yukon region, Canada. α- values were plotted against the estimated thermal age calculated

The relationship observed between deamidation and age in these samples can be split into four groups:

1) peptides that were observed in four or less of the samples, resulting in too few data points to interpret (*m/z* 836, 1690, 1706, 2056,)

2) peptides which showed little deamidation, regardless of age (*m/z* 1105)

3) peptides in which deamidation measured weakly correlated with age (*m/z* 2705, 2689, 3665, and 2072)

4) peptides in which deamidation did not appear to correlate with age (*m/z* 2088 and 3001).

As levels of deamidation measured in bone from Yukon showed relatively low levels of deamidation, especially in more stable peptides such as *m/z* 1105, further sampling of older bone material was carried out.

7.4 Analysis of bone material collected from: Tanner Row, Yukon, Victoria Cave, Stump Cross, Scladina, Joint Mitnor Cave, Ilford and Schöningen

In order to further investigate the relationship between time/thermal age and glutamine deamidation, samples from nine sites in total were measured (Table 12). The samples analysed in this section have age estimates ranging from ~800 to 300,000 years BP. While glutamine deamidation in peptide *m/z* 2705 was useful for comparisons in the Yukon material (Figures 76 and 77), in samples from Schöningen and Joint Mitnor Cave, levels of glutamine deamidation were high in this peptide, with α=values of 0.0 measured in some samples (Figure 79).Therefore in this section the slower-deamidating peptide *m/z* 1105 is used. For sites that have dates which span a range of MIS sub-stages, the α-values are plotted against the mid-point of the estimated chronological age. Neumark-Nord has been plotted as 5e (~123 ka) but it is acknowledged that it may be older (MIS 7, ~243 ka).

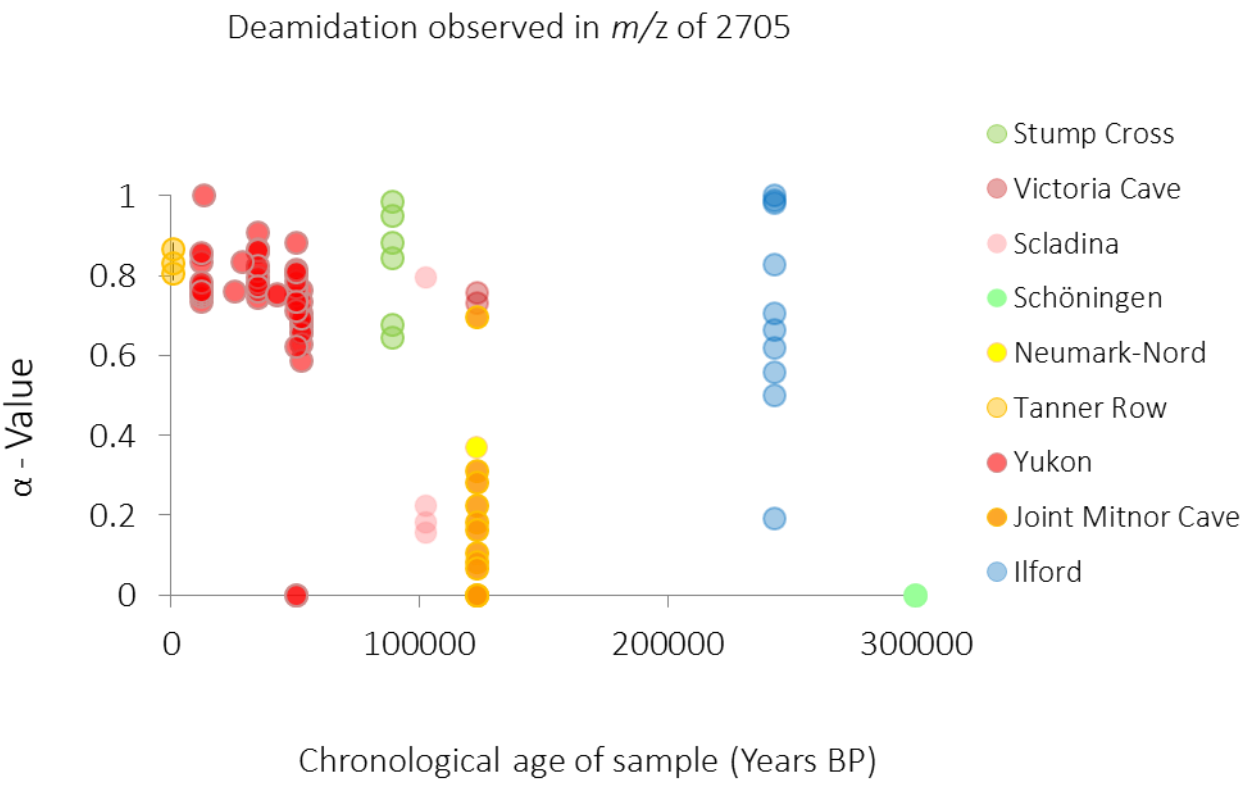


Figure 79: α- values obtained from peptide *m/z* 2705 observed in tryptic digests, of protein extracted from bone. For sites that have dates which span a rage of MIS sub-stages, the α-values are plotted against the mid-point of the estimated chronological age. Neumark-Nord has been plotted as 5e (~123 ka) but it is acknowledged that it may be older (MIS 7, ~243 ka)

As the nine sites differ significantly both in geographical location and age, the relationships between glutamine deamidation and both chronological and thermal age were investigated (Figure 80).

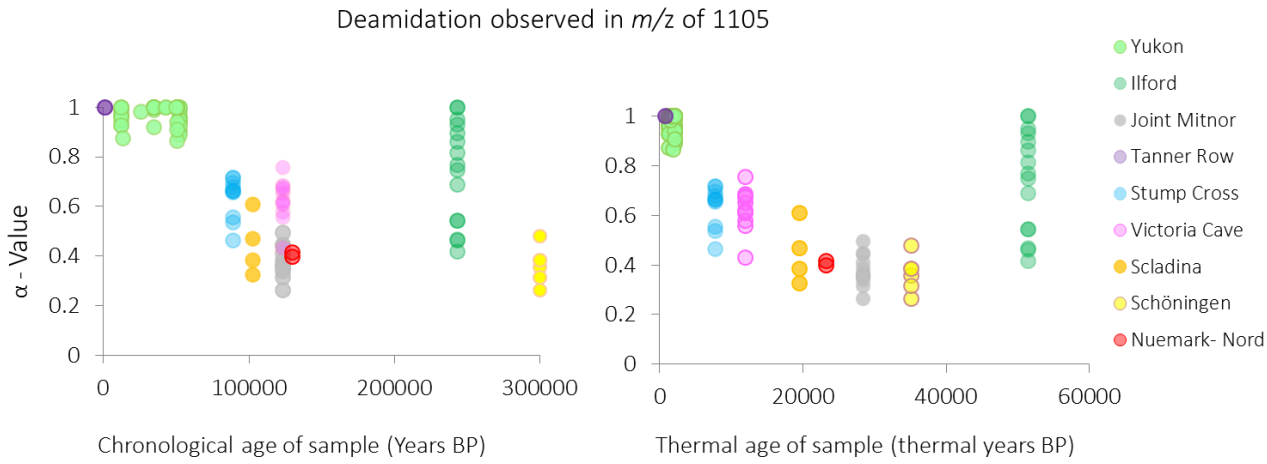
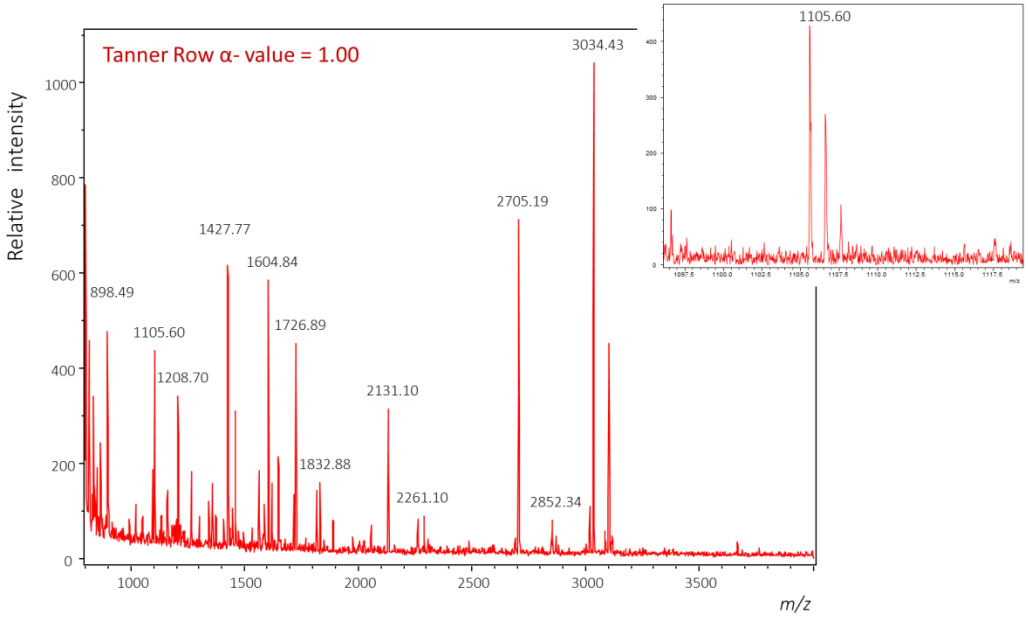


Figure 80: α- values obtained from peptide *m/z* 1105 observed in tryptic digests, of protein extracted from bone. For sites that have dates which span a rage of MIS sub-stages, the α-values are plotted against the mid-point of the estimated chronological age. Nuemark-Nord has been plotted as 5e (~123 ka) but it is acknowledged that it may be older (MIS 7, ~243 ka)(left). The corresponding thermal ages were also plotted (right)

It is clear that levels of glutamine deamidation in the samples presented here show a better correlation with the thermal age than chronological age. This is not surprising, as deamidation has already been shown to be temperature dependent. Taking into account the temperature of the burial environment over time helps to separate data from samples of similar ages, from different locations. However, with the level of variation observed in these samples, it would not be possible to estimate the age of a sample with any precision. Glutamine deamidation may be a useful tool to identify samples which are likely to be radiocarbon dead (i.e. older than ~ 50 ka BP), or to authenticate a genuine 'old' sample.

Peptide *m/z* 1105 measured in bones from Ilford showed a large variation of α-values. The α-values in these samples ranged from 0.42 to 1.00, with an average α-value of 0.74 and a standard deviation of 0.2. Ilford has been correlated with MIS 7 and therefore is likely to be ~ 243 ka BP; it has the highest calculated thermal age of the nine sites investigated in this chapter and therefore it is unlikely that no deamidation in peptide *m/z* 1105 has occurred. The spectra obtained from the Ilford samples were not high quality, in terms of the number of peaks observed. It is often difficult to clearly define parameters to distinguish 'good' from 'bad' spectra; especially when comparing biological extracts from different bones, excavated from different environments, and often with different species origins. Examples of 'good' spectra in this thesis are described in terms of the

number of peaks observed. Examples of 'good' spectra obtained from protein extracted from bones from two sites of different ages (Tanner Row and Joint Mitnor Cave) are shown in Figures 80 and 81. Spectra obtained for these samples contained over 30 peaks, and at least 25 of these had a S/N value of ≥ 6 (Figures 65 and 66). Currently, in order for a α-value to be calculated from a peptide distribution, a S/N of ≥ 6 is the minimum level required



:

Figure 81: Spectrum obtained from a tryptic digest of protein extracted from a piece of Tanner Row bovid bone. Collagen was extracted using 50 mM ammonium bicarbonate

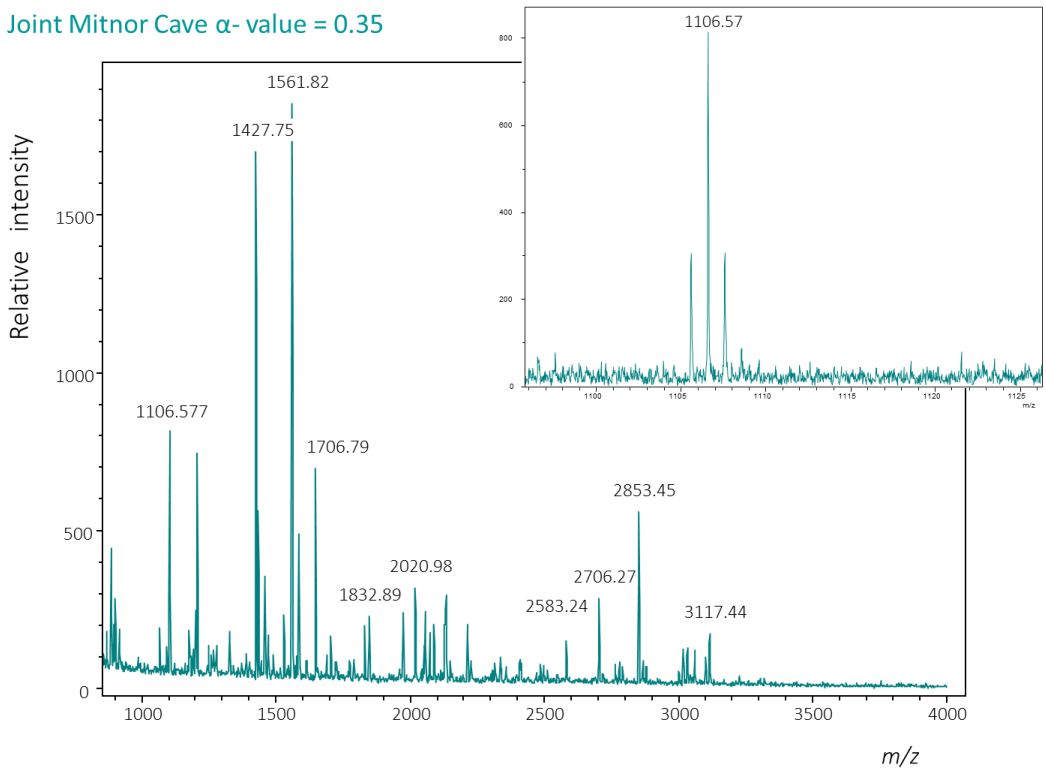


Figure 82: Spectrum obtained from a tryptic digest of protein extracted from a piece of Joint Mitnor bovid bone. Collagen was extracted using 50 mM ammonium bicarbonate

In addition examples of spectra obtained from the following sites: Joint Mitnor, Stump Cross, Victoria Cave, Scladina Cave, and Schöningen are shown in Figures 82-91.

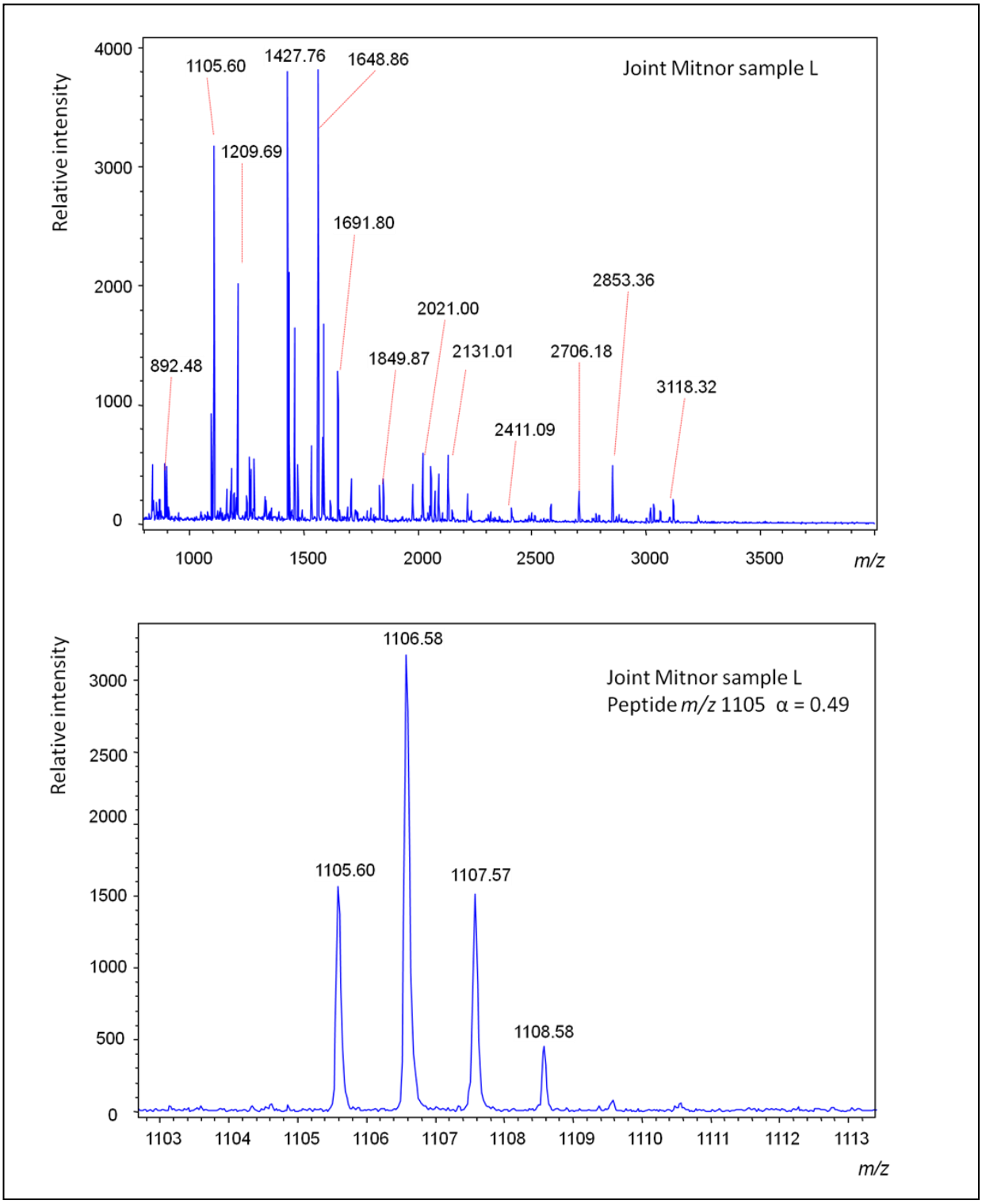


Figure 83: Spectrum obtained from a tryptic digest of protein extracted from a piece of Joint Mitnor bone (Joint Mitnor sample L). Collagen was extracted using 50 mM ammonium bicarbonate

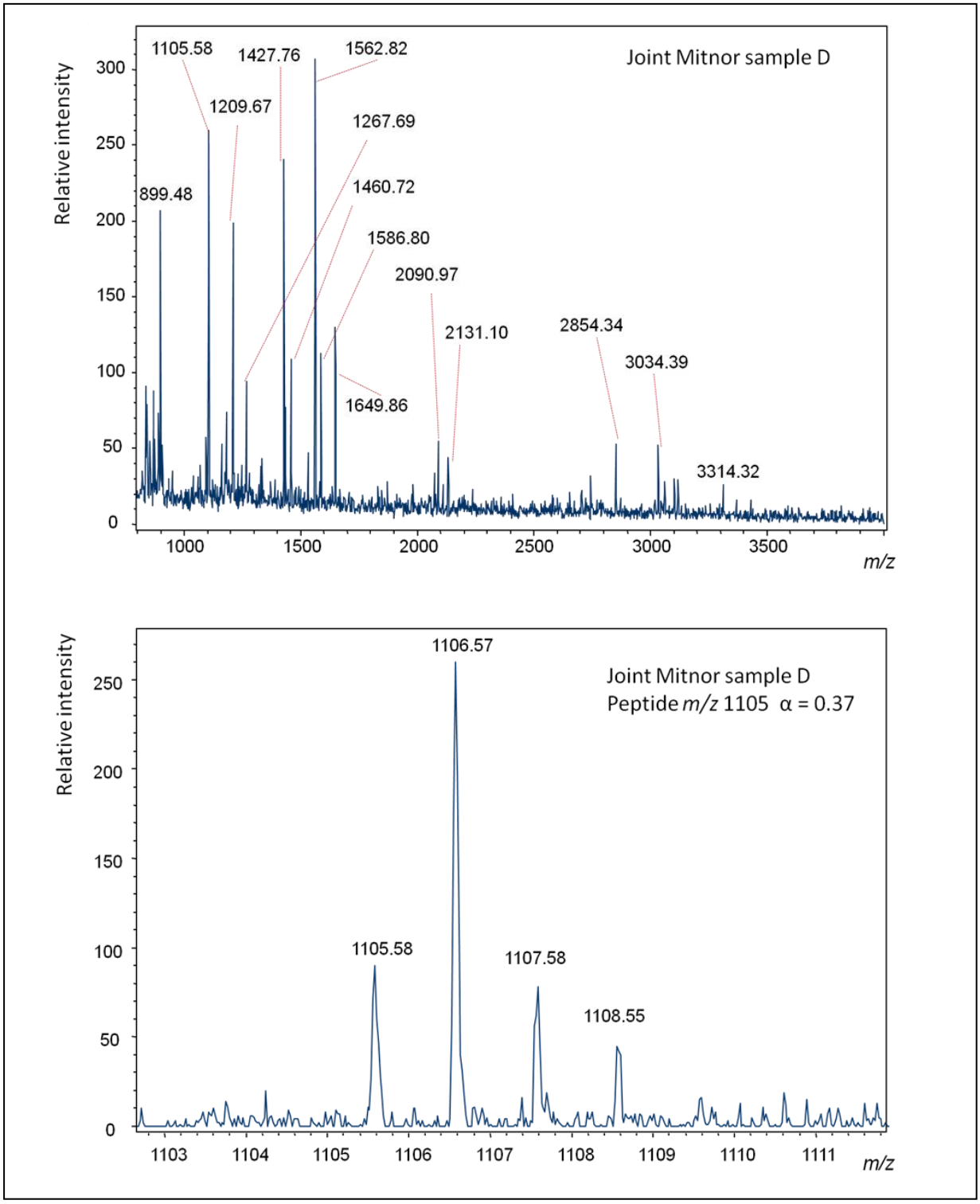


Figure 84: Spectrum obtained from a tryptic digest of protein extracted from a piece of Joint Mitnor (sample D) bone. Collagen was extracted using 50 mM ammonium bicarbonate

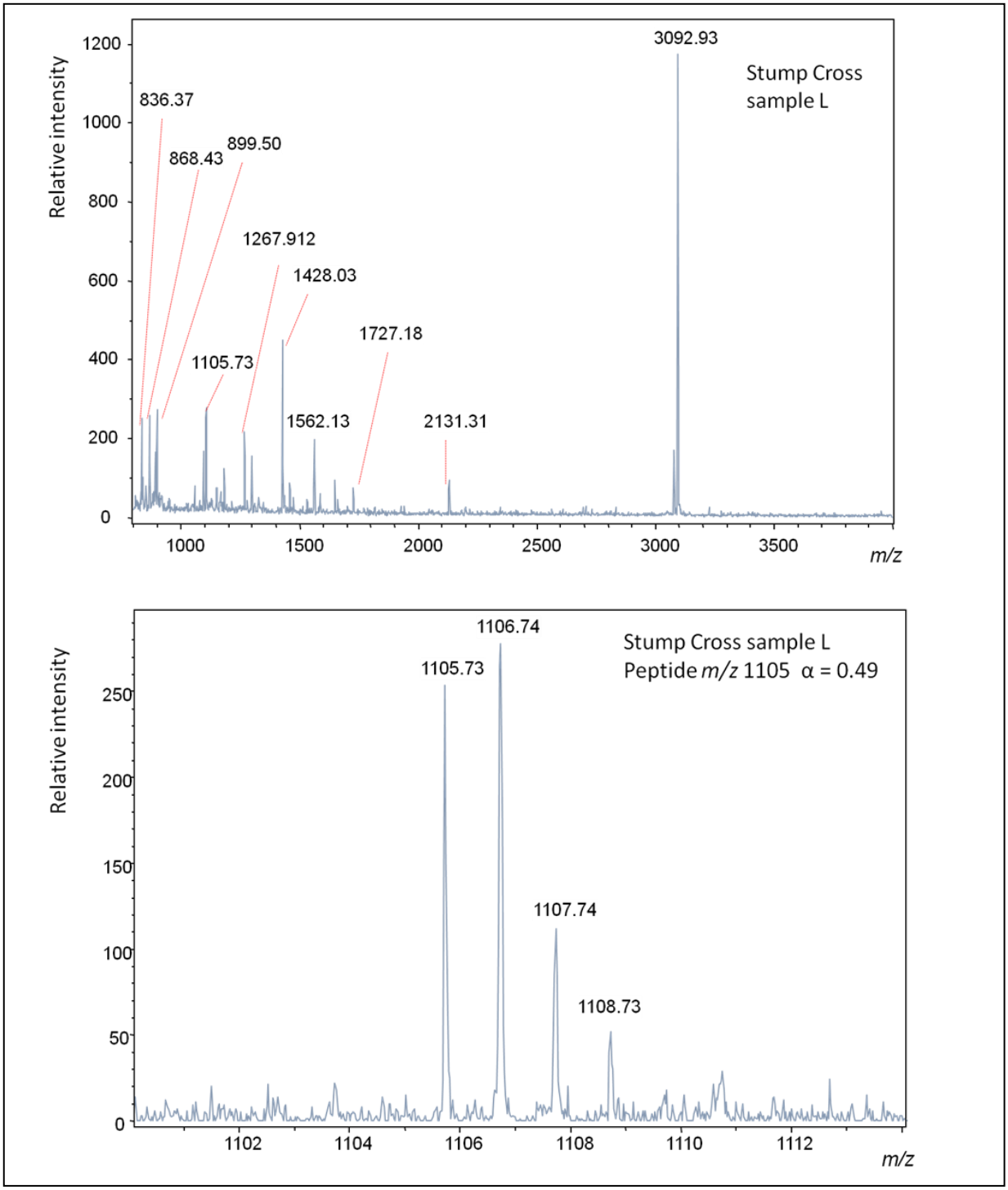


Figure 85: Spectrum obtained from a tryptic digest of protein extracted from a piece of Stump Cross (sample L) bone. Collagen was extracted using 50 mM ammonium bicarbonate

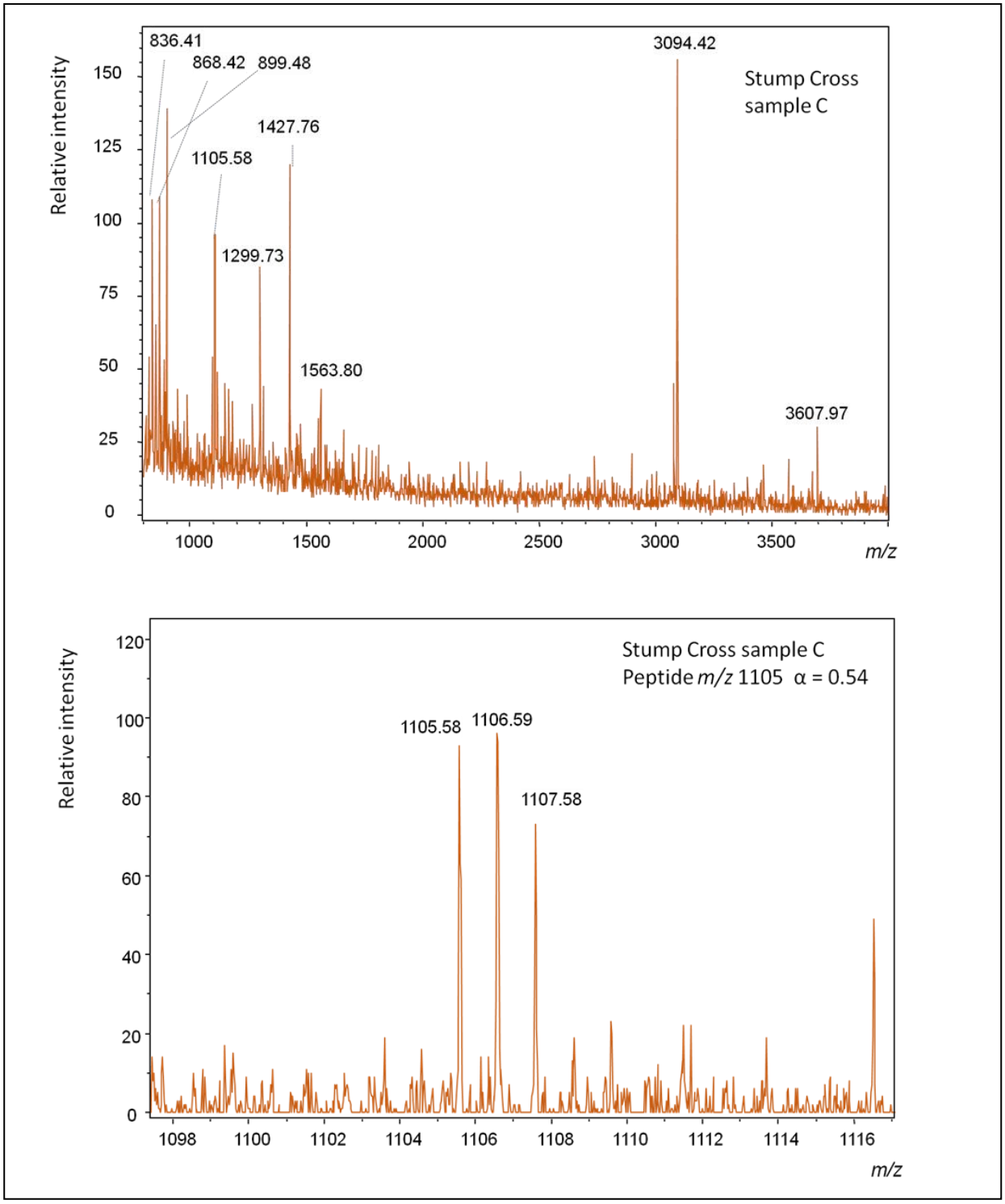


Figure 86: Spectrum obtained from a tryptic digest of protein extracted from a piece of Stump Cross (sample C) bone. Collagen was extracted using 50 mM ammonium bicarbonate

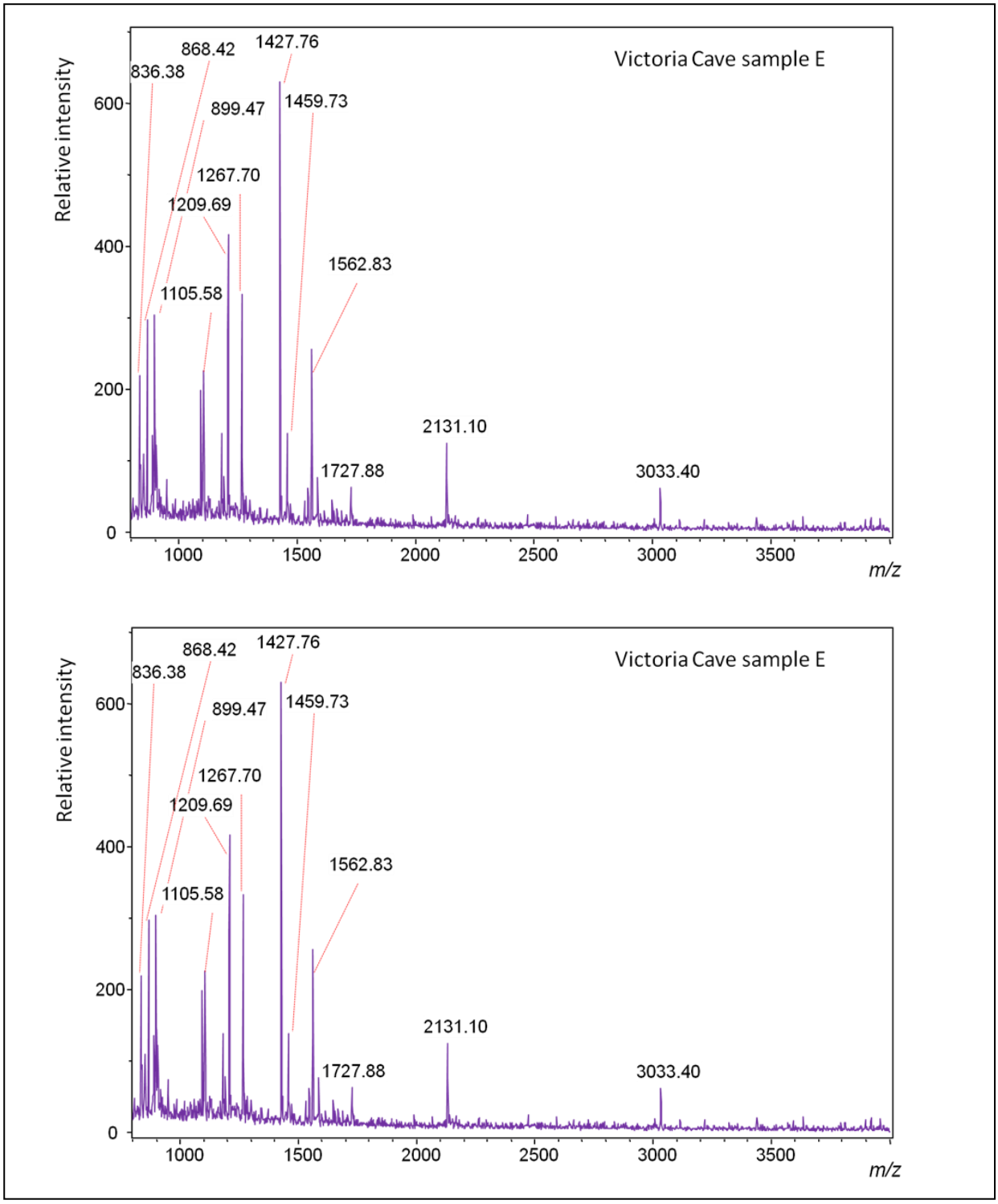


Figure 87: Spectrum obtained from a tryptic digest of protein extracted from a piece of Victoria Cave (sample E) bone. Collagen was extracted using 50 mM ammonium bicarbonate

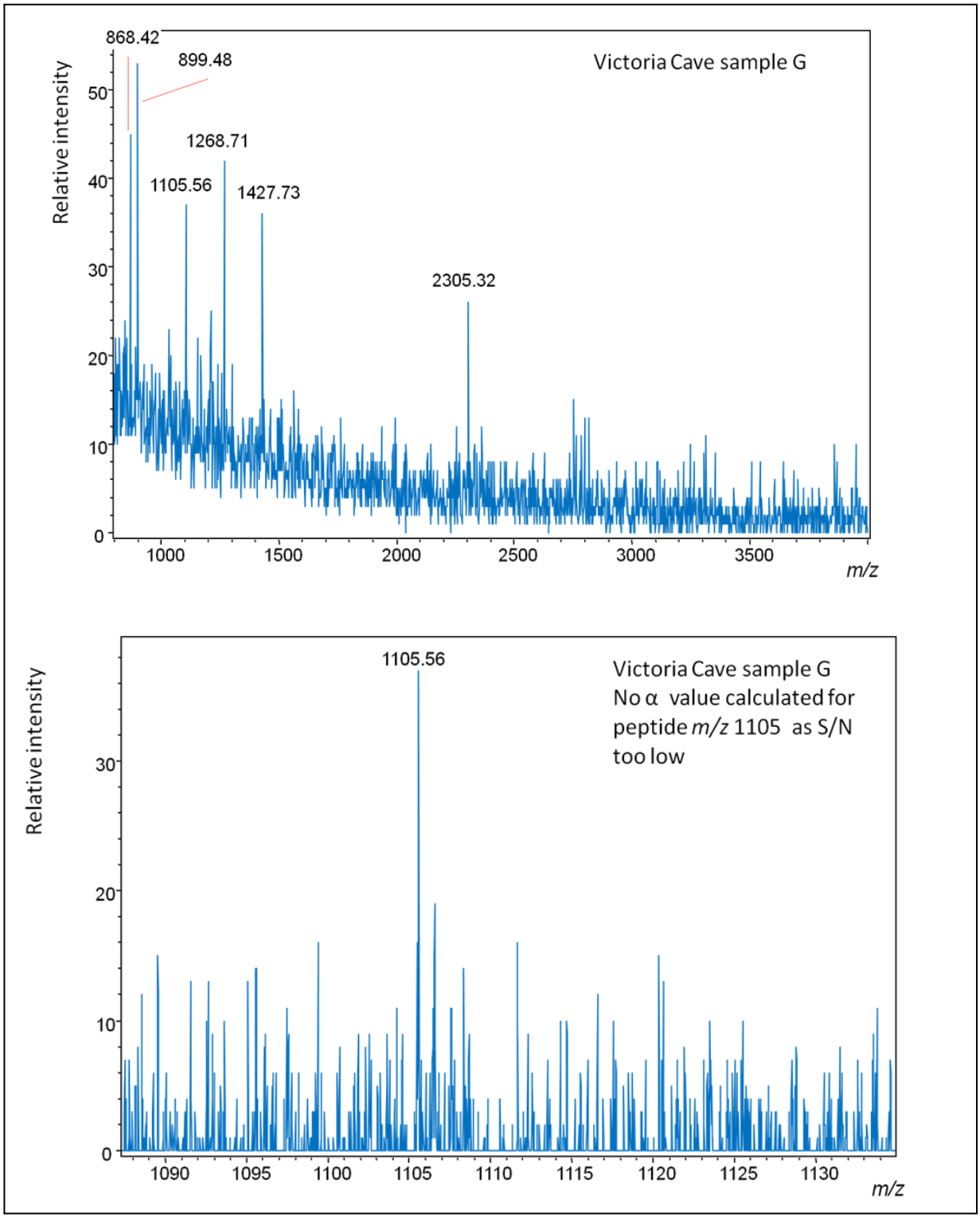


Figure 88: Spectrum obtained from a tryptic digest of protein extracted from a piece of Victoria Cave (sample G) bone. Collagen was extracted using 50 mM ammonium bicarbonate

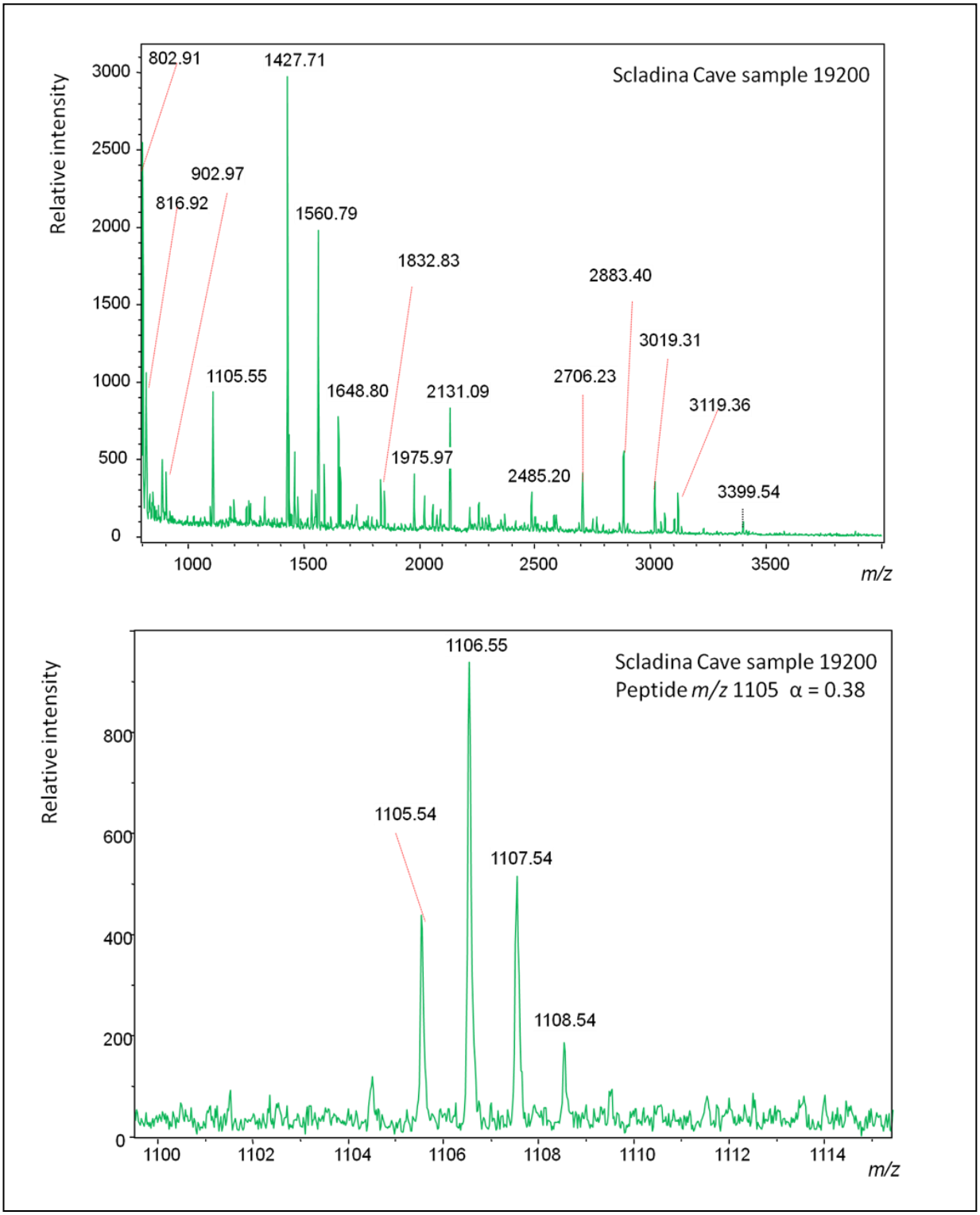


Figure 89: Spectrum obtained from a tryptic digest of protein extracted from a piece of Scladina Cave (Sample 19200) bone. Collagen was extracted using 50 mM ammonium bicarbonate

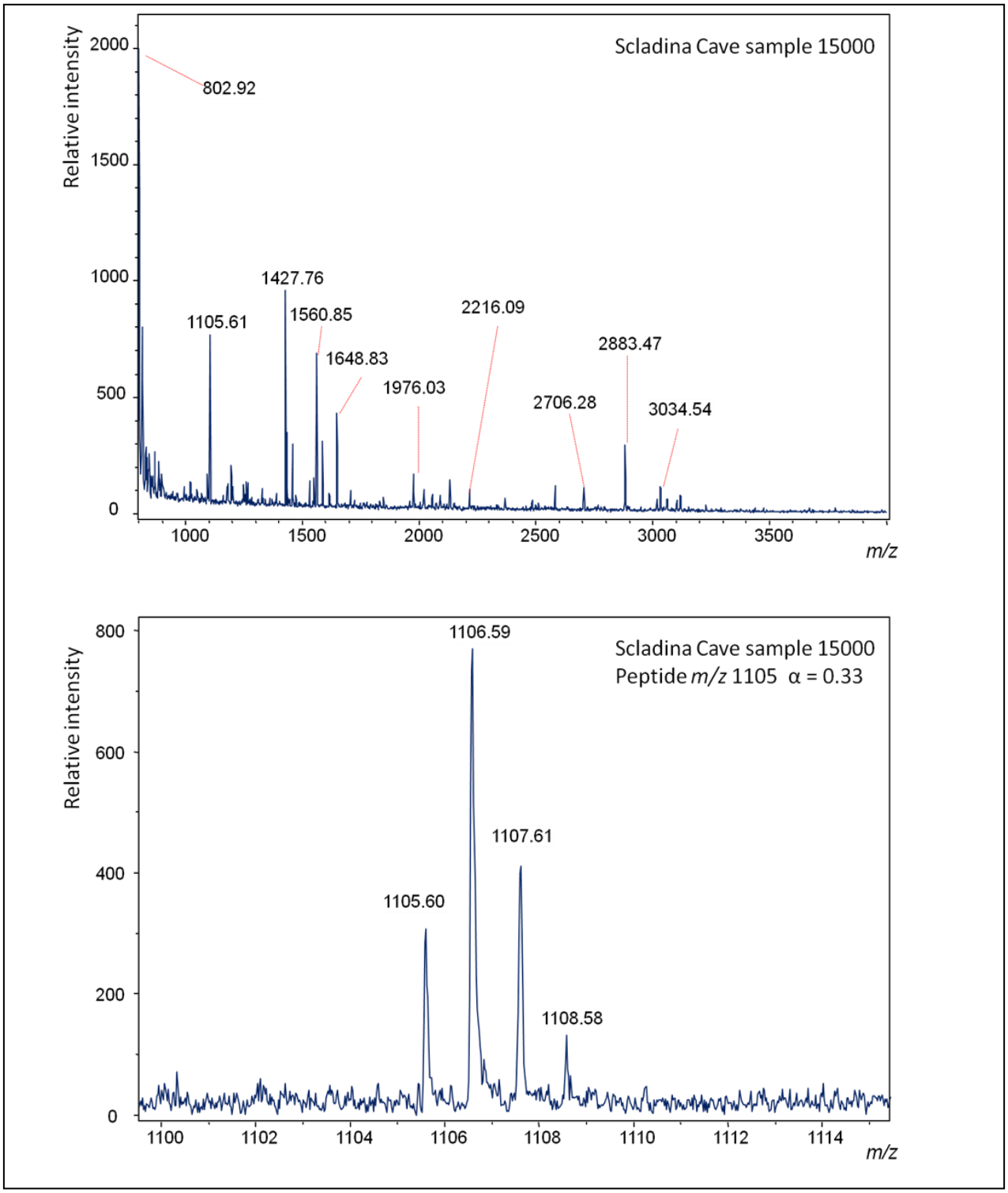


Figure 90: Spectrum obtained from a tryptic digest of protein extracted from a piece of Scldina Cave (sample 15000) bone. Collagen was extracted using 50 mM ammonium bicarbonate

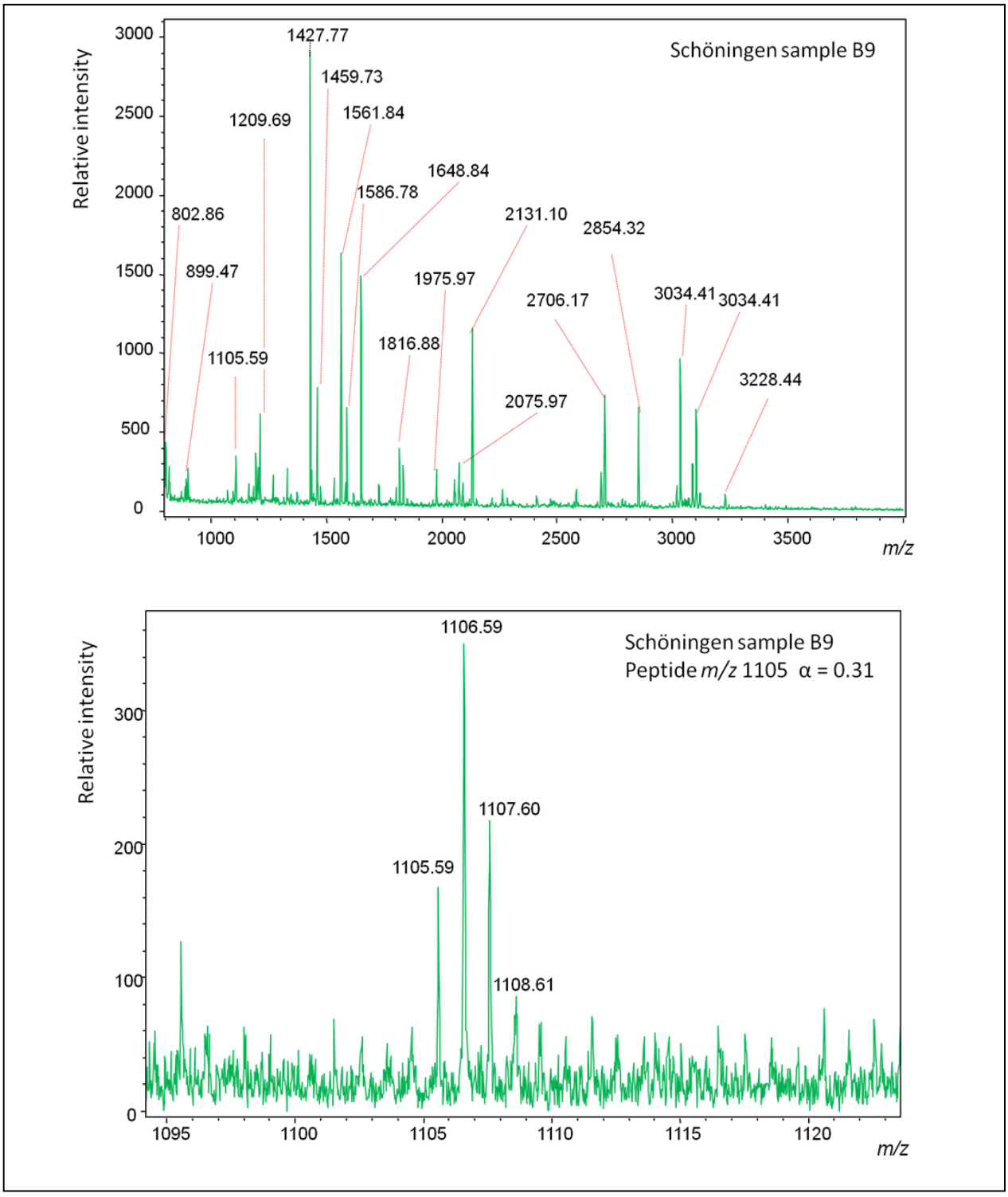


Figure 91: Spectrum obtained from a tryptic digest of protein extracted from a piece of Schöningen (sample B9) bone. Collagen was extracted using 50 mM ammonium bicarbonate

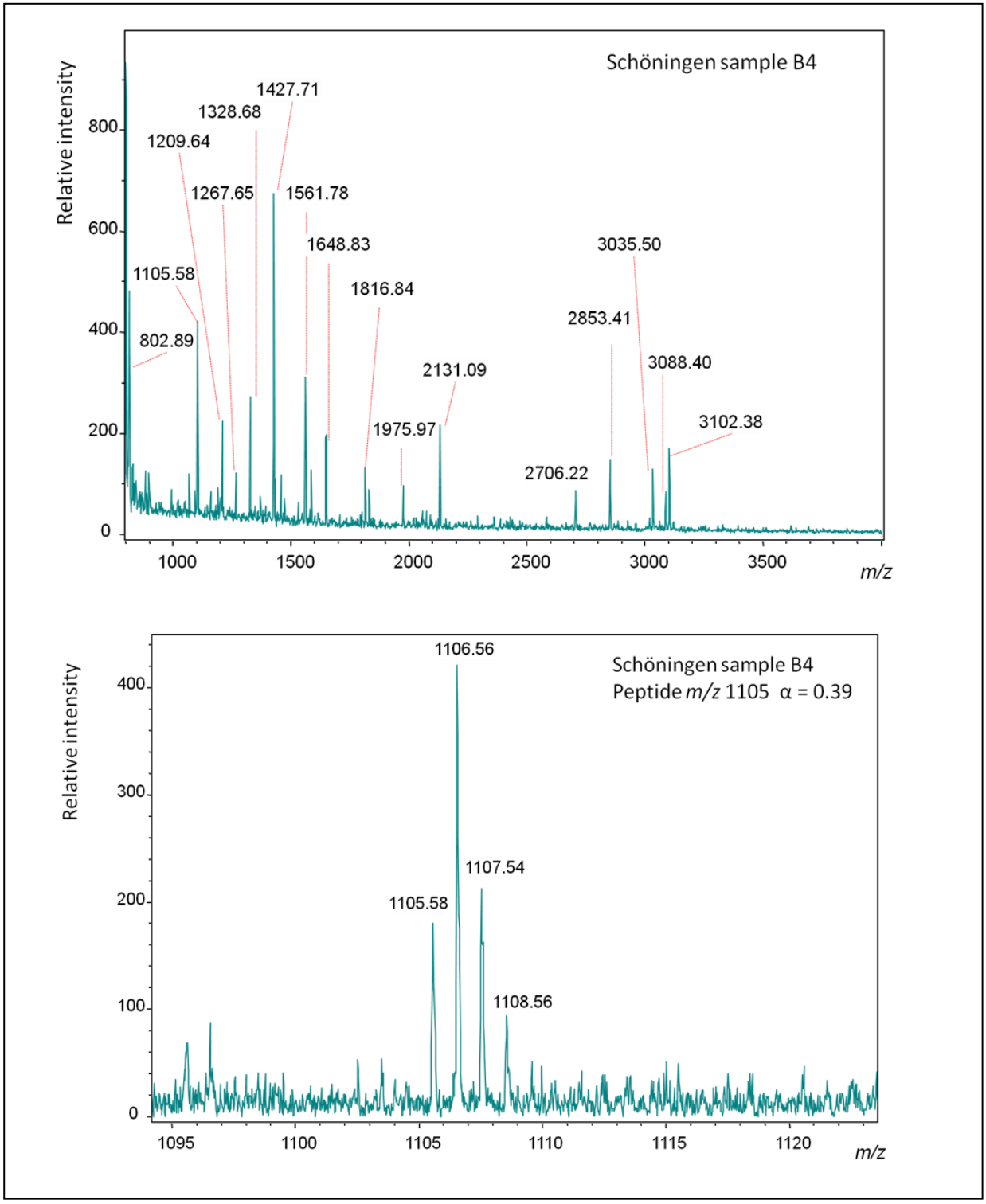


Figure 92: Spectrum obtained from a tryptic digest of protein extracted from a piece of Schöningen (sample B4). Collagen was extracted using 50 mM ammonium bicarbonate

The α-values obtained from the Ilford samples appeared abnormally high in comparison to the other eight sites. Therefore, the spectra obtained for the Ilford samples were investigated. For each spectrum, the following were recorded: the S/N ratio of *m/z* 1105, the total number of peaks in the spectrum with S/N ≥ 6, and the corresponding α-value calculated from the analytical triplicates (Table 15). Most samples from this site resulted in poor spectra with 12 of the 15 samples having fewer than 17 peaks, with a range of 6-17 peaks and an average of just 9. The α-values obtained for these samples ranged from 0.46 to 1.00. Two example spectra, one from each end of the range (sample with 6 (A), and 17 peaks (B)) are shown in Figures 92 and 93.

Table 16: Corresponding S/N ratios, and number of peaks present in spectrum with S/N ratios ≥ 6, with the corresponding calculated α-values.

Sample name	S/N of *m/z* 1105	No. of peaks with S/N ≥ 6	α-value	Sample name	S/N of *m/z* 1105	No. of peaks with S/N ≥ 6	α-value
Ilford01A.1	18.1	41	0.47	Ilford03A.1	13.9	14	0.75
Ilford01A.2	19.6	45	0.47	Ilford03A.2	nd	9	0.75
Ilford01A.3	15	31	0.47	Ilford03A.3	14.2	12	0.75
Ilford01B.1	6	7	1.00	Ilford03B.1	8.4	12	0.86
Ilford01B.2	6.9	8	1.00	Ilford03B.2	11	11	0.86
Ilford01B.3	16.6	22	1.00	Ilford03B.3	nd	9	0.86
Ilford01C.1	nd	6	0.90	Ilford03C.1	10.7	13	0.82
Ilford01C.2	5.4	6	0.90	Ilford03C.2	10.6	11	0.82
Ilford01C.3	5.4	6	0.90	Ilford03C.3	8.6	11	0.82
Ilford02A.1	7.1	6	0.90	Ilford03D.1	8.1	15	0.77
Ilford02A.2	*	4	0.90	Ilford03D.2	7.3	11	0.77
Ilford02A.3	7.1	5	0.90	Ilford03D.3	17.6	17	0.77
Ilford02B.1	6.5	5	0.54	Ilford05B.1	nd	9	0.69
Ilford02B.2	nd	5	0.54	Ilford05B.2	4.3	7	0.69
Ilford02B.3	8.3	6	0.54	Ilford05B.3	5.1	5	0.69
Ilford02C.1	7.8	6	0.46	Ilford06A.1	8.3	11	1.00
Ilford02C.2	4.7	7	0.46	Ilford06A.2	nd	7	1.00
Ilford02C.3	6.6	6	0.46	Ilford06A.3	nd	9	1.00
Ilford02D.1	19.5	11	0.52	Ilford07A.1	7.4	20	0.93
Ilford02D.2	*	8	0.52	Ilford07A.2	10.5	26	0.93
Ilford02D.3	9.4	9	0.52	Ilford07A.3	6.8	20	0.93
				Ilford07B.1	92.8	53	0.95
				Ilford07B.2	39.1	39	0.95
				Ilford07B.3	30.5	40	0.95

* *m/z* 1105 was not detected, however *m/z* 1106 was. nd indicates that the peptide was not detected

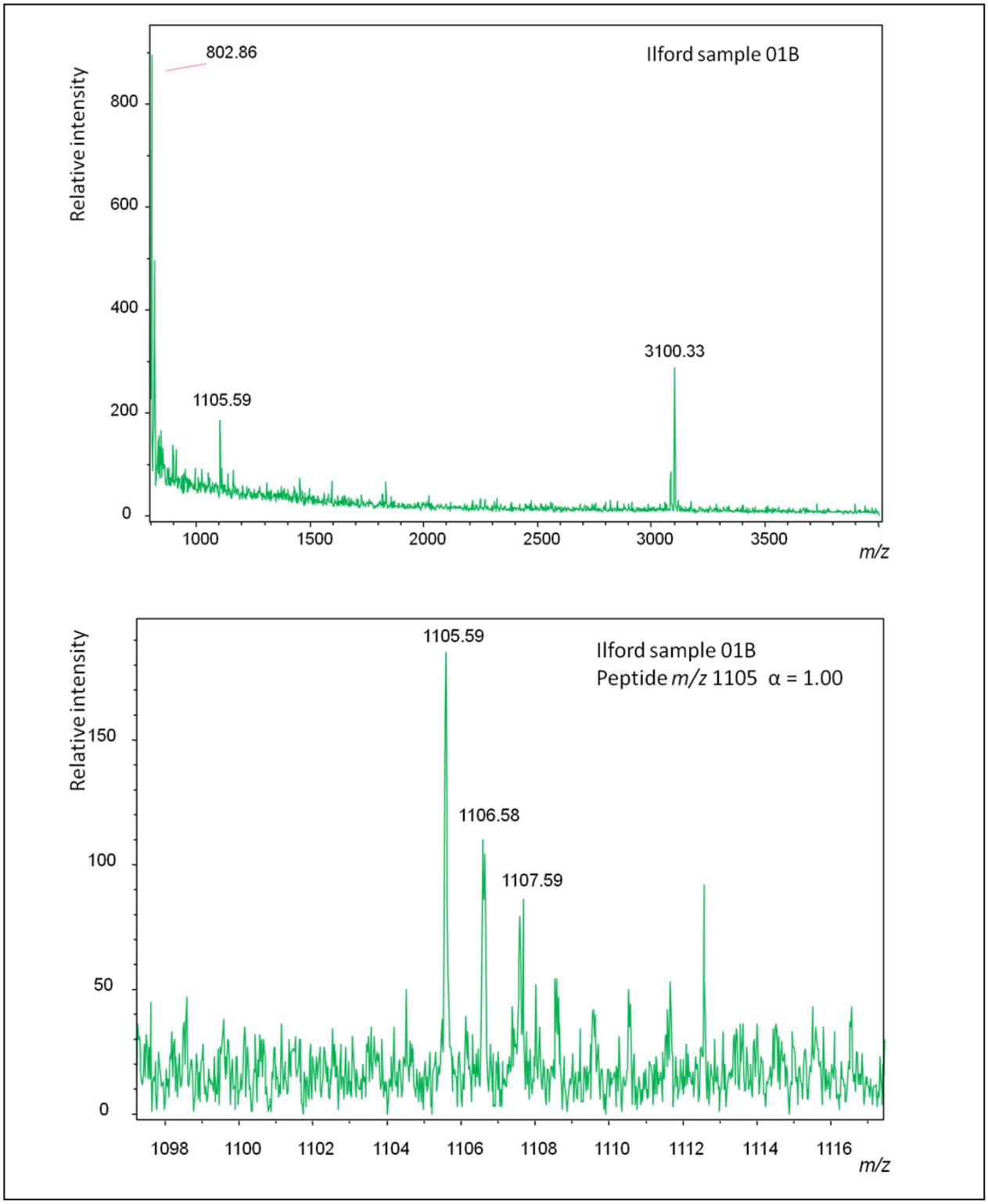


Figure 93: Spectrum obtained from a tryptic digest of protein extracted from a piece of Ilford (sample 01B) bone. Collagen was extracted using 50 mM ammonium bicarbonate

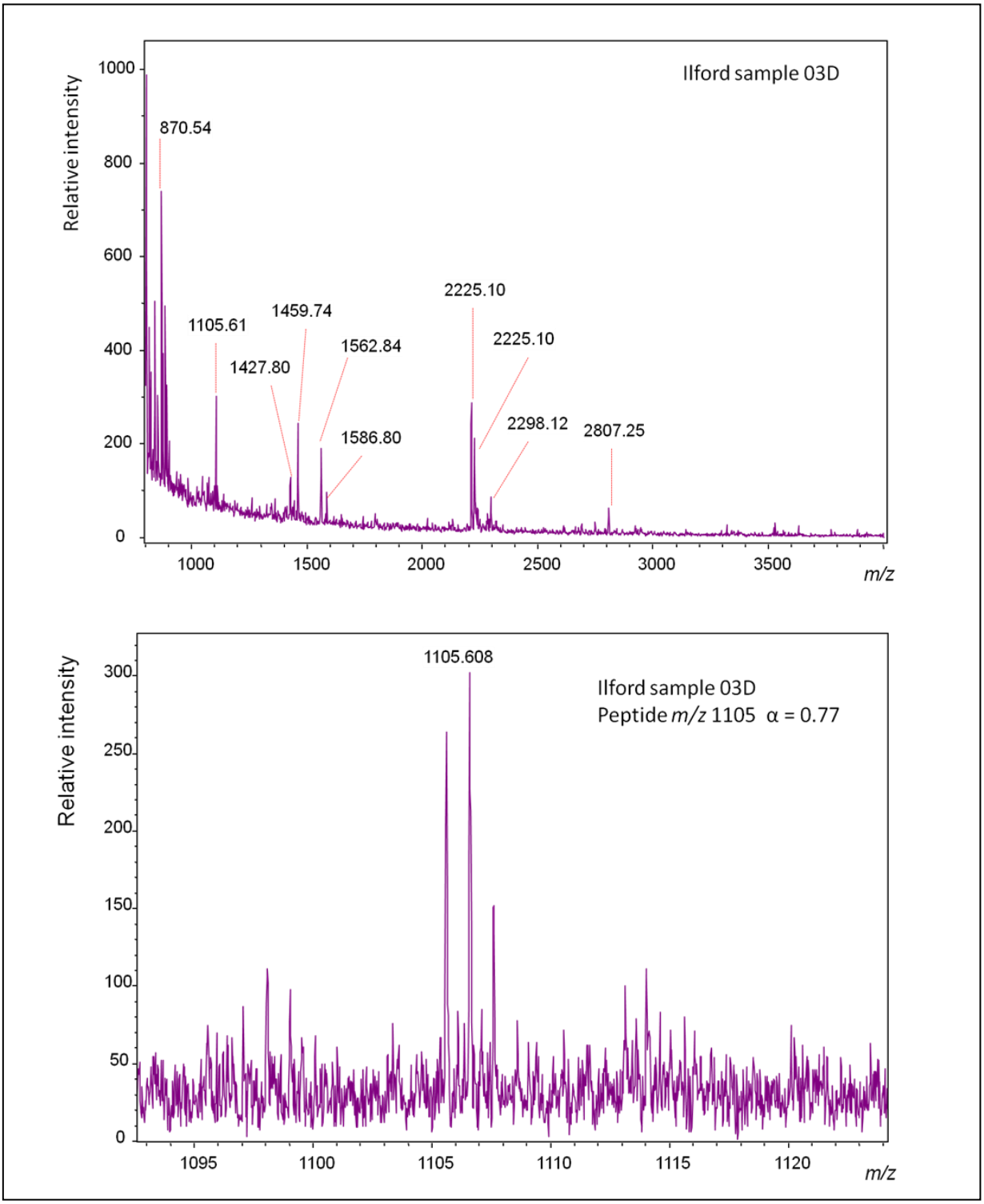


Figure 94: Spectrum obtained from a tryptic digest of protein extracted from a piece of Ilford (sample 03D) bone. Collagen was extracted using 50 mM ammonium bicarbonate

Figure 93 (Ilford01B.1) is representative of most of the samples from Ilford in terms of the quality of spectra obtained. Figure 94 is an example of a slightly better spectrum obtained from sample Ilford03D.3. It should be noted that even in the better of the two spectra (Figure 94), the biggest peaks in the spectrum are around *m/z* 800. This low *m/z* region of the spectrum is commonly associated with matrix peaks. If the peaks in this region are higher than those in the rest of the spectrum, this indicates a low amount of collagen has been measured.

Two of the Ilford samples (Ilford01A and Ilford07) produced good spectra in all three triplicates but gave very different α-values of 0.46 and 0.95 respectively. These two samples were also the only ones which had S/N ratios for peptide 1105 > 15 in all three analytical replicates. Sample Ilford01A had S/N values for *m/z* 1105 of 18.1, 19.6 and 15.0, with sample Ilford07 yielding much higher S/N ratios of 92.8, 39.1 and 30.5. One reason for the abnormally high α-value in sample Ilford07 may be due to contamination with younger bone material; this might also explain the higher S/N observed in this sample compared to other samples analysed for this site. Without further investigation of samples from this site, it is not possible to determine if the levels of deamidation observed in these samples are representative or of the levels of deamidation present in all bone from Ilford.

The α-values obtained in peptide *m/z* 1105 for Joint Mitnor and Victoria Cave (both correlated with MIS 5e) were compared to the S/N value of the *m/z* 1105 peptide. These sites were chosen for comparison as they are both dated to MIS 5e, but have different thermal ages, with Joint Mitnor being more thermally degraded. Despite the chalky appearance, and older thermal age of the bone from Joint Mitnor, the spectra obtained from this sample were better than those obtained from the visually well-preserved bone from Victoria Cave. Initially it was hypothesised that the quality of spectra obtained from the protein extract would decrease with the age/thermal of the sample, and that the level of deamidation would increase with the sample age/thermal age. The data from these sites showed the opposite relationship, with lower S/N corresponding with higher α-values (Figure 95). This, along with the data obtained from the Ilford samples implies that a false higher α-value may be obtained from poor spectra. From this data alone it is not possible to determine the effect of the S/N of a peptide and the calculated α-value, with further investigation required.

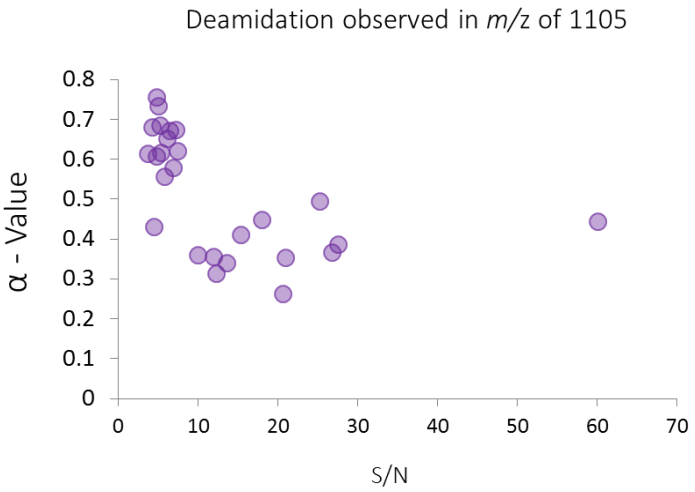


Figure 95: α- values obtained from peptide *m/z* 1105 observed in tryptic digests of protein extracted from bone from two MIS 5e sites, Joint Mitnor and Victoria Cave. The α-values are plotted against the corresponding S/N of the *m/z* 1105 peptide

7.5 Discussion

There are a number of possible causes for the variation observed in the data presented in this chapter. It may be that the level of glutamine deamidation is naturally variable, or that the preparation and/or analysis of the sample induces additional variation. It may be that the quality of the collagen present is too low, or the sample size was insufficient for this kind of semi-quantitative analysis. Since these studies were undertaken, it is clearer that MALDI may not be the best ionisation method for the analysis of deamidation (Chapter 4), and that ESI not only produces more accurate measurements, but also results in less variation between analytical sample replicates. In addition, the ammonium bicarbonate extraction method used in this chapter may not be suitable for the extraction of collagen from 'old' or heavily mineralised bone. The buffer soluble collagen is likely to be less structurally conserved and may be too degraded in thermally older material. It has been shown in heating experiments (Chapter 5) that the mineral bound collagen, and the buffer soluble collagen analysed here undergo deamidation at different rates at 65 °C. In order to obtain better quality spectra it may be necessary to perform EDTA de-mineralisation prior to the protein extraction. However this data may not be directly comparable to the levels of deamidation measured here, from the buffer soluble fraction. The data presented here also highlights the importance of checking raw data when using data processing algorithms, such as the *Q2E* program. At the moment the program decides if spectra are suitable for processing based on three main criteria, 1) the presence of the peptide which is being measured, 2) The S/N of the peak being measured (in this case a S/N cut off ≥ 6 has been used), and 3) the isotopic fit of the distribution fits the predicted distribution of the given peptide sequence. Additional parameters could be added to this, such as a minimum number of peptides with a S/N of ≥ 6 present in the spectrum or an increase in the S/N cut off, as well as ensuring all three triplicate spectra are of similar quality.

Another method of highlighting poor spectra could look at the ratio of the S/N peak commonly observed in the *m/z* region near to *m/z* 1105. In order to optimise this, further investigation into the effect of S/N on the measurement of deamidation is needed. This could be investigated further by analysing a dilution series of mixtures of the Gln and Glu-containing peptides described in Chapter 4. By looking at samples with known Q/E ratios, it may be possible to further investigate the effects of S/N on the observed level of deamidation, and help to establish limit of detection criteria.

7.6 Conclusion

The data obtained from nine sites has shown that there may be a relationship between glutamine deamidation and thermal age in peptide *m/z* 1105, but it is currently not possible to distinguish the ages of the sites purely by the levels of deamidation calculated using this method. Further investigation into the high α-values obtained from Ilford is required, with possible additional sampling and EDTA pre-treatment methods. The measurement of glutamine deamidation however may be a useful tool for identifying samples which are ^{14}C dead, thus saving money on analysis costs. Low levels of deamidation in collagen may also help well as helping to authenticate or support other dating methods looking at Pleistocene bone material.

Chapter 8 - Conclusions and future work

8.1 Conclusions

The work carried out in this thesis has addressed a number of questions on how glutamine degrades in bone over time, as well as how the sample preparation, analytical techniques and set up of kinetic experiments can affect the level of deamidation measured. These research outcomes can be summarised into four main findings.

- The effects of low pH (HCl) and neutral pH (EDTA/ammonium bicarbonate) extractions on levels of glutamine deamidation in bone (Chapter 6)
- Sample point variability and the effects of macroscopic degradation on glutamine deamidation (Chapter 6)
- A comparison of the accuracy of MALDI and ESI sources for the measurement of mixtures of Q and E-containing synthetic peptides (Chapter 4)
- The efficiency of heating experiments on modelling glutamine deamidation in bone (Chapter 5). Effects of low or neutral pH extraction methods on heated bone.

8.1.1 The effect of acidity during collagen pre-treatment

The treatment of bone with acidic solutions such as HCl prior to analysis of protein extracted from bone is common in archaeological science, for example when extracting collagen for isotopic analysis (including radiocarbon dating). The work presented here has shown that treatment of bone with HCl not only increases levels of glutamine deamidation, but also causes measurable swelling of the collagen fibril (Chapter 6). The observed swelling is more pronounced in older Pleistocene bone than in modern bone, highlighting the susceptibility of older/less well preserved bone to degradation induced during sample preparation. Analysis of tryptic collagen peptides from HCl-treated bone by mass spectrometry resulted in an absence of peaks corresponding to peptides containing aspartic acid. This may be of significance when looking for unique peptide markers, or when generating peptide mass fingerprints for species identification, such as distinguishing between cow and sheep, where the collagen peptide fingerprints differ by only a few peptides. If such peptides contain aspartic acid, then HCl treatment should be avoided. In addition, the heating of bone also resulted in the absence of the aspartic acid containing peptides and so should be minimised when trying to extract aspartic acid containing peptides.

8.1.2 Sample point variability

Levels of glutamine deamidation were found to be increased in bone sections which displayed macroscopic degradation, suggesting that glutamine deamidation may be a marker for bone preservation. In experiments described in Chapter 6, samples from five different slices taken from a single bone were used to investigate variability in deamidation levels. This investigation found that some peptides had relatively consistent levels of deamidation across the bone length, with deamidation in other peptides appearing to be more variable. Measuring deamidation in particular peptides at different positions in a bone could therefore help to create a map of the bone's preservation. The sample sizes required for MS analysis outlined in this thesis are very small (~ 10-30 mg of bone), which makes MS a good screening tool for precious samples, or samples of limited size. MS analysis could therefore be used to identify well-preserved bones from large sample sets, and/or areas within a single bone that may be more suitable for sampling for further analysis, such as DNA extraction or isotopic analysis. Levels of deamidation in bone artefacts could also help conservators identify samples which may be more susceptible to degradation during long term storage, as well as inform appropriate methods of conservation.

8.1.3 Using synthetic peptides to investigate suitability of MALDI and ESI sources for measuring deamidation

As the natural level of deamidation in biological material is generally unknown, it is difficult to quantify levels of deamidation induced during sample preparation and analysis without a reference standard. Standard peptide synthesis (Chapter 4) was used to investigate the suitability of two ionisation sources for the measurement of the ratios of Gln/Glu-containing peptides, and to test the accuracy of the genetic algorithm in deconvoluting the data obtained from a TOF instrument. The product peptides analysed in this thesis showed ionisation bias occurs when using MALDI, with product peptide E consistently underestimated. However, analysis of the same mixturesusing ESI showed no ionisation bias between the two peptide products and is therefore preferable for the measurement of glutamine containing peptides. The genetic algorithm performed well and may therefore be a good alternative where access to high-resolution mass spectrometers is not possible.

8.1.4 Heating experiments to induce deamidation

Unlike kinetic heating experiments carried out on single amino acids, the calculation of activation energies required for the deamidation of glutamine residues bound in proteins is problematic. In Chapter 5, bone chips were heated at a range of high temperatures (80 °C-140 °C) in order to assess the kinetic parameters for deamidation at different temperatures. However, these experimental samples produced poor quality spectra with few peptide peaks observed, typically with low S/N. This is most likely due to diagenetic changes other than deamidation occurring during

the heating process. In order to better observe collagen peptides after bone heating and tryptic digestion, bone chips were heated at a lower temperature of 65 °C over 75 weeks. Buffer-soluble collagen fractions were successfully extracted from these samples, and showed an increase in glutamine deamidation in the resulting tryptic peptides over time. However when the same bone chips were analysed using the EDTA pre-treatment method to extract the mineral-bound protein, little deamidation was observed. This suggests that the two collagen fractions (buffer-soluble and mineral bound) undergo deamidation at different rates, and that the buffer-soluble collagen fraction may not therefore be representative of the bone protein as a whole, which is important to consider when trying to use levels of deamidation in the buffer soluble fraction to estimate the age/preservation of a bone. Although low temperature kinetic studies proved to be gentle enough to preserve the buffer-soluble collagen fraction, temperatures this low require long experiment times in order to induce measurable deamidation. The adapted EDTA extraction method discussed in Chapter 5 may be beneficial when extracting collagen from bones heated at higher temperatures, where a buffer-soluble extraction is unlikely to be successful.

8.2 Future work

8.2.1 Choice of ionisation source

The data presented in this thesis shows that ESI is a better ionisation source for the measurement of ratios of Q and E-containing peptides in peptide mixtures. It should be noted however that these experiments were carried out in 'optimum' conditions, i.e. not taking into account sample matrix. Although the synthetic peptides were designed to be chemically representative of peptides commonly observed in collagen, they are unlikely to full represent the reactions possible in more complex biological samples. In order to further investigate the suitability of ESI over MALDI for the measurement of Q/E-containing peptides in complex biological mixtures, the synthetic peptides discussed in Chapter 4 could be mixed in known ratios and spiked into a number of different biological extracts,such as the tryptic peptides obtained from collagen extracted from bone. These mixtures could then be analysed using ESI and MALDI, enabling a better assessment of the effect of biological matrices on the ionisation of Q and E-containing peptides. Synthetic peptides could also be used to investigate the minimum levels of detection at which the genetic algorithm can still accurately determine deamidation levels. As well as being used as a standard to investigate the effects of ionisation, the peptides synthesised for this work could also be used to investigate the stability of peptides in different solutions over time at different conditions. Long term stability testing may help to identify if these peptides would be suitable internal reference standards, as well as enabling the investigation of the degradation pathways of glutamine in small peptides. Ideally a reference peptide mixture would contain a range of synthetic Q and E-containing peptides covering the mass range being analysed (in this case over a range of 900 – 4000 Da). In practice this would be difficult, as the longer the peptide being synthesised, the more steps required, resulting

in a lower purity of the final product. However having a reference standard containing a few peptides of known sequence and known levels of deamidation would greatly help to validate the methodologies being used and help to better investigate miscellaneous results.

8.2.2 Expanding the number of peptides measured

In the research carried out in the PhD, only 12 glutamine-containing peptides within collagen were investigated. In order to look at a wider range of peptides both, with and without glutamine, a *de novo* sequencing project could be carried out on commonly observed peaks in tryptic collagen peptides. Once these additional peptide sequences are confirmed, a larger scale assessment of peptides visible pre and post-heating and/or HCl treatment, as well as those present in young and old/degraded bone could be used to help further understand how the collagen molecule as a whole degrades in different conditions, as well as further in investigating how the collagen structure in bone is effected by age. This data could then be applied to annotate a 3D collagen chemical fibril map, to help identify regions within the protein which may be more or less susceptible to degradation. This would help to improve current methodologies that regularly utilise the analysis of collagen (such as peptide mass fingerprinting techniques or dietary investigations looking at isotope ratios), as well aiding biological studies on collagen deterioration or diseases which can occur *in vivo* (Takemoto and Boyle 1998; Masters, Bada and Zigler 1978).

8.2.3 Low temperature EDTA demineralisation

From the data obtained from heating experiments carried out in this PhD it was noted that pre-treatment of bone in EDTA at freezing temperature may be a suitable demineralisation method (Chapter 5). In order to investigate this further, a number of bone chips could be demineralised in EDTA at freezing temperatures for varying lengths of time to determine the optimum conditions. In addition this optimised demineralisation method could be carried out on a number of bones with differing levels of preservation. This would help to assess the suitability of this method on a wider sample range.

EDTA pre-treatment of bone at low temperature may be particularly useful when extracting collagen from poorly preserved or heavily mineralised bone, where the remaining protein is fragile and/or low in abundance. The effects of cold EDTA-demineralisation on the collagen fibril structure could also be investigated using the TEM methods described in chapter 6.

8.2.4 Heating experiments

The heating experiments carried out in this PhD were largely unsuccessful due to the fact that the heating of bone chips not only causes deamidation, but also causes breakdown of the soluble collagen fraction, resulting in a poor observation of collagen peptides extracted from bone post heating. Heating at low temperature (65 °C) resulted in successful extraction of collagen, but no observable levels of deamidation. It may be that heating experiments carried out at temperatures

higher than 65 °C may be more successful if the demineralised fraction is analysed using the low temperature EDTA extraction described in Chapter 5.

8.3 Can glutamine deamidation be used to date bone beyond the limits of radiocarbon dating?

In the nine sites chosen for this study (Chapter 7), deamidation in the buffer-soluble collagen fractions appears to correlate with thermal age, although variability within a sample is high. However, these measured levels of deamidation may not be representative of the bone as whole, as only the buffer-soluble fraction was analysed. The buffer-soluble collagen fraction and the mineral-bound collagen fraction are not mutually exclusive; as the bone degrades more collagen is likely to become buffer-soluble over time. This may be one of the reasons for the levels of variation observed in multiple measurements of the same bone fragments analysed in Chapter 7. An optimised version of the adapted EDTA method could be investigated further to assess the levels of deamidation in the mineral-bound fractions in bone from these sites. As protein in the mineral-bound fraction has been found to undergo deamidation more slowly than the buffer-soluble fraction (Chapter 5), the mineral-bound protein fraction may be useful when assessing the preservation of older material, such as bones from the Late Pleistocene period.

The main aims of the research described in this thesis were to investigate the suitability of glutamine deamidation as a marker of thermal age. The methods described in this thesis suggest that this may not be straightforward. However the relationship observed between deamidation and thermal age is promising and warrants further investigation. Re analysis of the samples from the nine sites discussed, using the adapted EDTA pre-treatment method as well as analysis using ESI instead of MALDI ionisation may result not only in more accurate measurements of deamidation, but also an increase in the number of observed Gln-containing peptides. It may be, however, that the variability observed is naturally occurring.

Although at present the extent of glutamine deamidation is too variable to be used as a precise direct dating tool for bone, it may be useful to screen bone material prior to radiocarbon dating, especially when analysing bone in deposits of mixed age. In order to look at deamidation as a suitable screening tool for samples that are radiocarbon dead, further investigation would be required, looking at a range of samples from sites of differing thermal ages. If samples could be screened prior to radiocarbon analysis then money could be saved on analytical costs, analysis time, as well as the amount of sample used.

Although some of the aspects investigated within this thesis (such as the effects of low pH on collagen structure, and ionisation bias when analysing peptide mixtures) are acknowledged problems, in both analytical and archaeological sciences, both HCl and MALDI are still routinely used to pre-treat and measure collagen peptides (Buckley *et al.*, 2009; Araki and Moini, 2011; Leo

et al., 2001; van Doorn *et al.*, 2011; Hurtado and O'Connor, 2012; Welker *et al.*, 2011). Therefore thorough investigation into exactly how these preparation and analysis methods may affect the measured outcome is important.

As the famous physicist Werner Heisenberg once said: 'What we observe is not nature itself, but nature exposed to our method of questioning' (Werner Heisenberg, 1962).

Abbreviations

Ka BP – thousand years before present

Cal BP – calibrated years before present

OSL –optically stimulated Luminescence

MS – Mass spectrometry

ESI – electrospray ionisation

MALDI – matrix assisted laser desorption/ionisation

TOF – MS - time of flight mass spectrometry

FT – ICR-MS –Fourier-transform ion cyclotron resonance

HPLC – High performance liquid chromatography

TEM – transmission electron microscopy

MW – molecular weight

m/z – mass to charge ratio

*t*R – retention time

References

Abraham, M.H., (2004). 100 Years of Chromatography—or Is It 171? *Journal of Chromatography A,* 1: 113-114.

Abrams, C., Bello, S.M., Di Modica, K., Pirson, S., and Bonjean, D., (2014). When Neanderthals used cave bear (Ursus spelaeus) remains: Bone retouchers from unit 5 of Scladina Cave (Belgium), *Quaternary International* 326-327: 274-287

Ambrose, S.H., and Norr, L., (1993). Experimental Evidence for the Relationship of the Carbon Isotope Ratios of Whole Diet and Dietary Protein to Those of Bone Collagen and Carbonate. *Prehistoric Human Bone*, 1–37.

Amit, K., and Ben-Tal, N., (2001). Introduction to Proteins, Structure, Function and Motion, CRC Press

Araki, N., and Moini, M., (2011). Age estimation of museum wool textiles from Ovis aries using deamidation rates utilizing matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 25 (22): 3396–400.

Ascough P, Cook G, Dugmore A., (2005). Methodological approaches to determining the marine radiocarbon reservoir effect. *Progress in Physical Geography* 29 (4): 532–47.

Bada, J.L., and Helfman, P.M., (1975). Amino acid racemization dating of fossil bones. *World Archaeology*, 7(2): 160–73.

Bada, J.L, and Protsch, R., (1973). Racemization reaction of aspartic acid and its use in dating fossil bones. *Proceedings of the National Academy of Sciences*, 70 (5): 1331-1134.

Bada, J.L., and Schroeder, R.A., 1975. (1975)Amino acid racemization reactions and their geochemical implications. *Die Naturwissenschaften*, 62 (2): 71-79.

Bada, J.L., Kvrnvolden, K., and Peterson., E, 1973). Racemization of amino acids in bones. *Nature*, 245: 308-310.

Bada, J.L., Gillespie, R., and Gowlett., (1984) Accelerator mass spectrometry radiocarbon ages of amino acid extracts from Californian palaeoindian skeletons. *Nature,* 312(5993): 442–4.

Bada, J.L., (1972). Kinetics of racemization of amino acids as a function of pH. *Journal of the American Chemical Society*, 94 (4): 1371-1373.

Bahr, U., Kras, F., and Hillenkamp, F., (1994). Analysis of biopolymers by Matrix-assisted Laser Desorption/Ionisation (MALDI) mass spectrometry. *Journal of Analytical Chemistry*, 348: 783-791.

Balter, M., (1014). The Killing Ground. *Science*, 344 (6188): 1080-1083.

Berisio, R., *and* Vitagliano, L*.,* (2002). Crystal structure of the collagen triple helix model 3. *Protein science,* 262-270.

Bahattacharyya, S.K., and Banerjee, A.B., (1974). D-amino acids in the cell pool of bacteria. Folia *Microbiologica*, 19: 43-50.

Bigi, A., Panzavolta, S., and Rubini, K., (2004). Relationship between triple-helix content and mechanical properties of gelatin films. *Biomaterials*, 25(25): 5675-5680.

Bocherens, H., Billiou, D., and Mariotti, A., (1999). Palaeoenvironmental and palaeodietary implications of isotopic biochemistry of last interglacial Neanderthal and mammal bones in Scladina Cave (Belgium). *Journal of Archaeological Science* 26: 599–607

Bridgland, D.R., (2003). Key Middle Pleistocene localities of the Lower Thames: site conservation issues, recent research and report of a Geologists' Association excursion, 8 July, 2000. *Proceedings of the Geologists' Association,* 114 (3): 211-225.

Bridgland, D.R., (1994). Quatrnary of the Thmes. In: Geological Conservation Review Series, vol, 7. Chapmand and Hall, London.

Brock, F., Ramsey, C. and Higham, T., (2007). Quality assurance of ultrafiltered bone dating. *Radiocarbon*, 49(2): 187–192.

Brock, F., Higham, T. and Ramsey, C.B., (2010). Pre-screening techniques for identification of samples suitable for radiocarbon dating of poorly preserved bones. *Journal of Archaeological Science*, 37(4): 855–865

Brooks, A., (1990). Dating pleistocene archeological sites by protein diagenesis in ostrich eggshell. *Science, 248 (248): 60-64.*

Brühl, E., and Mania, D., (2003). Neumark-Nord: a late middle Pleistocene lake shore with synchronous sites of different functional character. *Abstract Données récentes sur les modalités de peuplement en Europe au Paléolithique inférieur et moyen Rennes, 102*

Buckley, M., Collins, C., and Thomas-Oates, J., (2009). Species identification by analysis of bone collagen using matrix-assisted laser desorption / ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 23: 3843–3854.

Buckley, M*.,* Anderung, C., Penkman, K.E.H., Raney, B.J., Götherström, A., Thomas-Oates, J., and Collins, M.J., (2008). Comparing the survival of osteocalcin and mtDNA in archaeological bone from four European sites. *Journal of Archaeological Science* 35 1756-1764

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

Burjanadze, (2000). New analysis of the phylogenetic change of collagen thermostability. *Biopolymers,* **53(**6): 523-528.

Cameron, A.E., and Eggers Jr, D.F (1948). An Ion "Velocitron''. *Review of Scientific Instruments,* **19** (9): 605-607.

Candy, I., and Schreve, D., (2007). Land-sea correlation of Middle Pleistocene temperate sub-stages using high-precision uranium-series dating of tufa deposits from southern England. *Quaternary Science Reviews*, **26** (9-10): 1223-1235.

Capasso, S., Mazzarella, L., Sica, F., and Zagari, A., (1991). First Evidence of Spontaneous Deamidation of Glutamine Residue via Cyclic Imide to α- and γ-Glutamic Residue Under Physiological Conditions. *Journal of the Chemical Society, Chemical Communications*., **23**: 1667–1668.

Cleland, J.L., Powell, M.F., and Shire, S.J., and Turner-Walker, G., (1993). The development of stable protein formulations: a close look at protein aggregation, deamidation and oxidation. *Critical Reviews in Therapeutic Drug Carrier Systems,* **10**: 307-377.

Cole, R.B., (2010). Electrospray and MALDI mass spectrometry: fundamentals, instrumentation, practicalities, and biological applications. John Wiley and Sons,

Collins, M.J., Riley, M.S, Child, A.M., (1995). A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. *Journal of Archaeological Science*, **22**(2): 175–183.

Collins, M.J., Waite, E.R., and van Duin, A.C., (1999). Predicting protein decomposition: the case of aspartic-acid racemization kinetics. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* **354** :51-64.

Collins, M.J, Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R.V., Wess, T.J., Csapo, J., Millard, A.R.m and Turner-Walker, G., (2002). The survival of organic matter in bone: a review. *Archaeometry*, **44** (3): 383-394

Comisarow, M.B., and Marshall, A.G. (1974). Fourier transform ion cyclotron resonance spectroscopy. *Chemical physics letters*, **25** (2): 282-283.

Covington, A.D., Song, L., Suparano, O., Koon, H.E.C., and Collins, M.J., (2010). Link-lock: the mechanism of stabilising collagen by chemical reactions. *World Leather*, **25**(5): 35–43.

Crisp, M., Demarchi, B., Collins, M., Morgan-Williams, M., Pilgrim, E., (2013). Isolation of the intra-crystalline proteins and kinetic studies in Struthio camelus (ostrich) eggshell for amino acid geochronology. *Quaternary Geochronology*, **16**: 110-128.

Currant, A., and Jacobbi, R.M. (2001) A formal mammalian biostratigraphy for the late Pleistocene of Britain. *Quaternary Science Review*, **20**: 1701-1716.

Demarchi, B., Collins, M.J., Tomiak, P.J., Davies, B.J., and Penkman, K.E.H., (2013) Intra-crystalline protein diagenesis (IcPD) in Patella vulgata. Part II: Breakdown and temperature sensitivity. *Quaternary Geochronology*, **16**: 158-172

Dobberstein, R.C., Collins, M.J., Craig, O.E., Taylor, G., Penkman, K.E.H., and Ritz-Timme, S., (2009). Archaeological collagen: Why worry about collagen diagenesis? *Archaeological and Anthropological Sciences*, **1**(1): 31–42.

Dole, M., *Mack, L., and Hines, R.,* (1968). Molecular beams of macroions. *Journal of Chemical physics*, **49**: 2240-2249.

Duller, G.A.T., (2008). Luminescence Dating: guidelines on using luminescence dating in archaeology. Swindon: English Heritage

Edwards, *R.L., Chen, J.H., and Wasserburg, G.J.,* (1987). Precise timing of the Last Interglacial period from mass spectrometric analysis of ^{230}Th in corals. *Science,* **236**, 1547– 1553.

Emiliani, C., (1955). Pleistocene temperatures. *Journal of Geology* **63**: 585-99.

Ettre, L., and Sakodynskii, K.I., (1993). M . S . Tswett and the Discovery of Chromatography h Early Work (1899-1903). *Chromatographia, **35***(3/4): 223-231.

Fenn, J.B., (2002). Electrospray ionisation mass spectrometry: How it all began. *Journal of biomolecular techniques*, **13** (3): 101-118.

Fernandes, R., and Bergemann, S. (2012). Mussels with Meat: Bivalve Tissue-Shell Radiocarbon Age Differences and Archaeological Implications. *Radiocarbon*, **54** (3): 953-965.

Flaugh, S., Mills,. I., and King, J., (2006). Glutamine deamidation destabilizes human gamma D-crystallin and lowers the kinetic barrier to unfolding. *The Journal of biological chemistry*, 281 (41): 30782-93.

Gascoyne, M., Currant, A., and Lord, T., (1981). Ipswichian fauna of Victoria Cave and the marine palaeoclimatic record. *Nature*, **294**, 652–654.

Gaudzinski-Windheuser, S., (2013). The Eemian Interglacial lake-landscape at Neumark-Nord (Germany) and its potential for our knowledge of hominin subsistence strategies. *Quaternary International*, 1-8.

Geiger, T. and Clarke, S., (1987). Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *Journal of Biological Chemistry*, 262(2): 785.

Goodfriend, G.A., (1991). Patterns of racemization and epimerization of amino acids in land snail shells over the course of the Holocene. *Geochimica et Cosmochimica Acta,* Acta 55 (1):293-302.

Gordon, C.C., and Buikstra, J.E. (1981) Soil pH, Bone Preservation, and Sampling Bias at Mortuary Sites. *American Antiquity*, 46 (3), 566–571.

Gorres, K., and Raines, R., (2010). Prolyl 4-Hydroxylase. *Critical reviews in biochemistry and molecular biology*, 45 (2): 106-124.

Götherström, A., Collins, M.J., Angerbjörn, A., and Lidén, K., (2002). Bone preservation and dna amplification. *Archaeometry*,3: 395-404.

Grün, R., Aubert, M., Hellstorm, J., Duval, M., (2010). The challenge of direct dating old human fossils. *Quaternary International* 223-224: 87-93.

Grün, R., (2006). Direct Dating of Human Fossils. *Yearbook of Physical Anthropology*, 48: 2-48.

Hajdas, I., Ivy-ochs, S., Picker, R., and Preusser, F., (2007). Recent developments in Quaternary dating methods. *Geographica Helvetica*, 63: 1-4.

Hawkey, D.E. and Merbs, C.F., (1995). Activity-induced Muscu los keletal Stress Markers (MSM) and Subsistence Strategy Changes among Ancient Hudson Bay Eskimos. *International Journal of Osteoarchaeology*,.324 – 338.

High, K., Milner, N., Panter, I., and Penkman, K.E.H., (2015). Appetite for destruction: investigating bone degradation due to high acidity at Star Carr. *Journal of Archaeological Science*, *59*: 159-168.

Hill, R.L., (1965). Hydrolysis of proteins. *Advances in Protein Chemistry* 20: 37-107.

Holmgren, S.K., Taylor, K.M., Bretscher, L.E., and Raines, R.T., (1999). Code for collagen's stability deciphered. *Nature*, 392: 666-667.

Huntley, D.J., Godfrey-Smith, D.I., and Thewalt, M.L.W., (1985). Optical dating of sediments. *Nature* 313: 105-107.

Hurtado, P.P. and O'Connor, P.B., (2012). Deamidation of collagen. *Analytical chemistry*, **84**(6): 3017–25.

Huxtable, J., and Aitken, M.J., (1992). Thermoluminescence dating of burned flint and stalagmitic calcite from grotte de Sclayn (Namur). *Etudes et Recherches Archeologiques de l'Universite´de Liege* **27**: 175–177.

Iribarne, J.V., and Thomson, B.A., (1976). On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64 (6): 2287-2294.

Jacobs, Z., Duller, G.A.T., Wintle, A.G., and Henshilwood, *C.S.,* (2006). Extending the chronology of deposits at Blombos Cave, South Africa, back to 140 ka using opticaldating of single and multiple grains of quartz. *Journal of Human Evolution* **51**: 255-273.

Jagtap, R.N., and Ambre, A.H., (2005). Overview literature on matrix assisted laser desorption ionisation mass spectroscopy (MALDI MS): basics and its applications in characterizing polymeric materials. *Bulletin of Materials Scienc, **28** (6): 515-528.*

Jonsson, R., Tarkowski, A. and Klareskog, L., (1986). A demineralization procedure for immunohistopathological use EDTA treatment preserves lymphoid cell surface antigens. *Journal of Immunological Methods*, **88**: 109–114.

Johnson, B., and Miller, G., (2007). Archeological applications of amino acid racemization. *Archaeometry, **39** (2): 265-287.*

Joshi, A.B, (2005). Studies on the mechanism of aspartic acid cleavage and glutamine deamidation in the acidic degradation of glucagon. *Journal of pharmaceutical sciences,* **94** (9): 1912-1927.

Joshi, A., and Kirsch, L., (2004). The estimation of glutaminyl deamidation and aspartyl cleavage rates in glucagon. *International journal of pharmaceutics,* **273** (1-2): 213 -219.

Karas, M., Gluckmann, M., and Schafer, J., (2000). Ionisation in matrix-assisted laser desorption / ionisation: singly charged molecular ions are the lucky survivors. *Journal of Mass Spectrometry,* **35**: 1-12.

Karas, M., Bachmann, D., and Hillenkamp, F., (1985). Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry*, 57

Karas, M., *Backmann, D., Bahr, U., and Hillenkamp, F.,* (1987). Matrix- assisted ultraviolet laser desorption of non-volatile com- pounds. *International Journal of Mass Spectrometry Ion Processes*, **78**, 53.

Kasem, M.A, Russo, R.E., Harith, M.A., (2011). Influence of biological degradation and environmental effects on the interpretation of archeological bone samples with laser-induced breakdown spectroscopy. *Journal of Analytical Atomic Spectrometry,*26 (9): 1733-1739.

Kebarle, P., and Verkerk, U.H., (2009). Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass spectrometry reviews*, 898-917.

Kirby, D.P., Buckley, M., Promise, E., Trauger, S.A., and Holdcraft, T.R., (2013). Identification of collagen-based materials in cultural heritage. *The Analyst ,*138 (17): 4849-4858.

Kraut, A., *Marcellin, M., Adrait, A., Kuhn, L., Louwagie, M., Kieffer-Jaquinod, S., Lebert, D., Masselon, C.D., Dupris, A., Bruley, C., Jaquinod, M., Jerome, G.,and Gallagher-Gambarelli, M.,* (2009). Peptide storage: are you getting the best return on your investment? Defining optimal storage conditions for proteomics samples. *Journal of Proteome Research,* 8(7): 3778-3785.

Koon, H., O'Connor, T., and Collins, M.J., (2012). Identification of Cooked Bone Using TEM Imaging of Bone Collagen. *Forensic Microscopy for Skeletal Tissues*, 915: 249–261.

Koon, H.E.C., O'Connor, T.P. and Collins, M.J., (2010). Sorting the butchered from the boiled. Journal of Archaeological Science, 37(1): 62–69.

Kuitems, M., (2015). Carbon and nitrogen stable isotopes of well-preserved, Middle Pleistocene bone collagen from Schöningen (Germany) and their palaeoecological implications. *Journal of Human Evolution, 1-9.*

Kyte, J., and Doolittle, R., (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* 157 (1): 105-132.

Lang, J., *Winsemann, J., Steinnmetz, D., Polom, U., Pollok, L., Böhner, U., Seangeli, J., Brandes, C., Hampel, A., and Winghart, S.,* The Pleistocene of Schöningen Germany: a complex tunnel valley fill revealed from 3D subsurface modelling and shear wave seismics. *Quaternary Science Review*, 39: 1-20.

Leo, G., Bonaduce, I., Andreotti., Marino, G., Pucci, P., Colomnini, M.P., Birolo, L., (2011). Deamidation at Asparagine and Glutamine As a Major Modification upon Deterioration/Aging of Proteinaceous Binders in Mural Paintings. *Analytical Chemistry*, 2056–2064.

Li, X., Lin, C., and O'Connor, P.B., (2010). Glutamine deamidation: differentiation of glutamic acid and γ-glutamic acid in peptides by electron capture dissociation. *Analytical Chemistry*, 82(9): 3606–15.

Libby, W., (1960). Radiocarbon dating. Nobel lecture.

Lisiecki, L.E., and Raymo, M.E., (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic delta ^{18}O records. *Paleoceanography*, **20** (1): 1-17.

Lister, A.M., (1992). Mammalian fossils and Quaternary biostratigraphy. *Quaternary Science reviews*, *11*: 329-344.

Lowe, J., and *Walker, M., (1997). Reconstructing Quaternary Environments, Vol. 2. London, Longman*

Mamyrin, B., Karataev, V.I., Shmikk, D.V., and Zagulin, V.A., (1973). The mass-reflectron, a new non-magnetic time-of-flight mass spectrometer with high resolution. *Journal of Experimental and Theoretical Physics,* **64** (1): 82-89.

Mann, R.W. and Hunt, D.R., (2013). Photographic regional atlas of bone disease: A guide to pathologic and normal variation in the human skeleton, Charles C Thomas Publisher.

Marieb, M., and Hoehn, K., (2007). Human Anatomy and Physiology, Pearson Education.

Marom, A., Mccullagh, J.S.O., Higham, T.F.G., Sinitsyn, A.A., and Hedges, R.E.M., (2012). Single amino acid radiocarbon dating of Upper Paleolithic modern humans. *Pnas*, **109**(18): 1–4.

Masters, P., Bada, J., and Zigler, J., (1978). Aspartic acid racemization in heavy molecular weight crystallins and water insoluble protein from normal human lenses and cataracts. *Proceedings of the National Academy of Sciences of the United States of America*, 75 (3): 1204-1208.

McCullagh, J.S.O., Marom, A. and Hedges, R.E.M., (2010). Radiocarbon dating of individual amino acids from archaeological bone collagen. *Radiocarbon*, **52**(2): 620–634.

Mckerrow, J., (1979). None-Enzymatic, Post-Translational, Amino Acid, Modification in Ageing. A Brief Review. *Mechanisms of Aging and Delvelopment,* **10**: 371-377.

Mikšík, I., (2014). Prince Cangrande's Collagen: Study of Protein Modification on the Mummy of the Lord of Verona, Italy (1291–1329 AD). *Chromatographia*

Moreau, K., and King, J., (2012). Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends in Molecular Medicin,* **18** (5): 273-282.

Nicholson, A., (1998). Bone Degradation in a Compost Heap. *Journal of Archaeological Science*, 25: 393-403.

O'Connor, T., (2000). The archaeology of animal bones, Sutton Publishing Limited.

O'Connor, T., and Lord, T., (2013). Cave palaeontology. *Caves and Karst of the Yorkshire Dales.* Chapter 15.

Olsen, J., Ong, S., and Mann, M., (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular and cellular proteomics*, **3** (6): 608-614.

Orgel, J., Miller, A. and Irving, T., (2001). The in situ supermolecular structure of type I collagen. *Structure*, **9**(01): 1061–1069.

Ottani, V., Raspanti, A., and Ruggeri, A., (2001). Collagen structure and functional implications. *Micro,* **32** (3): 251-260.

Penkman, K.E.H., Kaufman, D.S., Maddy, D., and Collins, M.J., (2008). Closed-system behaviour of the intra-crystalline fraction of amino acids in mollusc shells. *Quaternary,* **3** (1-3): 2-25.

Penkman, K.E.H., Preece, R.C., Bridlandm D.R., Keen, D.H., Meijer, T., Parfitt, S.A., White, T.S., and Coliins, M.J., (2013). An aminostratigraphy for the British Quaternary based on Bithynia opercula. *Quaternary Science Reviews,* **61**: 111-134.

Pike, A., Hedges, R., and Van Calsteren, P., (2002). U-series dating of bone using the diffusion-adsorption model. *Geochimica et Cosmochimica Acta*, **66** (24): 4273-4286.

Piliang, H.,Ren, Y., Alpert, and A.J., Sze, S.K., (2001). Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *The American Society for Biochemistry and Molecular Biology*.

Pisskiewicz, D., Landon, and M., Smith., (1970). Anomalous Cleaveage of Aspartyl-Proline Peptide Bonds During Amino Acid Sequence Determinations. *Biochemical and Biophysical research Communications,* **40** (5): 1173-1178.

Radkiewicz, J.L. Zipse, H., Clarke, S., Houk, K.N., (1996). Accelerated Racemization of Aspartic Acid and Asparagine Residues via Succinimide Intermediates: An ab Initio Theoretical Exploration of Mechanism. *Journal of the American Chemical Society*, **118**(38): 9148–9155.

Radkiewicz, J.L. Zipse, H., Clarke, S., Houk, K.N., (2001). Neighboring Side Chain Effects on Asparaginyl and Aspartyl Degradation : An Ab Initio Study of the Relationship between Peptide Conformation and Backbone NH Acidity. *Journal of the American Chemical Society*, **124**: 3499–3506.

Railsback, L.B., Gibbard, P.L., Head, M.J., Riavo, N.Y., Voarintsoa, G., and Toucanne, S., (2015). An optimized scheme of lettered marine isotope substages for the last 1.0 million years , and the climatostratigraphic nature of isotope stages and substages. *Quaternary Science Reviews*, **111**: 94-106.

Ramsey, C. and Higham, T., (2007). Improvements to the pretreatment of bone at Oxford. *Radiocarbon*, **46**(1): 155–163.

Reimer, P., Bard, E. and Bayliss, A., (2013). IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon*, 55(4): 1869–1877.

Renfrew, C., and Clarke, R.M., (1974). Problems of the radiocarbon calendar and its calibration. *Archaeometry* 16 (1): 5-18.

Rho, J., Kuhn-Spearing, L., Zioipos, P., (1998). Mechanical properties and the hierarchical structure of bone. *Medical engineering and physics*, 20 (2): 92-102.

Rich, A. and Crick, F.H.C., (1961). The molecular structure of collagen. *Journal of Molecular Biology*, 3(5): 483–506.

Ritchie, R., (2011). The conflicts between strength and toughness. *Nature Materials*, 10 (11): 817-822.

Robinson, (2004). Molecular clocks: Deamidation of asparaginyl and gluatminyl residues in pepides and proteins. In Peptides. Althouse press, cave junction,1–11.

Robinson, A.B., McKerrow, J.H., and Cary, P., (1970). 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,* 66 (3): 753–7.

Schreve, D., (2001). Differentiation of the British late Middle Pleistocene interglacials: The evidence from mammalian biostratigraphy. *Quaternary Science Reviews*, 20 (16-17): 1693-1705.

Schreve, D., (1998). Mammalian biostratigraphy of the later Middle Pleistocene in Britain. Vol. 2.

Schwarcz, H., (2005). Uranium series dating in paleoanthropology. Evolutionary Anthropology: Issues, News, and Reviews, 1 (2): 56-62.

Scotchler, J., and Robinson., (1974). Deamidation of glutaminyl residues: dependence on pH, temperature, and ionic strength. *Analytical Biochemistry,* 59 (1): 319-322.

Sealy, J., Johnson, M., Richards, M., and Nehlich, O., *(*2014). Comparison of two methods of extracting bone collagen for stable carbon and nitrogen isotope analysis: comparing whole bone demineralization with gelatinization and ultrafiltration. *Journal of Archaeological Science*, 47: 64-69.

Shoulders, M. D. and Raines, R.T., (2009). Collagen structure and stability. *Annual Review of Biochemistry*, 78: 929–58.

Sinex, F., (1957). Aging and the lability of irreplaceable molecules. *Journal of gerontology*, 12 (2): 190-198.

Smith, C.I., Craig, O.E., Prigodich, R.V., Nielsen-Marsh, C.M., Jans, M.M.E., Vermeer, C., and Collins, M.J., (2005). Diagenesis and survival of osteocalcin in archaeological bone. *Journal of Archaeological Science*, **32**(1): 105–113.

Smith, C.I., Chamberlain, A.T., Riley, M.S., Stringer, C., and Collins, M.J., (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution,* **45**(3): 203–217.

Snedecor, G.W., (1934). Calculation and interpretation of analysis of variance and covariance,

Stapels, M.D., and Barofsky, D., (2004). Complementary use of MALDI and ESI for the HPLC-MS/MS analysis of DNA-binding proteins. *Analytical chemistry,* **76** (18): 5423-5430.

Stejskal, K., Potěšil, D., and Zdráhal, Z., (2013). Suppression of peptide sample losses in autosampler vials. *Journal of Proteome Research*, **12** (6): 3057-3062.

Stephens, W., (1946). A pulsed mass spectrometer with time dispersion. A*rchives of the Bulletin of the American Physical Society*, **21** (2):22

Stevens, S., (2008). Factors that contribute to the misidentification of tyrosine nitration by shotgun proteomics. *Molecular and cellular proteomics*, **7** (12): 2442-2451.

Stratton, L.P., Kelly, R.M., Rowe, J., Shively, J.E., Smith, D.D., Carpenter, J.F., and Manning, M.C., (2001). Controlling deamidation rates in a model peptide: effects of temperature, peptide concentration, and additives. *Journal of Pharmaceutical Sciences*, **90**(12): 2141

Stroop, S.D., (2007). A modified peptide mapping strategy for quantifying site-specific deamidation by electrospray time-of-flight mass spectrometry. *Rapid communications in mass spectrometry*, **21** (6): 830-836.

Stuart, A.J., (1976). The history of the mammal fauna during the IpswichianlLast interglacial in England. *Philosophical Transactions of the Royal Society of London*, **8276**: 221-250.

Sutcliffe, A.J., (1960). Joint Mitnor Cave, Buckfastleigh. *Transactions and Proceedings of the Torquay Natural History and Scientific Society,* **13**: 1–26.

Sykes, G.A., (1995). The significance of a geochemically isolated (intra-crystalline) organic fraction within biominerals. *Organic Geochemistry*, **23**: 1059-1066.

Takehara, T., and Takahashi, H., (2003). Suppression of Bcl-xL deamidation in human hepatocellular carcinomas, *Cancer Research*, **63** (12): 3054-3057.

Takemoto, L. and Boyle, D., (1998). Deamidation of specific glutamine residues from alpha-A crystallin during aging of the human lens. *Biochemistry*, **37**(39): 13681–5.

Taylor, R., *Payen, L.A., Prior, C.A., Slota, P.J., Gillespie, R., Gowlett, J.A.J., Hedges, R.E.M., Jull, A.J.T., Zabel, T.H., Donahue, D.J., and Berger, T.,* (1985). Society for American Archaeology Major Revisions in the Pleistocene Age Assignments for North American Human Skeletons by C-14 Accelerator Mass Spectrometry: None Older Than 11,000 C-14 Years B.P. *American Antiquity*, 50 (1): 136-140.

Thieme, H., (1997). Lower Palaeolithic hunting spears from Germany. *Nature*, 385 (6619): 807-810.

Towe, K.M., (1980). Preserved organic ultrastructure: An unreliable indicator for paleozoic amino acid biogeochemistry. *The Biogeochemistry of Amino Acids*, ed. PE Hare, TC Hoering, and K King JR, 65-74. Wiley, New York.

Trueman, C.N., and Martill., D.M., The long-term survival of bone: the role of bioerosion. *Archaeometry*, 44 (3): 371-382.

Turner-Walker, G., (2008). The Chemical and Microbial Degradation of Bones and Teeth. *Advances in Human Palaeopathology*.

Tuross, N., (2012). Comparative decalcification methods, radiocarbon dates, and stable isotopes of the VIRI bones. *Radiocarbon*, 54(3): 837–844.

Van der Plicht, J. and Palstra, S.W.L., (2014). Radiocarbon and mammoth bones: What's in a date. *Quaternary International*, 1–6.

van Doorn, N.L, Hollund, H., and Collins, M.J., (2011). A novel and non-destructive approach for ZooMS analysis: Ammonium bicarbonate buffer extraction. *Archaeological and Anthropological Sciences*, 3(3): 281–289.

Van Doorn, N.L., Wilson, J., Hollund, H., Soressi, M., and Collins, M., (2012). Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Communications in Mass Spectrometry,* 26(19): 2319–27.

Van Duin, A.C.T. and Collins, M.J., (1998). The effects of conformational constraints on aspartic acid racemization. *Organic Geochemistry*, 29(5): 1227–1232.

van Kleef, F.S.M., De Jong, W.W., and Hoenders, H.J., (1975). Stepwise degradations and deamidation of the eye lens protein a-crystallin in ageing. *Nature*, 254: 620-622.

Viguet-Carrin, S., Garnero, P. and Delmas, P.D., (2006). The role of collagen in bone strength. *Osteoporosis International*, 17(3): 319–36.

Welker, F., Collins, M.J., Thomas, J.A., Wadsley, M., Brace, S., Cappellini, E., Turvay, S.T., Reguero, M., Gelfo, J.N., Kramarz, J., Burger, J., Thomas-Oates, J., Ashford, D.A., Ashton, P.D., Rowsell, K.,

Porter, D.M., Kesser, B., Fischer, R., Baessmann, C., Kasper, S., Olsen, J.V., Kiley, P., Elliott, J.A., Kelstrup, C.D., Mullin, V., Hofreiter, M., Willerslev, E., Hublin, J., Orando, L., Barnes, I., and Macphee, R.D.E., (2015). Ancient proteins resolve the evolutionary history of Darwin 's South American ungulates. *Nature,* 522: 81-87.

Heisenberg, W., (1962). Physics and philosophy: The revolution in modern science.

Westgate, J.A., Preece, S.J., Froese, D.G., Pearce, N.J.G., Roberts, R.G., Demuro, M., Hart, W.K., and Perkins, W., (2008) Changing ideas on the identity and stratigraphic significance of the Sheep Creek tephra beds in Alaska and the Yukon Territory, north-western North Americ*a. Quaternary International,* 178(1): 183–209

Whitford, D., (2008). proteins, structure and function. In John Wiley and Sons Ltd, 92–101.

Wiley, W., and McLaren, I., (1995). Time-of-flight mass spectrometer with improved resolution. *Review of Scientific Instruments*, 26 (12): 1150-1157.

Wilson, J., van Doorn, N.L. and Collins, M.J., (2012). Assessing the extent of bone degradation using glutamine deamidation in collagen. *Analytical Chemistry*, 84 (21): 9041–8.

Wood, R.E., Bronk Ramsey, C., and Higham, T.F.G., (2010). Refining Background Corrections For Radiocarbon Dating of Bone Collagen at Orau. *Radiocarbon*, 52 (2): 600-611.

Wood, R., (2015). From revolution to convention: The past, present and future of radiocarbon dating. Journal of Archaeological Science, 56: 61-72.

Xie, M. and Schowen, R.L., (1999). Secondary structure and protein deamidation. *Journal of Pharmaceutical Sciences*, 88 (1): 8–13.

Yamashita, M., and Fenn, J.B., (1984). Negative ion production with the electrospray ion source. The *Journal of Physical Chemistry,* 750 (9): 4611-4615.

Zeng, J., (2001). Non-linear electrohydrodynamics in microfluidic devices. *International journal of molecular sciences,* 12 (3): 1633-1649.

Zenobi, R., and Knochenmuss, R., (1999). Ion Formation in MALDI Mass Spectrometry. 337-366