Exploring patterns of diagenesis in ancient proteins

Abigail Daisy Ramsøe

Doctor of Philosophy

University of York

Archaeology

September 2020

Abstract

Palaeoproteomics is a rapidly growing field of research that involves applying mass spectrometry to objects of archaeological and palaeontological interest in order to gain insight into the past. However, due to a lack of clear authentication criteria and understanding of the biomolecular characteristics of ancient proteins, there are suspicions about many remarkable results. This is of particular importance, as many samples are derived from museum collections, which may have been exposed to contamination, or had contaminants added as part of past conservation treatments. This thesis aims to (a) investigate the diagenetic forces acting on ancient proteins by quantifying the patterns of degradation in truly ancient proteins, and (b) use this knowledge in order to develop tools to authenticate results.

This thesis develops a novel tool, deamiDATE 1.0, for the authentication of ancient proteins using site-specific deamidation. This is relatively effective in broadly differentiating between modern and ancient proteins (with a focus on structural (collagen) and dietary (milk) proteins). Analysis of site-specific deamidation in ancient milk proteins reveals that although there is a modest trend of increasing deamidation over time, the deamidation of milk proteomes in the archaeological record is hugely variable. Therefore, other lines of evidence are recommended in order to assess the authenticity of low abundance proteins. Lastly, the patterns of collagen degradation in parchment are elucidated, showing that there is a region of the collagen structure that is a hotspot for hydrolysis. Taken together, the results of this thesis represent a novel contribution to the understanding of the diagenetic forces that act on proteins through time, and how to reliably authenticate proteomics data.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This work was supported by the Arts & Humanities Research Council (grant number AH/N005015/1).

I am extremely lucky to have been under the supervision of five wonderful supervisors throughout the course of this PhD. My first thanks must go to my original supervisory team - Matthew Collins, Camilla Speller, and Ian Barnes. Matthew, thank you for all your ideas, and for being extraordinarily enthusiastic about all my results (and non-results!), and for always making time to chat. Camilla, I am so grateful for your support throughout my PhD, thank you for being one of the most organised people I know - our meetings always left me feeling calm, motivated, and convinced I would finish this PhD. Ian, thank you for making this project possible, helping shape its focus, and for introducing me to the secrets of the Natural History Museum. Next, thank you to my fourth supervisor, Nathan Wales. Nathan, I can't thank you enough for taking me on as your student, even when our research areas did not have much intersection, and for continuing to make me feel welcome and appreciated in your group. Lastly, my sincere thanks to my fifth supervisor, Jessica Hendy. Jessie, even though you joined at a rather late stage of my PhD, your help and guidance during writing-up has been invaluable, and I am so grateful for your input and support. Thank you all being my supervisors - I could not have done it without any one of you.

One of the best times of my PhD was my secondment to the British Antarctic Survey. I would like to express my sincerest gratitude to my supervisors there, Melody Clark and Victoria Sleight. Thank you both for your kindness and patience, and the opportunity to collaborate on such exciting research.

I am also grateful for the support of my colleagues. In particular, I could not have finished this thesis without my office mates: Meaghan, Liam, and Bharath. Meaghan is probably very glad this is over, because there was not a day that went by where I wasn't asking her inane questions about when the Iron Age was, or how the mass spectrometer works. Thank you for always answering my questions. Thank you to Liam for always providing some entertaining distraction during the work day, and for buying me beer on every single train journey we took together. Lastly, thank you to Bharath, who, weeks after he arrived, already had more friends than all of us combined. Thank you for always being up for a laugh, and for being an incredibly helpful colleague and friend.

I would also like to thank Eleanor. We met at a compulsory event in our first year that neither of us wanted to be at, and immediately hit it off. Thank you for being my cheerleader throughout this journey, and for proofreading pretty much everything I have ever written. I hope I can do the same for you.

I am also profoundly grateful to my family. To my mum, who has always been at the end of a phone, ready to talk about everything and nothing, thank you for your unconditional support. Thank you for always being interested in my work, even the incredibly dull parts. Thanks also go to my brothers and sisters, Fred, Jessie, Jack, and Meg. Thank you for putting up with me during this time. I hope you will grow used to having to call me Dr Abby.

Lastly, I could not have done this without the support of my husband, Max. A PhD is often a lonely journey, but you have been by my side every step of the way. Thank you for celebrating the highs and commiserating the lows, and for listening to the minutiae of my research. Thank you for being the sounding board for ideas, and the proofreader of awkward emails. This would have been impossible without you.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

A form of Chapter 3 is published with the following reference:

Ramsøe, A., van Heekeren, V., Ponce, P., Fischer, R., Barnes, I., Speller, C., & Collins, M. J. (2020). DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *Journal of Archaeological Science*, *115*, 105080.

A form of Chapter 4 is published with the following reference:

Ramsøe, A., Crispin, M., Mackie, M. *et al.* Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. Sci Rep 11, 7795 (2021). https://doi.org/10.1038/s41598-021-87125-x

Additionally, during my PhD I participated in the following publications:

- Mackie, M., Hendy, J., **Lowe, A. D.**, Sperduti, A., Holst, M., Collins, M. J., & Speller, C. F. (2017). Preservation of the metaproteome: variability of protein preservation in ancient dental calculus. *Science and Technology of Archaeological Research*, 3(1), 74–86.
- Mays, S., Roberts, D., Marshall, P., Pike, A. W. G., van Heekeren, V., Bronk Ramsey, C., Dunbar, E., Reimer, P., Linscott, B., Radini, A., **Lowe, A.**, Dowle, A., Speller, C., Vallender, J., & Bedford, J. (2018). Lives before and after Stonehenge: An osteobiographical study of four prehistoric burials recently excavated from the Stonehenge World Heritage Site. *Journal of Archaeological Science: Reports*, 20, 692–710.
- Charlton, S., **Ramsøe, A**., Collins, M., Craig, O. E., Fischer, R., Alexander, M., & Speller, C. F. (2019). New insights into Neolithic milk consumption through proteomic analysis of dental calculus. *Archaeological and Anthropological Sciences.* https://doi.org/10.1007/s12520-019-00911-7

- Wilkin, S., Ventresca Miller, A., Taylor, W. T. T., Miller, B. K., Hagan, R. W., Bleasdale, M., Scott, A., Gankhuyg, S., **Ramsøe, A.**, Uliziibayar, S., Trachsel, C., Nanni, P., Grossmann, J., Orlando, L., Horton, M., Stockhammer, P. W., Myagmar, E., Boivin, N., Warinner, C., & Hendy, J. (2020). Dairy pastoralism sustained eastern Eurasian steppe populations for 5,000 years. *Nature Ecology & Evolution*, 4(3), 346–355.
- Morton-Hayward, A. L., Thompson, T., Thomas-Oates, J. E., Buckley, S., Petzold, A., **Ramsøe, A.**, O'Connor, S., & Collins, M. J. (2020). A conscious rethink: Why is brain tissue commonly preserved in the archaeological record? Commentary on: Petrone P, Pucci P, Niola M, *et al.* Heat-induced brain vitrification from the Vesuvius eruption in C.E. 79. N Engl J Med 2020;382:383-4. DOI: 10.1056/NEJMc1909867. *STAR: Science & Technology of Archaeological Research*, 1–9.

During my PhD I was lucky enough to undertake a Researcher Employability Project (REP), funded by WRoCAH. I decided to spend this time working at the British Antarctic Survey. During my time there, my supervisors and I authored the following publication, which has no relevance to this thesis, but relevance to my PhD journey:

Ramsøe, A., Clark, M. S., & Sleight, V. A. (2020). Gene network analyses support subfunctionalization hypothesis for duplicated hsp70 genes in the Antarctic clam. *Cell Stress & Chaperones.* https://doi.org/10.1007/s12192-020-01118-9

Chapter Contributions

All chapters which are not explicitly mentioned below were the sole work of AR.

Chapter 3: deamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes:

- Data processing, analysis, interpretation, software tool development and manuscript preparation were completed by AR.
- Paola Ponce contributed skeletal material
- Vivian van Heekeren performed the proteomic extraction and initial data screening.

- Camilla Speller and Matthew Collins assisted with the interpretation of the results and the preparation of the manuscript.

Chapter 4: Milking it? Assessing the degradation of ancient milk proteomes:

- Data processing, analysis, interpretation and manuscript preparation were completed by AR.
- Beatrice Demarchi provided samples for the contamination experiment.
- Mia Crispin and Krista McGrath performed the extraction of the contaminated samples.
- Meaghan Mackie and Rosa R. Jersie-Christensen performed the extraction of the skimmed milk powder.
- Roman Fisher performed the laboratory experiments.
- Camilla Speller and Jessica Hendy assisted with result interpretation and manuscript preparation.

Chapter 5: Collagen type I degradation in limed skins:

- Data processing, analysis, interpretation and manuscript preparation were completed by AR.
- Sean Doherty designed and performed the liming experiment, and assisted with manuscript preparation.
- Sarah Fiddyment performed laboratory experiments.
- Matthew Collins assisted with data processing, analysis and manuscript preparation.

Chapter 1: Research Framework and Aims

1.1 Research Context and Significance

The field of palaeoproteomics provides invaluable insight into the past, with recent studies shedding light on many facets of archaeology and palaeontology, such as past diets and diseases. However, as palaeoproteomics is a relatively new field of study, there is no commonly agreed upon standard for the authentication of results. As a consequence of this, some studies report remarkable findings which may contradict archaeological understanding. Therefore, it is of the utmost importance to develop tools and methodologies that provide authentication of ancient proteomics results.

This can only be achieved through the analysis of ancient and modern proteomes in order to study the patterns of diagenesis that have affected ancient proteins through time. Increased understanding of the characteristics of truly ancient proteins can therefore be used to differentiate between ancient and modern proteomes.

1.2 Aims and objectives

This thesis explores the molecular traces of diagenesis acting on proteins in the archaeological record, in order to:

1) Understand these diagenetic forces, and
2) Use this knowledge to develop tools and methodologies to authenticate ancient proteomes

To achieve these aims, this thesis will:

A. Develop and demonstrate the effectiveness of site-specific deamidation tools (Chapter 3)

B. Investigate the deamidation in archaeological and modern milk proteomes, and discuss whether milk can be reliably authenticated using deamidation alone (Chapter 4).
C. Analyse the degradation of collagen in parchment (Chapter 5)
D. Discuss the challenges and potentials for using damage patterns to authenticate ancient proteins (Chapter 6)

1.3 Thesis structure and Outline

This thesis is structured in the form of an introduction (Chapter 2), followed by three standalone research articles (Chapters 3, 4, and 5), and a discussion section that brings the thesis strands together (Chapter 6).

1.3.1 Chapter Outlines

Chapter 2: introduces the key elements of the thesis, including an overview of proteins, with a focus on ancient proteins. It goes on to discuss the methodological approaches and applications of proteomics in archaeology. Lastly, it describes the challenges the field faces.

Chapter 3: deamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes proposes a novel method to authenticate ancient protein results, by differentiating genuine ancient proteins from novel contaminants using site-specific deamidation.

Abstract: Contamination is a potential problem in the study of ancient proteins, either from prior handling of the sample, laboratory consumables, or cross-sample carryover from mass spectrometers. Recently, deamidation of glutamine has been proposed as a measure for assessing the degradation of ancient proteins. Here, we present deamiDATE 1.0, a method for the authentication of ancient proteins using measure of site-specific deamidation rates. We test this approach on shotgun proteomic data derived from bone collagen from modern, archaeological and extinct taxa. We further demonstrate how this method may be used to

differentiate between modern contaminants and authentic ancient proteins using a case study from Neolithic dental calculus.

Chapter 4: Milking it? Assessing the degradation of ancient milk proteomes comprises a comparative analysis of three experimental milk datasets and 274 previously published samples containing archaeological milk. Its focus is whether milk proteomes can be reliably authenticated using deamidation alone.

Abstract: The origins, prevalence and nature of dairying have been long debated by archaeologists. Within the last decade, new advances in high-resolution mass spectrometry have allowed for the direct detection of milk proteins from archaeological remains, including ceramics, dental calculus, and preserved dairy products. However, proteins recovered from archaeological remains are susceptible to pre- and post-excavation and laboratory contamination, a particular concern for ancient dairying studies as milk proteins are potential laboratory contaminants. Here, we examine how site-specific rates of deamidation can be used to elucidate patterns of peptide degradation, and authenticate ancient milk proteins. First, we characterize site-specific deamidation patterns in modern milk products and experimental samples, confirming that deamidation occurs primarily at low half-time sites. We then compare this to previously published ancient proteomic data from six studies reporting ancient milk peptides. We confirm that the extent of deamidation, on average, are more advanced in beta-lactoglobulin recovered from ancient dental calculus and pottery residues. Nevertheless, deamidation displayed a high degree of variability, making it challenging to authenticate samples with relatively few milk peptides. We demonstrate that site-specific deamidation is a useful tool for identifying modern contamination but highlight the need for multiple lines of evidence to authenticate ancient protein data.

Chapter 5: Collagen type I degradation in limed skins: assesses the degradation of collagen through a time series of skins that have been subjected to liming; with a focus on what drives collagen diagenesis.

Abstract: As the primary writing medium in Europe, North America (post-Columbian) and the Near East for over 1000 years, parchment (processed animal skin) is one of the most abundant resources available for the study of past societies and cultures. Tens of millions of parchment skins were manufactured, many of which survive today in archives, libraries and

private collections. It is therefore important that they are conserved in such a way that the textual and biological information they record can be preserved. However, the degradation of collagen in parchment is poorly understood. Here, we elucidate the patterns of degradation in experimentally limed parchment in order to better understand the diagenetic process. It is observed that during liming the peptide count, coverage and deamidation patterns progress at a predictable rate. Furthermore, a hotspot of hydrolytic activity is discovered, and its mechanisms are explored. Lastly, it is shown that the skin removed from stillborn lambs is dramatically different from that of flayed older sheep, but that after liming, there is no detectable difference on the molecular level. This study represents the most in-depth biomolecular analysis of collagen parchment to date, and sheds light on the driving forces behind its degradation.

Chapter 2: Introduction

This chapter introduces the methodological approaches of this thesis and the current state of palaeoproteomic research. Ancient proteins will first be introduced (2.1), with special emphasis on the diagenetic forces they undergo - the understanding of which forms the basis of this PhD. The methods for palaeoproteomics and their development that made this thesis possible will be introduced in Chapter 2.2 In Chapter 2.3, applications of palaeoproteomics will be reviewed, with reference to the different archaeological questions it has the potential to address. Lastly, Chapter 2.4 will discuss the challenges palaeoproteomics faces.

2.1 Ancient proteins

Proteins are made of one or more polypeptides, which are chains of amino acids that are bound together with peptide bonds. Amino acids are organic compounds that contain an amine (NH2) group attached to an alpha-carbon atom next to a carboxyl (COOH) group, and a side chain (R) that varies between the 20 (DNA encoded) amino acids. Proteins are organised into several structural levels. Primary structure refers to the sequence of amino acids in the peptide chain. The secondary structure is the local folded structures within a polypeptide, determined by the pattern of hydrogen bonds between peptide groups. Tertiary structures are groups of one or more secondary structures with a polypeptide backbone. Lastly, quaternary structures are composed of several tertiary subunits. Proteins carry out a diverse set of functions within organisms, including acting as enzymes, facilitating cell signaling, transporting other molecules, and providing structure (Murray *et al.*, 2009).

The study of ancient proteins refers to the retrieval of peptide sequences (chains of amino acids) that are identified from objects in the archaeological record (Cappellini *et al.*, 2014). This section will describe the physical and chemical changes that occur to proteins after deposition, and then go on to describe the role of different factors on protein survival.

2.1.1 Protein diagenesis

Although there are many different ways in which proteins degrade over time, this section will highlight the ones that are most relevant for studies of ancient proteins.

2.1.1.1 Hydrolysis

Hydrolysis refers to the chemical reaction in which a molecule of water breaks a chemical bond, as shown in Figure 2.1. Hydrolysis can act on the peptide bonds which link amino acids together. It therefore splits the protein into smaller peptides, and consumes a molecule of water. This leads to measurable amounts of amino acids that are free - i.e. unbonded to any other amino acid. As the protein undergoes peptide bond hydrolysis increasingly smaller peptides become prevalent, and it is thought that the hydrolysis affecting smaller peptides acts at a different rate than that of the intact protein (Hare 1976).

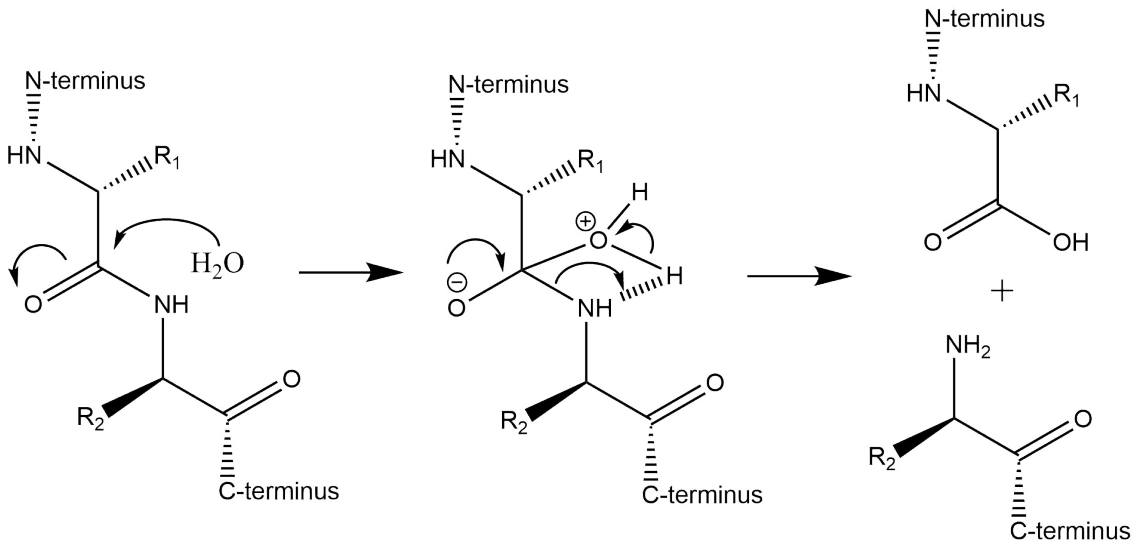


Figure 2.1: *Peptide bond hydrolysis (Image credit: F. Di Gianvincenzo)*

As hydrolysis cleaves peptides into small, unrecognisable fragments, understanding its rate is key to comprehending sequence survival. The rate of hydrolysis is dependent on several factors. Firstly, it is by definition dependent on the availability of water, as water is consumed in the reaction. It also depends on many other factors, such as time, temperature, and environmental pH (Collins *et al.*, 2002). Studies agree that higher temperatures lead to an exponential increase in the rate of hydrolysis (Collins *et al.*, 2002; Hare 1976), though it has been shown that pressure has a differing effect, relative to temperature: below 200-

220°C the rate of hydrolysis is faster at 265 atm (44.1 KJ/mol) than at water steam pressure (98.9 KJ/mol), whereas above this temperature the reverse is true (Qian *et al.*, 1993).

In the early diagenetic stages peptide bond hydrolysis is a pseudo-first order reaction, meaning that the concentration of one of the reactants (water) remains essentially constant (Kriausakul and Mitterer 1980). However, in the late stages of diagenesis or in closed system, this slows to a second-order reaction, either due to a lack of water, or because of a remaining bound fraction of amino acids that are resistant to hydrolysis (Penkman *et al.*, 2008; Demarchi *et al.*, 2013). The effects of temperature on peptide bond hydrolysis is typically described using the Arrhenius equation (although this was strictly developed to explore gas phase reactions) (Arrhenius 1889). Hare's (1976) work on the relative rates of peptide bond hydrolysis show that there are significant differences in the reaction rates of the various amino acids, but very few differences in their activation energies (i.e. temperature dependence), with the exception of valine and isoleucine, whose activation energies are slightly higher. This concurs with the observation that the rate of hydrolysis also depends on the neighbouring amino acid residues: peptide bonds adjacent to valine and isoleucine are known to be much more stable than those next to amino acids such as glycine and serine (Hill 1965). More recently, Tomiak and co-authors (2013) proposed a new method for estimating kinetic parameters, using a model free scaling technique that uses observed rates at different temperatures, and makes no assumption of linearity. They tested this technique using heating experiments and naturally aged coral, and comparing conventional methods of determining the activation energies. It was found that their model free technique was a good predictor of racemisation reaction rates, but overestimated the activation energy of hydrolysis. It was suggested that this difference could be due to the denatured proteins in the artificially aged (heated) samples, which implies that such artificial aging experiments to determine the underlying kinetics of diagenetic reactions may not be valid.

Crisp *et al.* (2013) also compared this model free approach to other models in order to derive the kinetic parameters of racemisation and hydrolysis in ostrich eggshell. They found good agreement between the models on the activation energies for hydrolysis, with the exception of asparagine & aspartic acid, and phenylalanine. It was shown that many amino acids undergo hydrolysis and racemisation at a predictable rate, and that these may be used to estimate the relative age of ostrich eggshell. However, without independent age information, a conversion to numerical age necessitates the accurate calculation of the

Arrhenius parameters, which, in turn, requires further understanding of the underlying mechanisms.

For example, Cappellini and colleagues (2010) characterised the biomolecular composition of archaeological grape seeds found in waterlogged environments. Seeds were recovered from two sites: a 700-800 years CE pit in York, UK, and a 14th to 15th century well in northern Italy. They were able to detect lignin and cellulose compounds in the York seeds, and nLC-MS/MS revealed the presence of six proteins based on 30 peptides. Furthermore, ancient DNA amplification was successful for two of the York samples and one from Italy. They found that the seed storage proteins were particularly affected by hydrolysis, and that specific motifs were particularly prone to cleavage - for example between asparagine and proline (also mentioned by Capasso *et al,.* (1996)), and asparagine and glycine (mentioned by Bada (1985), and Collins, Walton, and King in "The geochemical fate of proteins" in Stankiewicz and van Bergen (1998)).

Hydrolysis is both a driving force behind other diagenetic mechanisms (such as deamidation and racemisation - discussed later in this chapter), and a limiting factor in the detection of peptides using LC-MS/MS. Therefore further understanding of the mechanisms behind hydrolysis will shed light on both how proteins decay, and how best to detect them. In particular, future studies could examine how secondary and higher order structures play a role in facilitating (or protection from) hydrolysis.

2.1.1.2 Deamidation

Deamidation refers to the post-translational modification (discussed in Section 2.1.2 Synthesis) that the amino acids asparagine and glutamine undergo in which they become aspartic acid and glutamic acid respectively. Chemically, this involves the addition of a water group and the removal of an amino group, which results in a mass shift of +0.98402 Da.

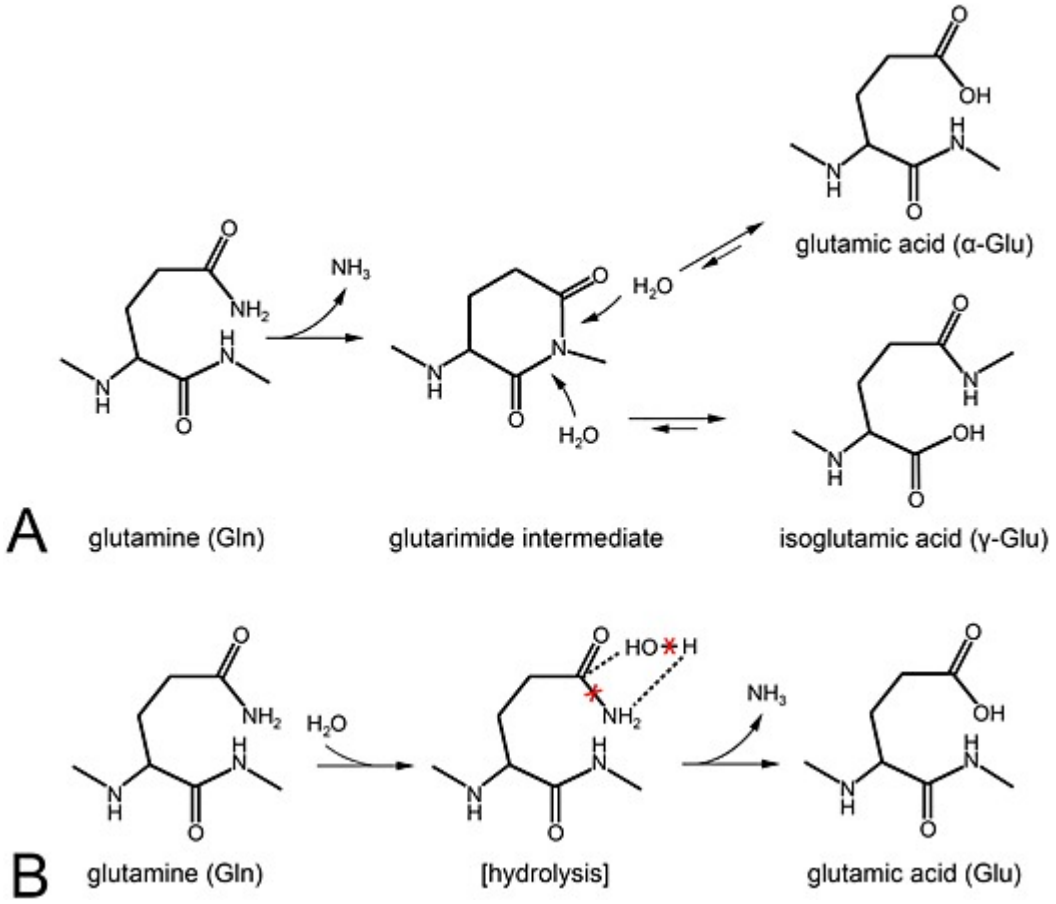


Figure 2.2: *Deamidation of glutamine to glutamic acid through A) a succinimide intermediate and B) Direct hydrolysis. Figure from Schroeter and Cleland (2016).*

The primary pathway involved in the deamidation of asparagine occurs via a succinimide intermediate: when the C-terminus residue's backbone nitrogen atom attacks the asparagine's side chain amide group carbon atom. This forms an intermediate succinimide ring, which then undergoes hydrolysis and forms aspartic acid or isoaspartic acid (Jia and Sun 2017). This process is shown (for glutamine) in Figure 2.2. Asparagine deamidation can also occur via side-chain hydrolysis (Catak *et al.*, 2009).

Asparagine deamidation is well studied because it is known to contribute to aging, disease, and affects the shelf-life of pharmaceuticals (Li *et al.,*, 2010). Glutamine deamidation studies are more scarce, as glutamine deamidates at a much slower rate (the average asparagine deamidation half-time is 81.4 days; whereas the mean half-time for glutamine deamidation is 6,000 days (Robinson and Robinson 2001)), and therefore has little relevance in biomedical studies. However, this slower deamidation makes glutamine deamidation one of the most important reactions in the study of ancient proteins, and has been hypothesized to act as a molecular clock (Robinson and Robinson 2004).

Throughout this thesis, the time it takes for half of a specific amino acid to deamidate will be referred to as its "half-time", rather than "half-life". Although both are used in the literature, half-time was selected as it is used in the Robinson papers (Robinson and Robinson 2001; Robinson *et al,*, 2004) that form the basis for deamiDATE, introduced in Chapter 3.

Glutamine can also deamidate via an intermediate, though this reaction is much slower as the formation of the glutarimide intermediate is less favoured than the succinimide intermediate of asparagine. Because of this, the primary pathway involved in glutamine deamidation is direct hydrolysis (Li *et al.*, 2010), depicted in Figure 2.2. The rate of deamidation is influenced by steric hindrance and charged residues near deamidation sites, and therefore primary (Robinson and Robinson 2001) secondary (Xie and Schowen 1999), and tertiary (Kossiakoff 1988) structure, temperature (Stratton *et al.*, 2001), and pH (Hao *et al.*, 2011; Wilson *et al.*, 2012).

Glutamine's slow rate of deamidation implies that it may be able to assign relative ages to samples, while avoiding further analytic steps (e.g. dating by amino acid racemization or radiocarbon dating). As such, the palaeoproteomics community has shown great interest in this possibility. One of the first studies into this was van Doorn and colleagues' (2012) paper in which they used ZooMS to quantify the deamidation in 911 different bone collagen samples from 50 different sites. By comparing the extent of glutamine deamidation per sample with the thermal age, they showed that the youngest samples had the lowest level of deamidation, while the (thermally) oldest samples portrayed the greatest extent of deamidation. The intermediate samples, however, showed a wide range of deamidation. They propose that glutamine deamidation can act as a litmus test as to whether a sample is thermally ancient or modern.

A follow up study from Wilson *et al.* (2012) used MALDI to assess the extent of glutamine deamidation in 87 bone samples from different sites, with different ages and archaeological considerations. They found that some peptides were always found highly deamidated, and some were almost never detected, irrespective of the sample. Other peptides showed a clear correlation with thermal age, just as in van Doorn *et al.* (2012). Moreover, it was shown that peptides detected in more than one hydroxylation state had deamidation that was highly correlated with thermal age, which implies that the hydroxylation of proline does not affect the rate of deamidation. Interestingly, they detected no meaningful differences relative to the tertiary and higher order structures of the collagen, including between the gap and overlap regions.

Solazzo and colleagues' (2014) paper examined deamidation in eight different peptides in wool textiles using MALDI-TOF-MS. Their dataset included 17 Medieval samples (nine of which were from England, and the remaining 8 from Iceland), and also modern controls, subsets of which were dyed, buried, and thermally aged. They found that the Icelandic samples showed the most deamidation, despite being more recent (also in terms of thermal age) than the English samples. The authors suggest that this could have been due to the hot springs near the excavation site subjecting the artefacts to increased temperatures. Moreover, it was found that five of the eight peptides studied were stable markers, while the remaining three either deamidated too quickly (and were therefore undetectable), or too sensitive to processing methods.

However, Schroeter and Cleland (2016) challenged the conclusion that glutamine deamidation can be used as a rough proxy for relative age. Using data from five previous palaeoproteomic studies, they showed that the high variability of deamidation precludes its use as a method of authentication. They suggest that, as the rate and extent of deamidation is influenced by environmental conditions (a full history of which cannot be known), and the primary and secondary structures of each protein, it cannot be used as a "one size fits all" proof of endogeneity. It is instead suggested that it must be used on a case-by-case basis.

Welker's (2016) study came to similar conclusions as Schroeter and Cleland. They used ZooMS screening to analyse 543 bone and tooth samples from Quinçay, France in order to quantify the spatial and temporal variation in glutamine deamidation in a single site. It was found that there was low or absent chronological resolution, but spatial variation is high. This implies that glutamine deamidation cannot successfully be used as a marker for

archaeological age (and therefore authenticity), but could shed light on specimens that have undergone different diagenetic histories.

Taken together, these results imply that although deamidation of some peptides may be a proxy for age, these marker peptides would need to be carefully curated through studies such as Solazzo's, as they show a high variability. As Schroeter and Cleland suggest, there is more than likely no "one size fits all" method for using glutamine deamidation to authenticate proteins. However, it should also be noted that many of these seminal papers discussed used MALDI, so they may have suffered from the recently reported bias against the detection of glutamic-acid containing peptides (Simpson *et al.*, 2019). One of the most important factors that speaks towards the calculation of relative glutamine deamidation is that it requires no more analysis than regular workflows - there is no extra sampling, nor analytical steps needed. However, as discussed above, more refined understanding into the mechanics of deamidation and the forces that affect it is needed if it is to be used as a cut and dried measure of authenticity.

2.1.1.3 Oxidation

Oxidative post-translational modifications most commonly involve the addition of one or more oxygen atoms to an amino acid side chain. Although all amino acids are susceptible to oxidation (Berlett and Stadtman 1997), methionine, tryptophan, histidine, and cysteine are the most commonly oxidized (Berrill *et al.*, 2011).

Oxidation occurs over time, and is especially well researched in the context of museum studies. Historical artefacts stored in dry, well-lit environments are susceptible to oxidation (Ahmed and Darwish 2012). For example, Solazzo and colleagues' (2013) study found a loss of most amino acids in both historical textiles and artificially aged modern samples, but an increase in the amount of acidic amino acids and cysteic acid - which results from the oxidation of cysteine and cystine (Vanden Berghe 2012). This formation of cysteic acid comes hand in hand with a loss of structural integrity and strength in the textile, as the disulphide bonds have been cleaved apart.

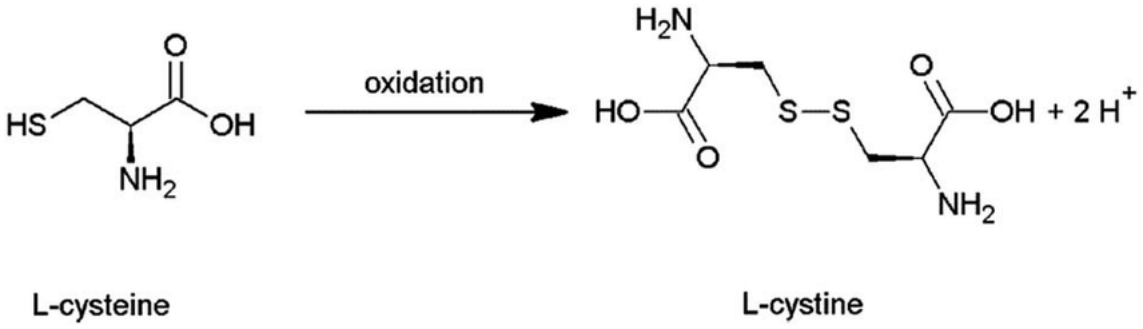


Figure 2.3: Oxidation of cysteine to cystine. Figure from Frank *et al.* (2017).

More studies describing oxidative damage and oxidation are discussed in Section 2.2.2.3.3, as it revolves around how protein sequences are preserved in oxygen-depleted environments.

2.1.1.4 Racemization

All amino acids except glycine have one or more chiral carbon centers .Those with one can exist in two forms of optical isomers: D and L (Demarchi and Collins 2014) . The term racemization is the term used to refer to the interconversion between the two forms, and derives from the term 'racemic' which describes the equilibrium constitution. The rate of racemization is affected by many factors, such as the presence of water, temperature, pH conditions (Liardon and Ledermann 1986), time, and chelation of specific metal ions (Bada 1972). The rate also depends on whether the amino acid is free, at a terminal position, or within a peptide chain, as the rate of most peptide bound amino acids is less variable, and has been shown to be over ten times faster than that of free amino acids (Collins and Riley 2000). Penkman and colleagues (2008) suggested that amino acid geochronology is at its most accurate when analysing just the intra-crystalline fraction of amino acids - as these are less susceptible to rate affecting factors, contamination, and leaching.

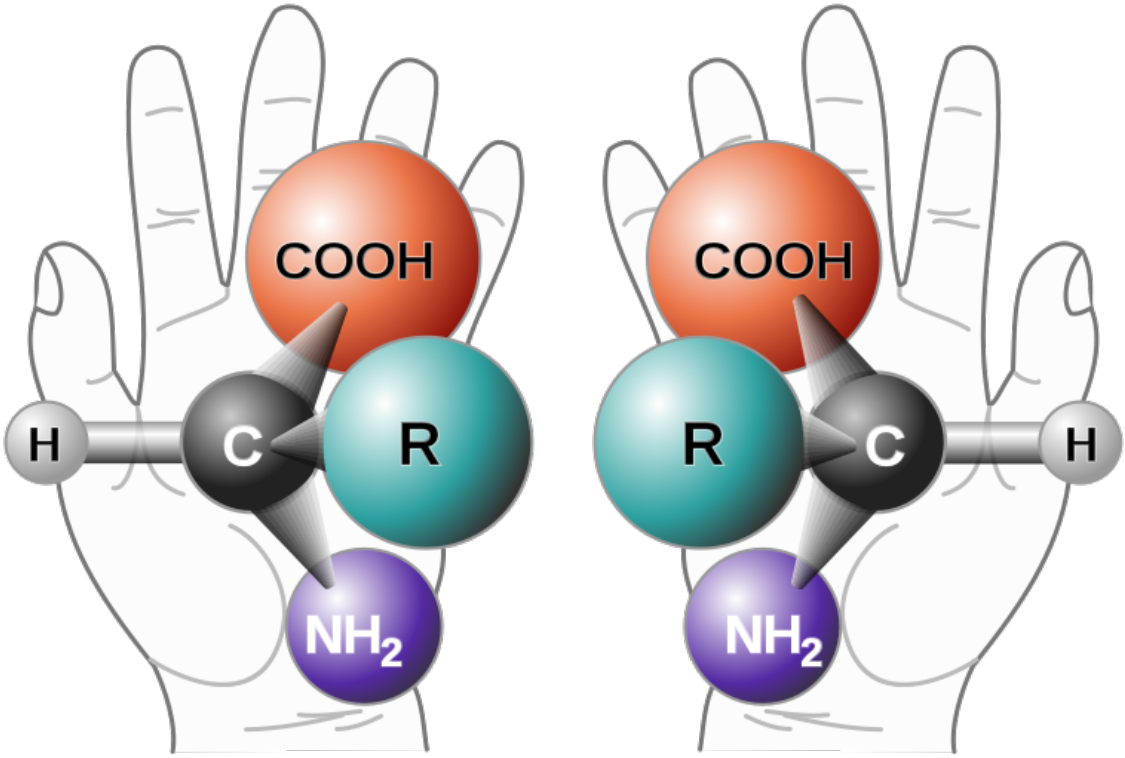


Figure 2.4: A generic amino acid in the L- and D-isomer form (left and right respectively) (Wikimedia Commons).

Uses of amino acid racemization (AAR) include paleothermometry and dating. If the age of a sample is known, the increase in the D-isomer form can be used to make inferences on paleotemperature (Miller *et al.*,1997; Collins and Riley 2000). An example of such is a study by Penkman *et al.*, (2011). The authors analysed 740 examples of the operculas from the gastropod *Bithynia*, from 74 British sites spanning the Quaternary. Racemization in glutamine/glutamic acid, asparagine/aspartic acid, serine, alanine, and valine was quantified, and it was shown that these racemize at different rates, therefore enabling increased resolution in dating. This approach added resolution to the timings of the interglacial stages. Lastly, this was combined with data of the occurrence of different types of tools, which showed that human occupation of Britain occurred in distinct phases, and provided more evidence to the hypothesis that humans were absent from Britain during the Last Interglacial.

Conversely, if the age of the samples is not known, the ratio between the L- and D-isomer form can be used as a relative dating technique. A (2019) paper by Dickinson and colleagues explores this in tooth enamel, calculating the extent of racemization in fossil and modern heated enamel. They found that intra-crystalline amino acids exhibiting closed-system behaviour can be isolated from enamel, and that these follow predictable trends of

racemization under heat treatment. This has huge implications for the field of archaeology, as it implies AAR can be used as a relative dating technique for timescales over 2.8 million years, which far eclipses radiocarbon dating's 50,000 year limit. However, the authors also found that different patterns were observed in the fossil enamel, which, as Tomiak and colleagues also noted (2013), implies that there are different mechanisms dominating at different temperatures, and that the correct interpretation of heating studies necessitates comprehensive understanding of the forces of diagenesis.

In this way, AAR is also a method for the authentication of samples, as modern contaminants are unlikely to have similar D/L values to genuinely ancient proteins (Demarchi and Collins 2014). Additionally, as many amino acids appear to racemize at a predictable rate in closed systems, measures of AAR are less variable than those of deamidation (another proposed measure for authenticity and relative age, discussed above). However, a disadvantage is that AAR cannot discriminate between proteins in complex mixtures. For example, given a single sample, deamidation can be calculated for each protein group - exogenous and endogenous to the source - while AAR represents the sample as a whole. Although AAR represents a useful complementary tool, it has not been focused on in this thesis.

2.1.2 Preserving sequence

Proteins are under a constant barrage from various pathways of degradation. However, there are also many mechanisms that preserve protein sequences, including protein structure, surface, and environmental factors. Understanding these is key to understanding why some proteins are exceptionally preserved in the archaeological record.

2.1.2.1 Role of protein structure

Structure plays a large role in protecting sequences from chemical degradation (Ma *et al.*, 2008). For example, collagen, widely regarded as the ancient protein with the most longevity in mammals (van Doorn *et al.*, 2012), is uniquely stable, due to its triple-helical arrangement and intra- and inter-molecular crosslinks (Némethy and Scheraga 1986; Wojtowicz *et al.*, 1999; San Antonio *et al.*, 2011). This tight structure means that collagen is less susceptible to damage. For example, in intact collagen glutamine cannot readily

deamidate via condensation (i.e. via a cyclic intermediate) due to the interatomic distance required for the heterocycle to form (van Doorn *et al.*, 2012; Kato *et al.*, 2019). However, the breakdown of the triple-helix structure leads to a catastrophic breakdown of the collagen chains (Smith *et al.*, 2007).

Despite its obvious influence in collagen preservation, no studies have as of yet performed a comprehensive investigation into how secondary and higher order structures affect protein preservation. This is sorely lacking, as proteins of different structures are highly likely to be differentially preserved. However, designing such a study would not be simple. As discussed above, the kinetics acting on proteins in artificial aging (e.g. heating) experiments are likely to be markedly different to those at low temperature.

2.1.2.2 Role of surface

Proteins that bind to mineral surfaces can exhibit exceptional preservation. The oldest (widely accepted - see (Asara *et al.*, 2007) and (Buckley *et al.*,2008)) peptide sequence ever recovered was from a Tanzanian ostrich egg shell protein discovered by Demarchi and colleagues (2016) dated to 3.8 million years old. Unlike collagen proteins (described above) that survive due to their rigid structure, it is thought that this peptide survived due to its flexibility allowing it to tightly bind to the mineral surface of the egg shell.

Discussed in detail in sections 2.3.1.2.1 Dental Calculus and 2.3.2.1 Pottery respectively, dental calculus and ceramic vessels also represent mineral substrates commonly used in palaeoproteomics studies. Both represent well-tapped sources of dietary proteins, yet to date there has not been a study exploring the way in which proteins bind to these substrates.

2.1.2.3 Role of environment

The immediate environment of the artifact also plays an important role in the preservation of the proteins associated with it, including the presence of water, oxygen, the temperature, and the special case of tar pits.

2.1.2.3.1 Environments that restrict water

Environments that restrict water have been shown to dramatically improve DNA preservation in bone (Gotherstrom *et al.*, 2002). Arid environments have also been found to lead to exceptional protein preservation. For example, Solazzo and colleagues' (2016) study found bovine collagen and blood proteins, and plant-origin proteins in artefacts and remains excavated from Nahal Hermar cave in Israel (8200 - 7300 BCE), which is in a semi-arid desert area. Moreover, small leucine-rich repeat proteoglycans (SLRPs) were identified in two samples. This is remarkable, as these small proteins (which are inherently less likely to survive in the archaeological record) have not been detected in samples from many other well-preserved sites (including in a waterlogged environment (Bleicher *et al.*, 2015), but were found in permafrost bones (Cappellini *et al.*, 2012; Orlando *et al.*, 2013). Further examples include the incredibly well preserved Bronze Age mummies in the Tarim Basin in China, the DNA analysis of which shed light on ancient population movements (Li *et al.*, 2015).

As discussed in Section 2.1.1.1 Hydrolysis, in environments without water, hydrolysis proceeds as a second-order reaction. As hydrolysis facilitates other diagenetic mechanisms, it is likely that this is the driving force behind exceptional preservation of proteins in dry environments. However, it does not follow that samples in wet or waterlogged environments are not well preserved, as these conditions are anoxic.

2.1.2.3.2 Oxygen

Oxygen is a reagent in many mechanisms of degradation (Höss *et al.*,1996). For example, carbonylation can lead to a reduction in tryptic peptide recovery, as it converts arginine (and proline) and lysine to glutamic semialdehyde and aminoadipic semialdehyde respectively, which removes the motif that trypsin recognises (Cappellini *et al.*, 2010).

Cappellini and colleagues (2012) used LC-MS/MS on a 43,000 year old mammoth bone, and were able to characterise 126 unique proteins. Moreover, they found multiple lines of evidence of oxidative damage to the bone. Firstly, tryptophan oxidation products, such as kynurenine, were identified, which have previously been detected in photo-yellowed wool

(Bringans *et al.*, 2006). Furthermore, they also detected aminoadipic acid from lysine, which increases *in vivo* in septicemia as a reaction to the invading microorganisms - its presence in the mammoth bone could therefore be associated with the decomposition of the mammoth or its deposition in the soil. Another example of oxidative damage in artefacts is Mackie and colleagues' (2018) study of one of the layers of a 14th century Italian wall painting: Ambrogio Lorenzetti's *A Group of Four Poor Clares*. Samples from an unknown proteinaceous layer were analysed with LC-MS/MS, and it was found that half of all the amino acids had undergone oxidation, including all of the chromophoric (light-absorbing/reflecting) residues. Additionally, as with the mammoth, evidence of the oxidation products of tryptophan was found.

Just as oxygen and oxidation reactions can cause damage to proteins, the lack of oxygen can therefore protect protein sequences. For example, Cappellini and colleagues (2010) paper on archaeological grape seeds (also discussed in 2.2.1.1 Hydrolysis) found exceptionally preserved proteins, due, in part, to the anoxic environment of the waterlogged seeds. von Holstein and colleagues (2016) investigated 83 textile and 59 bone samples in order to ascertain whether the original isotopic composition is preserved when samples have been preserved by anoxic waterlogging (and therefore the suspension of aerobic microbial activity). They found a loss of hydrophilic residues, and an increased identification of hydrophobic amino acid residues. It was also found that this (and other) degradation caused by waterlogging did not alter the isotopic composition, nor did it obscure the geographical origin of the samples. However, wet environments are not always conducive to preservation. Excavations in the 1950s (Clark 1954) and 1980s (Mellars and Dark 1998) of the wetland Mesolithic site of Star Carr have uncovered exceptionally preserved organic finds. However changes in the water table in the 2000s caused fluctuating levels of water in the previously waterlogged site. This contributed to extensive and rapid deterioration of organic artefacts at the site (High *et al.*, 2015; High *et al.*, 2016).

These combined studies suggest that oxygen and oxidation reactions are a major pathway of diagenesis in archaeological proteins, and that environments that restrict or exclude oxygen, either through waterlogging or other mechanisms, can lead to exceptional preservation of organic material. However, as at the Star Carr site, many of these environments are under threat due to changes in the use of the surrounding land.

2.1.2.3.3 Tar pits

An example of a unique environment that both restricts oxygen and water are tar pits. In particular, the La Brea tar pits in California represent a collection of over a million Late Pleistocene bones from 231 species (Stock 1930). As well as being anoxic and hydrophobic, samples buried in tar rich sediments are also protected from the effects of weathering or scavengers (Gold *et al.*, 2014).

Although these seem to be exceptional conditions for biomolecular preservation, the reagents that are typically used to remove the asphalt from the samples have been found to also hinder the retrieval of DNA (Gold *et al.*, 2014). This has also been reported to be a problem for proteomic and lipid analyses of the La Brea specimens. A recent study by Colleary and colleagues (2021), who aimed to detect proteins and lipids in mammoth bones of varying ages and depositional environments, including one from La Brea. However, surface mass spectrometry (TOF-SIMS) was hampered by the layer of tar covering the sample, which was also detected as heavy fluorescence in its Raman spectrum. Unfortunately, the presence of the tar layer also precluded the possibility of lipid analysis.

However, other recent studies have had slightly more success. A (2021) paper by Perri *et al.* attempted to shed light on the evolutionary history of dire wolves, which included the analysis of samples originating from La Brea. Although they were unable to extract any DNA, they reported that one sample yielded a collagen sequence. Lastly, and on a more promising note, 34 purified collagen extracts left over from the radiocarbon dating of bone from La Brea (and 133 other bone samples), were analysed by Presslee and colleagues (2021). They found that these collagen extracts, that would normally be discarded, all yielded high quality collagen sequences. Although analysis of non-collagenous proteins was impossible due to the purified nature of the extracts, this study shows that the retrieval of biomolecular data from La Brea (and other tar pits) is possible.

2.1.2.3.4 Temperature

The lower the temperature an artefact is stored at, the better the preservation of the proteins. This is because the reaction speed of the mechanisms that are responsible for diagenesis is lowered. Smith and colleagues (2003) suggest that the "thermal age" (chronological age corrected to a constant 10˚C) of a sample is a useful screening tool for determining the preservation of any DNA within it. While useful as a generality, this suggestion may be too simplistic, at least in the case of ancient proteins, as we have seen the multitude of factors that can contribute to a protein's degradation or lack thereof.

Nevertheless, it is undeniable that lower temperatures can lead to exceptional biomolecular preservation. The oldest DNA ever sequenced is was extracted from a 1.1 to 1.65 million year old mammoth tooth preserved in the eastern Siberian permafrost (van der Walk *et al.*, 2021). The oldest undisputed collagen sequence also originates from the northern hemisphere - the 3.5 million year old giant camel found on Ellesmere Island, Nunavut (Rybczynski *et al.*, 2013). Lastly, ancient RNA has also been recovered from a 14,300 year old canid from the village of Tumat in Siberia, Russia (Smith *et al.*, 2019).

Mackie and colleagues' (2017) paper investigated the differences in preservation of proteins entrapped in dental calculus in three different Roman-period sites - two in England, and one just outside Rome, Italy. As the Italian samples would have been subject to greater temperatures over (the same period of) time, one could expect that they would be more poorly preserved. However, the authors found no systematic trends in protein variability between sites, as it was outweighed by individual variability. They suggest that their study did not have a large enough sample size to elucidate differences driven by climatic factors, and conclude that more systematic studies must be conducted in order to identify the impact of *in vivo* mechanisms and post-depositional degradative processes.

This section has introduced the concept of proteins surviving in the archaeological record, and the main forces of diagenesis behind their decay, including hydrolysis, deamidation, oxidation, and racemization. It went on to discuss the ways that protein sequences are preserved, and the role of different factors involved, including structure, surface, and environment. Understanding the mechanisms behind protein breakdown and survival is key to understanding the characteristics of truly ancient proteins

2.2 Methods

This section will discuss the methods involved in studying ancient proteins, from the inception of the field until now. This will take the form of two main subsections: the methods that existed before mass spectrometry was applied to ancient proteins, and mass spectrometry itself.

2.2.1 Analysis of fossil protein prior to mass spectrometry approaches

The first ancient amino acids were isolated by Abelson in 1954, who showed that amino acids in shells were stable over geological time-scales (Abelson 1954). By the next decade, Hare and Abelson (1968) were able to detect that amino acids in fossil shells existed in two isomers: L and D. This formed the basis for the amino acid racemization. Amino acids were further detected in fossil cephalopods by de Jong *et al.* (1974), who used immunological cross-reactivity to identify epitope matches. However, none of these approaches uncovered protein sequence information.

The first ancient protein sequence was obtained by Huq and colleagues (1990), where they were able to sequence the first 16 residues of osteocalcin from a well preserved moa bone using Edman degradation. This is the only ancient protein sequence ever recovered using Edman degradation, as it required a high concentration of purified and undamaged proteins. As discussed in the earlier sections of this thesis, ancient proteins exist in low abundance, in complex mixtures, and in various states of degradation, so the use of Edman degradation is unfeasible for most samples (Cappellini *et al.*, 2018).

2.2.2 Mass spectrometry

Mass spectrometry, which allows the reliable identification of peptide sequences, was first applied to ancient proteins in 2000 (Ostrom *et al.*, 2000). In mass spectrometry, peptides are ionised and can be identified due to differences in their behaviour in a vacuum (Steen

and Mann 2004). It is more sensitive than Edman degradation, and more robust, and therefore more suitable for ancient samples (Cappellini *et al.*, 2014).

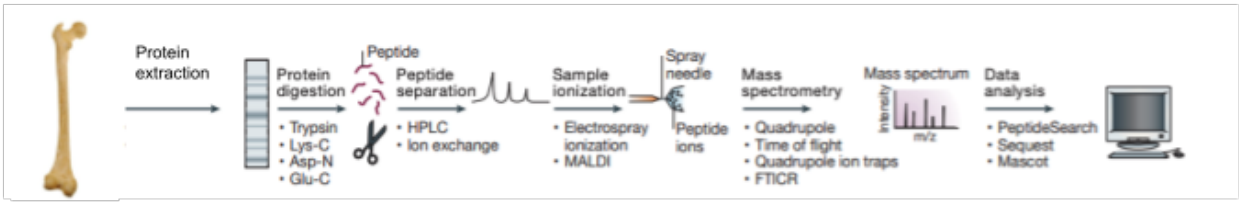


Figure 2.5: Shotgun proteomics workflow. Modified from: (Steen and Mann 2004)

A typical shotgun proteomics workflow is shown in Figure 2.5. Firstly, proteins are extracted from the sample. Typically these extraction protocols are specialised depending on the substrate, and there is a considerable amount of work going into optimising palaeoproteomic sample preparation (Cleland 2018; Le Meillour *et al.*, 2018; Cersoy *et al.*, 2019; Cleland 2018; Schroeter *et al.*, 2019). As the method of extraction will affect the amount and type of proteins identified (Schroeter *et al.*, 2016), specific methods must be chosen based on the type of substrate and end goal. Samples with mineral components must typically also undergo demineralisation in order to liberate the proteins from the mineral.

Proteins are often then fragmented into peptides in order to allow for more accurate analysis in the mass spectrometer, as whole proteins do not all behave the same way in solution due to their secondary, tertiary, and quaternary structure. The most common method is digestion by trypsin, which cleaves peptides at the carboxy side of arginine and lysine (Burkhart *et al.*, 2012). This creates peptides that are around 14 amino acids long, and therefore in the preferred mass range for accurate analysis in the mass spectrometer (Steen and Mann 2004). However, the digestion step is not used in all workflows. For example, digestion-free extraction in enamel has been found to result in higher rates of protein recovery (Cappellini *et al.*, 2019; Welker *et al.*, 2019; Welker *et al.*, 2020).

There are three main components to mass spectrometry: ionization, the mass analyser, and the detector. Purified peptides are separated, and then ionized through a number of different methods, including matrix assisted laser desorption ionization (MALDI) or electrospray ionization (ESI). Once the charge fragments reach the mass analyser, they are separated by size and charge (mass to charge, *m/z* ratio). There are many different strategies to achieve this, including calculating the time accelerated ions take to reach the detector (time-of-flight, TOF), the stability of their trajectories in the electric fields

generated by four parallel rods (quadrupole), or by trapping ions in orbit around a central electrode (orbitrap). The detector records the charge of each ion, and the resulting masses are represented by spectra, where each peak is one *m/z* ratio.

Peptide mass fingerprinting (PMF) is the technique in which proteins are identified just from this spectra, which can be matched with a database of known spectra in order to determine the origin protein. This technique can be used for species identification from bone collagen (Buckley *et al.*, 2009), also known as zooarchaeology by mass spectrometry (ZooMS). This technique works well as collagen is the main protein in bone, but when identifying peptides in complex mixtures of less abundant peptides it lacks diagnostic accuracy (Naihui *et al.*, 2021).

This is where tandem mass spectrometry, also known as MS/MS comes into play. The fragment ions - the original spectra - are further fragmented in another mass analyser, either by collision-induced fragmentation (CID) or electron-transfer dissociation (ETD). This generates a series of masses (product ions) that reveal the sequence of the peptides (Steen and Mann 2004).

There are two ways of identifying the peptides from LC-MS/MS based approaches. The first is similar to peptide mass fingerprinting in that it involves comparison with a reference database. The reference database is provided by the researcher (after retrieval from published databases, e.g. UniProt (UniProt Consortium 2015)), and should contain sequences of all proteins that may feasibly be identified in the sample (though this is a tricky balance, too large a database needlessly increases search space and can obscure important results; whereas too small a database can cause spurious matches or mismatches). These sequences undergo theoretical, or *in silico*, digestion by the same enzyme used in the real digestion process. All i*n silico* peptides that match the precursor ion charge (including possible mass changes caused by post-translational modifications) are then further fragmented into all possible product ions. These theoretical masses are then compared to the product ion spectra from the MS/MS, and the most likely match is found (Steen and Mann 2004).

Sometimes, especially in palaeoproteomics, researchers are interested in identifying sequences that are not yet known, for example in the study of extinct animals (Welker *et al.*, 2015). Here, it is possible to choose sequences from modern species that are phylogenetically related to the target species, but that can introduce errors due to single

amino acid substitutions (SAPs) between the species. Instead, *de novo* sequencing reconstructs the peptide sequence directly from the product ion spectra. By calculating the mass between the peaks it is possible to match these mass differences to amino acids, and therefore infer the sequence (Presslee *et al.*, 2019). However, as sometimes spectra are partial or incomplete, or could theoretically involve too many different post-translational modifications, sometimes this is intractable (Steen and Mann 2004).

2.3 Applications

This section will discuss the applications of palaeoproteomics, with particular focus on the different substrates analysed in studies to date, and the conclusions researchers can draw from them. I divide this into two main categories: biological remains (2.3.1) and archaeological artefacts (2.3.2).

2.3.1 Biological Remains

Biological remains, both human and animal, store a wealth of information about the past. This section will focus on two main subcategories: bones and teeth. Although other substrates, such as brains (Maixner *et al.*, 2013), have also been studied, mineralised tissues such as bones and teeth are the most commonly surviving biological material on archaeological sites (Gilbert *et al.*, 2005).

2.3.1.1 Bones

Bones are composed of two main components: the organic (primarily collagen) and inorganic fractions (primarily the mineral hydroxyapatite) (Collins *et al.*, 2002). The analysis of archaeological bones by mass spectrometry can provide clarity on the species the bone originated from. This is often not possible with traditional archaeological analyses (such as the examination of morphological differences) where the species in question are too closely related to exhibit measurable differences (Buckley *et al.*, 2010), bones have been worked and have lost morphological features (Bradfield *et al.*, 2019) or the bones in question are too fragmented (Buckley and Collins 2011; Welker *et al.*, 2015). This is

commonly done by the peptide mass fingerprinting method ZooMS, as the differences in the collagen spectra are often sufficient to identify the species. For example, Buckley and colleagues' (2014) study analysed 50 bone samples, spanning 6000 years of history, from British sites on the Atlantic coast. These bones had all been previously identified as either cetacean (the infraorder that comprises whales, dolphins, and porpoises) or more generally as marine mammals. Buckley and colleagues performed ZooMS on the samples and were able to provide genus and species level resolution for a majority of them. Of the samples identified as cetacean, one was found to have been originally misidentified, and was actually a walrus. Of the samples identified as general marine mammals, four of them were actually bovine. This highlights the importance of biomolecular analyses as a tool to complement traditional methods, and therefore provide more knowledge about the species in question, their exploitation (hunting, farming) by people, and the ecological conditions of the region at the time (Murray 2008).

Another area that the proteomic analysis of bone can provide clarity on is phylogenetic relationships. Presslee and colleagues' study (2019) provided clarity on sloth phylogenetic relationships using MALDI-TOF and LC-MS/MS through the analysis of 120 bone samples from 24 different taxa within the suborder Folivora. By interrogating the single amino acid polymorphisms (SAPs) in collagen type 1, researchers were able to uncover the relationship of and between tree sloths, the composition of two superfamilies (Megatherioidea and Mylodontidae) and the date of divergence of the major ingroups. This study was validated by an additional phylogenetic analysis by sequencing the mitochondrial genomes of ten ancient sloths (Delsuc *et al.*, 2019).

Lastly, the extraction and analysis of proteins from ancient human bone can provide information about any disease the person may have suffered from. There are many diseases that leave physical traces on bones, such as cancer (Binder *et al.*, 2014), tuberculosis (Masson *et al.*, 2013; Müller *et al.*, 2016), and arthritis (Entezami *et al.*, 2011). If palaeoproteomic analysis of samples taken from ancient bone could identify proteins unique to disease causing agents, such as viruses, bacteria, or parasites, this would provide evidence that the person suffered from the disease. However, this is not usually possible, as the organisms in question either would not be located in or on the bone, or would not be sufficiently preserved (Müller *et al.*, 2016). Nevertheless, a recent study by Mühlemann and colleagues' (2020) confidently identified smallpox DNA in the skeletal and dental remains of 13 Viking Age individuals - so with the increase in sequencing and analytical power, such discoveries may become more widespread.

However, it is seemingly possible to identify proteins in bone that are involved in the biological response to disease. Sawafuji and colleagues' study (2017) involved the analysis of rib bones from eight people who died in 17th century Tokyo. Proteins deriving from leukocytes (which are involved in mounting immune response) were confidently identified in many of the individuals, implying that they could have been suffering from disease when they died. This is in accordance with other archaeological evidence from Tokyo at the time, namely that it was overcrowded and inhabitants had a low life expectancy (Yamamoto 1989). Additionally, there was known to be an outbreak of parasitic infection in Japan at the time (Maki *et al.*, 2014).

2.3.1.2 Teeth

Teeth are incredibly important sources of information in the archaeological record. Broadly, teeth are made up of a hard layer of enamel that cover the surface, below this is dentin which is softer, and the tissue inside the tooth is known as the dental pulp. Patterns of tooth wear can indicate diet, including the type and form of food (Mahajan 2019); and whether the individual used their teeth as tools (Blakely and Beck 1984). Lastly, analysis of teeth is key to ascertaining the oral health, and therefore general health, of individuals (Smith 2019).

Sex determination in humans is interesting to archaeologists as a tool for reconstructing past societies in terms of demography, identity, and epidemiology (Stewart *et al.*,2017). It is also of interest to determine the sex of animal remains, as this can shed light on ritual importance of animals, human and animal relationships (Nistelberger *et al.*, 2019), hunting, husbandry and farming (Davis *et al.*, 2018). Skeletal sex determination methods typically rely on morphology, but require specific bones to be in tact, and are not reliable for young individuals. Using palaeoproteomic methods, sex determination of remains can be achieved through the analysis of the enamel protein amelogenein. Amelogenin exists in two isoforms - X (AMELX) and Y (AMELY) (Bansal *et al.*, 2012), with AMELX encoded on the X chromosome and Y on the Y. Therefore, the reliable identification of AMELY is diagnostic for biological sex.

Luigi and colleagues (2019) analysed teeth from the individuals in the famous grave "Lovers of Modena", which are the remains of two adults buried hand-in-hand in Modena, northern

Italy, 1600 years ago. It was impossible to tell the sex of the remains due to poor preservation, but the conventional wisdom was that they were a heterosexual couple. However, Luigi and colleagues analysed the enamel from both individuals with LC-MS/MS and confidently identified AMELY proteins in both of them, meaning they were, in fact, both men, leading to a reinterpretation of this presumption.

As with bone, peptides preserved in teeth can also be used to infer phylogeny. One of the most exciting examples of this is Welker and colleagues' *Gigantopithecus* study (2019). *Gigantopithecus blacki* is an extinct hominid species that existed in Asia between the Early Pleistocene and the Late Middle Pleistocene. Welker and team extracted enamel and dentine proteins from one molar and characterised them using LC-MS/MS. Endogenous proteins were found in the enamel, which, in itself, represents the most thermally old (i.e. oldest in terms of thermal age, age corrected to a constant 10°C temperature) mammalian proteins ever recovered. By comparing the sequence of the proteins to the enamel proteome of extant apes, *Gigantopithecus* was found to be an early radiation of pongines, and the absence of AMELY suggests that the individual in this study was female. However, sexing in this manner is not conclusive, as it is impossible to prove the absence of a protein.

Lastly, palaeoproteomic analysis of teeth can also provide insight into disease. Barbieri and colleagues (2017) sampled dental pulp from 16 different individuals that were buried in two cemetery sites in 18th century France. One of these cemeteries had anthropological and historical evidence of the plague, and one did not. The samples were analysed with LC-MS/MS, and four peptides between three samples from the plague site were found to be an identical match to *Y. pestis,* the organism that causes the plague. None of the samples from the non-plague site showed any evidence of the *Y. pestis* proteome. Although this study only identified a few peptides (which is below the suggested guideline of two unique peptides per sample set (Hendy *et al.*, 2018)), and many of the identified *Y. pestis* peptides share identity with other bacteria; it represents an interesting first step into the study of bloodborne pathogens surviving in dental pulp.

2.3.1.2.1 Dental Calculus

While not a component of teeth in itself, dental calculus adhering to teeth has also been shown to preserve endogenous ancient proteins. Dental calculus is the mineralised biofilm

that forms on the surfaces of teeth, mostly made up of plaque and saliva, and is almost ubiquitous in people who lived before modern day dental practice (Warinner *et al.*, 2015). As dental calculus mineralises, it traps molecules, including dietary proteins (Warinner *et al.*, 2014), and the oral microbiome (Warinner *et al.*, 2014), including endogenous proteins and bacteria.

Warinner and colleagues (2014) were the first to publish a detailed analysis of the oral microbiome in ancient dental calculus using palaeoproteomics. They sampled dental calculus from four individuals with mild to severe periodontal disease from a German site from ca. 950 - 1200 CE, and dental calculus from nine modern dental patients. These samples were analysed with LC-MS/MS. The proteomes of three well known periodontal pathogens (*Tannerella forsythia, Porphyromonas gingivalis,* and *Treponema denticola*) were recovered in the ancient dental calculus, and were all found in much higher frequency than in modern healthy controls. Bacteria that cause bacteremia and infective endocarditis (*Aggregatibacter actinomycetemcomitans, Streptococcus mutans*, and *S. mitis*) were additionally identified, as were the bacteria that cause bacterial meningitis and gonorrhea (*Neisseria meningitidis* and *N. gonorrhoeae*), though there was no other evidence that the individuals suffered from any of these conditions.

Studying ancient diets provides invaluable insight into past cultures, including hunting and farming, and therefore the origin of dairying, which has long intrigued archaeologists. Hendy and colleagues (2018) analysed 100 samples from dental calculus from the Iron Age up until 19th century England (some of which (n=38) had been previously published (Warinner *et al.*,2014)) and 14 samples from modern dental patients with LC-MS/MS. A total of 59 dietary proteins were identified (31 across the archaeological samples). These represented beta lactoglobulin(which is a major protein in milk), one non-dairy ruminant protein; and a suite of plant-origin proteins, including oats (*Avena sativa*), peas (*Pisum sativum*), and cruciferous vegetables (*Brassica spp.*).

Biological remains offer a wealth of information about the past. Through the analysis of bones, researchers can find out the species of the sample, clarify phylogenetic relationships, and even detect evidence of disease and immune response. Teeth also offer an insight into disease, and expression of different proteins in enamel can be used to sex individuals. Lastly, dental calculus, formed on teeth *in vivo* and preserved through time, can also trap information about diet and disease.

2.3.2 Artefacts

This section will address previous palaeoproteomic analyses of archaeological artefacts, focusing on three main classes: pottery, parchment and skins. Sometimes artefacts are made from biological remains. For example, projectile points (Knecht 1997) or drinking vessels (Bello *et al.*, 2011) can be made from bone due the availability of this material, the ease of manufacture,, or symbolic or status reasons. Additionally, parchment is made from the skins of sheep, goats, or cows, and has been used throughout history as a writing material (Ryder 1964). As humans have transformed these objects from biological remains into objects of some use, for the purpose of this section they will be referred to as archaeological artefacts.

2.3.2.1 Pottery

The study of pots is an area of great interest in archaeology, as it can give insights into technology, social organisation, trade, art, and food (and therefore farming, hunting, and the history of other species). It can also be used to track the spread of an archaeological culture, which have often been defined by the presence of particular pottery types. For example, the Corded Ware Culture, which existed in Bronze Age Europe, was defined by the cord-like impressions on its pottery (Beckerman 2015). Therefore, finds of this certain pattern of pottery indicates presence of the Corded Ware Culture at a site, or can be used to date other artefacts recovered in the same context.

Pottery can also be sampled and subjected to palaeoproteomic analysis in order to shed light on ancient diets. This was first achieved by Solazzo *et al.* (2008), who identified pinniped and cetacean-specific myoglobin from a potsherd fragment from Point Barrow, Alaska, dating to 1200 - 1400 CE. Hendy and colleagues (2018) sampled ceramic sherds from the West Mound of Çatalhöyük, which dates back to 6000–5600 cal BC (Orton *et al.*, 2018). Three main types of sherds were analysed with LC-MS/MS - calcified deposits, and the inner and outer ceramic walls of the pots. The majority of confidently identified dietary proteins originated from the calcified deposits, and the minority from the inner ceramic wall. This analysis revealed that the pots used to hold a mix of plants (including cereals and pulses), dairy (mostly from sheep and goat), and meat proteins (again, mostly from sheep

and goat). This, combined with complementary studies, elucidates the diet and farming practices of the people who inhibited Çatalhöyük during this time.

2.3.2.2 Parchment and skins

Animal skins also hold a wealth of archaeological information, whether they remain unprocessed, or are prepared and used as parchment. They not only have the potential to uncover information about the animal used, and therefore farming practices, but also give insights into the technologies used at the time (Brandt *et al.*, 2014).

Brandt and colleagues' study (2014) on animal skin garments is a prime example of this. They retrieved 12 samples from garments found in Danish peat bogs, aging from 920 BCE to 775 CE. As these samples were waterlogged and physically protected by the peat, it would follow that they would exhibit exceptional preservation. They analysed the samples with LC-MS/MS and found proteins unique to bovine, sheep, and goat. Additionally, peptides unique to bovine haemoglobin subunit beta were found in one sample, which is a protein that is expressed in the foetus and up to three months after birth. This allows the researchers to pinpoint the time the animal was slaughtered, and therefore adds new perspectives to prehistoric animal husbandry.

Palaeoproteomics is also now often applied to the study of parchment. A key example of this is Teasdale and colleagues' (2017) study. The York Gospels, thought to have been started around the 11th century and added to throughout the years, represent one of the only surviving pre-conquest Gospel books in the UK (Alexander 1986). Teasdale and colleagues sampled eraser rubbings (discussed in more detail in 2.5.2 Destructive sampling) from each of the 167 folia, from both the original gospels, and the later additions (14th century). Using ZooMS, they identified that all but one folia of the original gospels were calf, and all of the later additions were sheep. This sheds more light on medieval animal management and parchment making practices.

2.4 Challenges

This thesis has hitherto focused on many of the important insights palaeoproteomics can bring; however, the discipline still faces many challenges. This section will outline these current challenges, and discuss how they may be tackled.

2.4.1 Protein survival

Proteins do not survive forever. As discussed in 2.2., there are lots of mechanisms that prolong survival, however these just delay the inevitable loss of sequence information. In terms of biomolecular archaeology in general, the survival age range of proteins is actually a strength - DNA's survival is far more limited. The oldest DNA ever sequenced is sampled from a mammoth tooth, thought to be around a million years old; whereas the oldest peptides ever recovered are from Demarchi and colleagues' 2016 (Demarchi *et al.*, 2016), dating to 3.8 million years ago. Current estimates (Cappellini *et al.*, 2014) predict that proteins should survive in the fossil record about ten times longer than DNA does. However, the relationship between protein survival and successful aDNA amplification is not straightforward (Wadsworth *et al.*, 2017).

If the oldest DNA that can ever be recovered is at the million year mark, this means that proteins should be recoverable from around 10 million years ago. If a million years if the upper limit for DNA survival (which, of course, it may not be), it would follow that it should be possible to recover proteins of up to ten million years old.

However, this ignores the issue of thermal age (discussed in 2.1.2.3.4 Temperature). The mammoth tooth that is the source of the oldest DNA was recovered from the Siberian permafrost - meaning its thermal age would actually be much lower than one million years old. This is the opposite case for the 3.8 million year old egg-shell proteins - as they originate from Tanzania, where the average temperature is well above 10°C, their thermal age is actually much older - 16 million years old (Ma)@10 °C. Welker and colleagues' *Gigantopithecus* molar is also thermally ancient, at 11.8 million years old (Ma)@10 °C. The oldest DNA by thermal age was recovered from the Middle Pleistocene site of Sima de los Huesos, Spain (Meyer *et al.*, 2016) - at 250,000 years old (0.25 Ma @10 °C). This is 64 times

younger than the thermally oldest proteins - so perhaps either there are thermally older DNA sources yet to be tapped, or the rule of thumb of proteins having ten times more longevity than DNA breaks down when applied to thermal ages. It is also more than likely that we are yet to find the oldest peptides.

Lastly, an important question in the hunt for the most ancient proteins is simple - why bother? There is a diminishing amount of information that can be discovered as peptide bond hydrolysis occurs. Collagen type 1 is thought to be the protein that can survive the longest in bone (Wadsworth and Buckley 2014), and in some very ancient samples it would follow that collagen will be the only sequence identified. Collagen can provide useful phylogenetic information through analysis of single amino acid polymorphisms (SAPs), so its importance should not be downplayed. However, it is far from a full suite of archaeologically informative information (as discussed in 2.4 Applications), and can come at the cost of the destruction of remains or artefacts.

2.4.2 Destructive sampling

Destructive sampling is defined as "any procedure that causes a permanent change to a specimen". It does not always refer to the total destruction of the artefact - although this could be defined as the far end of the destructive sampling scale, where minimally-destructive methods are at the other end (Squires *et al.*, 2019).

Many see traditional archaeologists and biomolecular archaeologists (particularly those in the field of ancient DNA) as being on opposite sides of the destructive sampling debate. Some archaeologists and curators would argue that no destructive sampling should ever be carried out; whereas biomolecular archaeologists argue that by destroying an infinitesimally small piece of an archaeological artefact, it is possible to learn so much more about the past (Kemp 2015). However, the truth is not so black and white. As sequencing (both by mass spectrometry and next generation DNA sequencing) is largely fast, accurate, and relatively affordable (albeit only possible in a select few centres (Makarewicz *et al.*, 2017)), archaeology is going through somewhat of a rush to analyse samples deemed to be of high-interest (Colwell 2018). For some, the aim is to sequence as much as possible as fast as possible, with little regard for the archaeological implications. However, that being said, the development of non-destructive methods has been the subject of increased focus in

recent years (Lee *et al.*, 2014; Papadopoulou *et al.*, 2007; Thomsen *et al.*, 2009; Shapiro *et al.*, 2012; Rohland *et al.*, 2004); and it is increasingly clear that destructive techniques should only be applied to answer important questions, and when there is a high likelihood of analytical success (Green and Speller 2017).

There are many palaeoproteomic studies that make exceptional use of minimally-destructive sample methods. For example, Demarchi and colleagues (2020) extracted proteins from the surface of an Egyptian mummy using mixed bed chromatographic media embedded on ethylene vinyl acetate membranes (EVA). The EVA membrane was put in contact with the mummy's skin and left for 45 minutes. After it was removed, the skin was inspected, and no alteration of surface (either by eye or microscope) was detected. Many keratin and collagen peptides were detected in the sample taken, as were some bacteria - for example many peptides unique to the family Sphingomonadaceae, some of the members of which can digest aromatic compounds. Additional analysis by GC-MS found the presence of plant resin on the mummy, so it follows that Sphingomonadaceae may have been present, as some members of this family are able to degrade aromatics. This study shows that non-invasive sampling can detect both high copy number proteins (such as collagen and keratin), and also metaproteomes.

Another study is Fiddyment and colleagues' (2015) paper on triboelectric protein extraction from parchment, which analysed over 500 samples of uterine vellum. Uterine vellum is tissue thin, and its composition has long been debated - some believe it is the skin of foetal calves or sheep (Clarkson 1992), whereas others think it is more likely that it is made from smaller animals, such as rabbits or squirrels (Thompson 1956). The sampling was done by the conservators at 54 different libraries and archives, and is simply achieved by gently rubbing the parchment with a polyvinyl chloride (PVC eraser) (which is already a widely accepted conservation method for removing dirt) - this creates an electrostatic charge which mobilises the proteins. The samples underwent species identification using ZooMS. Fiddyment and colleagues found equally strong, if not better, identifications using eraser rubbings than the destruction of parchment - possibly due to the fact that the eraser removes contaminants from the parchment, which are then not analysed. They identified that 68% of the samples were calf, 26% were goat, and 6% were sheep - with no trace of rabbit or squirrel. This study not only provided an insight into uterine vellum and parchment production across Europe; but also pioneered this sampling method, which requires no specialised equipment, knowledge, or sample storage.

Lastly, McGrath and colleagues (2019) used a similar principle in performing species identification using the bags in which the objects were stored. The frictional contact between the plastic bag and the proteins on the artefacts strips the electrons from the protein and results in charge electrification, allowing loose peptides to be pulled away from the bone. They analysed 7 bone points and 8 bone artefacts using four different methods: extracting collagen from the bag the artefact was stored in; extracting collagen from a new storage bag with the artefact in it; rubbing the artefact with a PVC eraser (as above, (Fiddyment *et al.*, 2015)), and destructive analysis. They found that most of the time, analysing the bag alone would produce an identification that agreed with the destructive analysis, although sometimes only identified as a "probable" match, or at a family level, whereas the destructive analysis would detect the exact species. Despite not being perfect, this plastic bag method represents a huge step forward for non-destructive sampling, especially for artefacts that could not undergo destructive sampling.

2.4.3 Maximising information

Where non destructive methods are not possible, researchers have an obligation to maximise the data obtained from the sample (Green and Speller 2017). However, this is not simple - palaeoproteomics is a relatively young field, and as such, researchers have not yet come to an agreement on how to maximise results.

However, this is an ongoing field of research. Considerable work has gone into determining the "best" extraction protocol for ancient samples (Schroeter *et al.*, 2016; Jiang *et al.*, 2007; Le Meillour *et al.*, 2018; Wadsworth and Buckley 2014; Lanigan *et al.*, 2020). For example, two extraction methods used in this thesis are both based on modern proteomics protocols: filter-aided sample preparation (FASP) (Cappellini *et al.*, 2014) and gel-aided sample preparation (GASP) (Fischer and Kessler 2015). A 2018 (Hendy *et al.*, 2018) study was the first to perform comparisons between GASP and FASP for dental calculus. The authors used both GASP and FASP to extract proteins from the dental calculus of two modern humans. They found that GASP led to significantly more peptide identifications than FASP (n=1159, n=1324 vs n=159, n=112 respectively). Moreover, GASP extractions led to a greater diversity of recovered proteins, and identified almost all proteins that were identified by FASP.

However, there is likely not one single best protocol for palaeoproteomics as a whole, due to the diversity of substrates. Cleland and colleagues (2018) put forward a new protocol called single-pot, solid-phase-enhanced sample preparation (SP3), specialised for human bone. They found that FASP was at a disadvantage as it uses a molecular weight cutoff, which means potentially missing small peptide fragments. SP3, however, is not molecular weight dependent, and therefore maximises the number of protein detections in ancient bone.

Maximising information does not only refer to maximising proteomic information. Mackie and colleagues proposed an extraction protocol for the simultaneous extraction of both DNA and proteins from dental calculus (Mackie *et al.*, 2017). A recent study by Fagernäs *et al.,* (2020) proposed another new combined extraction method, and found that although there was lower DNA recovery, the metagenomic composition of the results was unaffected. There was no change in the protein recovery, apart from a slight bias towards hydrophobic proteins. Protocols like these are undoubtedly the way forward, as although they are destructive, they reduce the overall sampling demand.

Lastly, one of the most important ways of maximising information from a single sample is open access sharing of both raw and processed data (along with full bioinformatic workflows) on sites such as ProteomeXchange (Vizcaíno *et al.*,2014). This not only allows researchers to double-check a study's veracity using their own systems, but crucially also invites future analyses with no additional destruction of valuable artefacts or remains (Hendy *et al.*, 2018c).

2.4.4 Analytical Challenges

Palaeoproteomics suffers from several non-trivial analytical challenges at every stage, from sampling to data analysis. The type and location of the sample is often dictated by factors that cannot be controlled (such as selective sampling from objects on display in museums). However, variations in protein recovery within the same bone, driven only by different sampling locations, have been shown (Simpson *et al.*, 2016; Procopio *et al.*, 2017). As mentioned in 2.4.3 Maximizing Information, the extraction protocol should be chosen based on the substrate and the research question, though there are no conclusive guidelines for this choice (Hendy *et al.*, 2018c). Demineralisation is typically (though not always, see (Cleland and Vashishth 2015)) a part of sample preparation, as it separates the organics from the mineral component in samples such as bones or teeth. However, this process has

been found to reduce the diversity of the recovered proteome, and analysis of the demineralised fraction resulted in richer diversity and coverage per weight of sample analysed than the soluble fraction (Schroeter *et al.*, 2016). Furthermore, demineralisation by hydrochloric acid (HCl) has been shown to induce deamidation (Simpson *et al.*, 2016), which is a potential measure of protein authenticity (discussed in detail in sections (2.1.1.2 Deamidation and 2.4.5.1 Authentication).

As the samples typically analysed in proteomic workflows are intrinsically finite and precious, method development is challenging. As discussed throughout section 2.1 Ancient proteins, truly ancient proteins have undergone a barrage of different mechanisms of degradation, which are impossible to unravel. Therefore, developing protocols for ancient samples based on their modern equivalent is far from ideal, as ancient samples are distinct down to the molecular level. Some researchers are able to use ancient samples for method development - an example of which is Lanigan and colleagues' (2020) study on parallel proteolysis using combinations of six different proteases on three Pleistocene bones, one of which was a hominin. The alternative to using "genuine" samples is laboratory induced aging. An example discussed in 2.1.1.1 Hydrolysis is Tomiak *et al.*'s (2013) paper on kinetic parameters, in which they compared the degradation of naturally aged and heated coral. They concluded that the underlying mechanisms of degradation may be different at high temperatures. Heating is not the only way to degrade samples, as discussed in section 2.1.2.3.3 Oxygen, High and colleagues (2016) created laboratory microcosms to mimic the environment of the Star Carr site, showing the rapid degradation of bone and wood.

The final analytical challenges take place in silico - the great advantage of which is that the analyses can be repeated and fine-tuned with no material cost. As this challenge is one of the cornerstones in this thesis, it is discussed in depth in section 6.3.2 on challenges. In short, there are many choices, such as the search software, database selection, modification choice, and digestion mode that are the driving force behind peptide identification, yet this process is complicated, again with no best practice guidelines (Hendy *et al.*, 2018).

2.4.5 Contamination

Another key challenge of ancient protein analysis is contamination. Contamination - the addition of proteins not genuinely associated with the sample - can happen at any stage, from burial to excavation to analysis in the mass spectrometer. For example, many conservation methods rely on protein-based materials, such as animal collagen-based glues (Lopez-Polin 2012; Nicholson *et al.*, 2002), plant proteins in natural resins, and insect proteins in shellac (Hendy *et al.*, 2018). However, these contaminants are often impossible to avoid if the conservation has already happened. Therefore, the form of contamination that is easiest to control, and thus most discussed in the palaeoproteomics community is laboratory contamination, either by humans (McCoy *et al.*, 2019; Oró and Skewes 1965) or laboratory reagents (Hendy *et al.*, 2018a; Hendy *et al.*,2018c).

Hendy and colleagues (2018c) proposed a set of best practices in order to avoid laboratory contamination. They include wearing synthetic clothing, covering exposed skin, the use of a dedicated clean facility, and the use of blank extractions and injection blanks to monitor contamination. On the bioinformatics side, it is possible to discard contaminant proteins by searching against a database that includes common laboratory contaminants (e.g. cRAP (cRAP protein sequences)), as well as adopting bioinformatic strategies that enable an assessment of authenticity.

2.4.5.1 Authentication

Authentication of proteomic results is an ongoing topic of discussion in the palaeoproteomics community (Hendy *et al.*, 2018c). There are three key issues in authentication: a) proper peptide identification, b) correct peptide assignment, and c) protein origin.

The first issue, peptide identification, refers to the assignment of the "correct" peptide sequence to the spectra. This is normally handled by the searching software, which matches theoretical peptide masses to the results generated by the mass spectrometer (discussed in 2.3.2), common examples of which include MaxQuant (Cox and Mann 2008), PEAKS (Ma *et*

al., 2003), Mascot (Hirosawa *et al.*, 1993), and Metamorpheus (Wenger and Coon 2013). However, failure to adjust parameters accordingly can lead to incorrect results - in the form of both false negatives and false positives. Database choice (as mentioned briefly in 2.3.2) is the driving force for peptide identifications - if a protein is not in the database the software uses to match against, the spectra will either not be associated with a peptide sequence, or assigned one that is a poor match. Therefore, the smaller the database, the more "forced" matches. However, searching against a needlessly large database wastes analytical time. The size of the database is increased by the number of variable modifications searched for - there is a fine balance between searching for all the modifications likely to occur and needlessly increasing the search space. Protein sequence identification softwares also include many measures for assuring good quality matches - for example automatic searching against a decoy database, false-discovery rate cut-offs, and score assignment. Using these tools, it is possible to ascertain whether a given peptide genuinely represents the target spectra.

Correct peptide assignment refers to the question of whether a given peptide genuinely originates from the specified protein. Because many peptides are not unique, and are either shared by many proteins, or many species, it is important to ascertain that an identified peptide is sufficiently unique. This can be done using BLASTp (Altschul *et al.*, 1990), and is a common criterion used by palaeoproteomic studies (Buckley *et al.*, 2019; Tsutaya *et al.*, 2019).

The last issue is the question of protein origin. The guidelines set out in the previous two points allow researchers to confidently identify peptide sequences, and therefore proteins, but there is no guarantee that these proteins genuinely originate from the artefact in question - i.e. whether they are contaminants. Currently, there is no widely accepted tool to differentiate between genuine ancient proteins and modern contaminants.

Although it is possible to exclude contaminants bioinformatically, sometimes the proteins researchers are interested in are also common contaminants. For example in studies of ancient diet, casein (a milk protein) is also a reagent in western blot analysis; bovine serum albumin (from blood plasma) is a quantitative standard and a reagent in immunological assays; and lysozyme and ovalbumin (egg proteins) are used as molecular weight markers (Hendy *et al.*, 2018a). Therefore, if these proteins were to be excluded, researchers may miss their genuine presence in artefacts.

Therefore it is necessary to authenticate proteins, that is, to differentiate between contaminants proteins and proteins that genuinely originate from the sample in question. In ancient DNA studies, this is done (in part) by assessing the extent of degradation in sequenced molecules - with the understanding that only truly ancient DNA will exhibit specific patterns of damage (Cooper and Poinar 2000; Gilbert *et al.*, 2005; Ginolhac *et al.*, 2011; Jónsson *et al.*, 2013). This logic is also applicable to the study of ancient proteins - as diagenetic forces (discussed in 2.2.1) act over time. For example, the deamidation of glutamine has been proposed to act as a molecular clock (Robinson and Robinson 2004), and is often reported in palaeoproteomic studies, as modern contaminants are unlikely to portray significantly advanced levels of glutamine deamidation (Welker *et al.*, 2016; Welker *et al.*, 2017; Mackie *et al.*, 2018; Demarchi *et al.*, 2020). However, as discussed in Section 2.1.1.2, and later in Chapters 3 and 4, this is highly variable.

Another method of authentication, proposed by Bleasdale and colleagues (2021) relies on the signature of the recovered proteome. Developed for dental calculus, their Oral Signature Screening Database (OSSD) contains a subset of proteins representing the oral microbiome. If the sample contains a wide range of proteins in the OSSD, other proteins of interest from the same sample, e.g. food-origin proteins, are assumed to be endogenous rather than contaminant.

Authentication can also be achieved through dating of the sample. Methods such as radiocarbon dating, which measures the amount of ^{14}C in a sample, can be used to assign relative ages. Discussed in 2.1.1.4 Racemization, the extent of amino acid racemization can also be used to date samples by comparing the ratio of D- and L- enantiomers (Demarchi *et al.*, 2011), but also to provide authentication to palaeoproteomic data (Presslee *et al.*, 2019; Cappellini *et al.*, 2019; Welker *et al.*, 2019). Both of these methods, however, require separate analytical steps. Furthermore, dating is not a perfect proxy for authentication. Take, for example, dietary proteins recovered from dental calculus. Even if the skull or tooth is dated to the correct time period, individual protein results may still arise from contamination, as dating methods cannot differentiate between proteins in mixtures. Another issue is cross contamination with other ancient artefacts or remains, which would result in a mixture of two or more genuinely ancient proteomes.

The field of ancient DNA handles this case with relative ease, as small sequence differences between reads at the same region of the genome can elucidate cases of cross contamination

(Peyrégne and Prüfer 2020). However, as proteomes, unlike genomes, are often highly conserved between individuals, it is impossible to detect a mixture of proteomes in this way.

Lastly, an often overlooked method of authentication is archaeology itself. Proteomic findings should be interpreted with reference to other lines of evidence from traditional (and other biomolecular) archaeological methods. For example, findings of milk proteins in dental calculus could be strengthened by archaeological evidence of pastoralism, animal burials, or lipids in pottery (Wilkin *et al.*, 2020). Of course, as the aim of biomolecular archaeology is to gain insights that would not have been possible with traditional methods, some palaeoproteomic findings may contradict current archaeological wisdom.

2.5 Chapter Summary

This chapter introduced the field of palaeoproteomics, including the various ways proteins are degraded (2.1.1) and the factors influencing their preservation (2.1.2). It went on to discuss the development of methodology in the field, with a focus on current mass spectrometry methods (2.2.2). Subsequently, the applications of mass spectrometry to archaeology were discussed, with reference to the different substrates and archaeological questions (2.3). Finally, the challenges the field currently faces were outlined (2.4).

Chapter 3: DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes

Abigail Ramsøe[1,2], Vivian van Heekeren[1], Paola Ponce[3], Roman Fischer[4], Ian Barnes[2], Camilla Speller[1,5], Matthew J. Collins[6,7]

Corresponding author: Abigail Ramsøe

Affiliations:

1. BioArCh, Department of Archaeology, University of York, York, UK
2. Department of Earth Sciences, Natural History Museum, London, UK, .
3. PalaeoHub, Department of Archaeology, University of York, UK,
4. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK
5. Department of Anthropology, University of British Columbia, Vancouver, Canada
6. EvoGenomics Section, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark.
7. McDonald Institute for Archaeological Research, Downing St, Cambridge CB2 3ER, UK

Relevance: This chapter seeks to address the problem of authentication of ancient proteins by proposing a method that quantifies the deamidation in recovered proteins.

Abstract: Contamination is a potential problem in the study of ancient proteins, either from prior handling of the sample, laboratory consumables, or cross-sample carryover from mass spectrometers. Recently, deamidation of glutamine has been proposed as a measure for assessing the degradation of ancient proteins. Here, we present deamiDATE 1.0, a method for the authentication of ancient proteins using measure of site-specific deamidation rates. We test this approach on shotgun proteomic data derived from bone collagen from modern, archaeological and extinct taxa. We further demonstrate how this method may be used to differentiate between modern contaminants and authentic ancient proteins using a case study from Neolithic dental calculus.

3.1. Introduction

The study of ancient proteins offers new insights into the past, and is increasingly utilised where DNA studies will fail, i.e. where the DNA in a given sample will have completely degraded, but the protein sequences remain somewhat intact (F. Welker *et al.*, 2015; Cappellini *et al.*, 2019; Welker *et al.*, 2019). Additionally, proteins are tissue specific and can identify more than just the presence of the organism, e.g. tissue specific proteins in foods such as seeds, meat or milk (Hendy *et al.*, 2018b), or those expressed in particular disease states (Jersie-Christensen *et al.*, 2018).

The long history of ancient DNA research has highlighted the importance of minimising and identifying modern contamination in analyses of ancient biomolecules. Like ancient DNA, ancient proteins are also susceptible to modern contamination either from prior handling (including conservation treatments), or through exposure to modern protein sources within the laboratory. Prior to the advent of ancient genomics, authentication of ancient data had been conducted through a variety of criteria, including the use of dedicated ancient DNA workspaces, reproducibility and replication of results by independent research groups, and the phylogenetic analysis of the resulting data (Cooper and Poinar, 2000; Gilbert *et al*., 2005). In particular, Cooper and Poinar (2000) and Gilbert *et al.* (2005) both list 'appropriate molecular behaviour' as one of their criteria for authentication of aDNA, with the amplification of long DNA fragments being treated with caution. Hendy *et al.* (2018c) argued for the use of a similar logic to assess the likelihood of claims for ancient proteins, highlighting the need to critically assess the authenticity of obtained results.

With the advent of next-generation sequencing, authentication measures for ancient genomic data have focused predominantly on the analysis of sequence length distributions (to demonstrate DNA fragmentation through hydrolysis) and observations of post-mortem degradation, particularly the deamination of cytosine at 5'-overhangs. As illustrated in the work of Wall and Kim (2007) it is possible to discriminate between ancient and modern sequences based upon DNA fragment length and the probability of cleavage. In the case of damage induced misincorporations, cytosine deamination is concentrated in regions of DNA most vulnerable to degradation - single stranded overhangs. Software packages such as mapDamage (Ginolhac *et al.*, 2011) , mapDamage 2.0 (Jónsson *et al.*, 2013), and bammds (Malaspinas *et al.*, 2014) were developed to estimate DNA damage and measure cytosine deamination, and the average length

of overhangs and nick frequencies. The power of these approaches have been effective for isolating contaminating DNA from authentic signals (Wall and Kim, 2007; Skoglund *et al.*, 2014) and these have been incorporated into standard strategies for ancient DNA authentication (Key *et al.*, 2017).

In the case of proteomic analysis of modern tissues, researchers already consider potential sources of contamination (Bell *et al.*, 2009); indeed it is standard to include in the search database a list of contaminants commonly found in proteomics experiments and, as such, a number contamination databases already exist e.g. cRAP (the common Repository of Adventitious Proteins), and the MaxQuant contaminant database (Cox and Mann, 2008). However, palaeoproteomics is a relatively new discipline, and the best standard measures to avoid protein contamination or to authenticate the antiquity of identified proteins are still being discussed within the community (Schroeter *et al.*, 2017; Hendy *et al.*, 2018c).

3.1.1 Strategies to authenticate ancient proteins

DNA is composed of four nucleotides and studies of DNA damage have primarily focused on the deamination of cytosine (Lindahl, 1993). However, there are instead 20 amino acids found in proteins, many of which are subsequently modified after translation, offering a much larger canvas upon which to explore degradation patterns. In the case of ancient proteins this is especially important, as damage patterns have been used to argue for or against authenticity. For example, Buckley *et al.* (2008) argue against the persistence of dinosaur proteins, in part, based on the grounds that, unlike the reported Mastodon peptides (Asara *et al.*, 2007), the claimed *T. rex* peptides lacked evidence of deamidation, seemingly contra to the antiquity of the sample. A common rule is that higher levels of deamidation imply older, or at least more damaged, samples, and therefore bulk deamidation has been used as a tool to assess the relative age of samples (van Doorn *et al.*, 2012; Wilson, van Doorn and Collins, 2012; Orlando *et al.*, 2013; F. Welker *et al.*, 2015; Hill *et al.*, 2015; Welker *et al.*, 2017).

Deamidation refers to the loss of an amide group, which converts asparagine to aspartic acid, and glutamine to glutamic acid, resulting in a +0.98 Da mass shift. There are two chemical pathways involved in deamidation of asparagine (N, Asn) and glutamine (Q, Gln): condensation to a cyclic intermediate; and side chain hydrolysis. Rates of deamidation are fairly regular in constant conditions, and have been hypothesized to act as a molecular clock (Robinson and Robinson, 2004). The rate of deamidation is influenced by steric hindrance and charged residues near

deamidation sites, and therefore primary (Robinson and Robinson, 2001), secondary (Xie and Schowen, 1999), and tertiary (Kossiakoff, 1988) structure, temperature (Stratton *et al.*, 2001), and pH (Hao *et al.*, 2011; Wilson, van Doorn and Collins, 2012) all affect deamidation rates.

In order to assess the effect of primary structure on the rate of deamidation, Robinson *et al.,* (2004) used MALDI-TOF to detect deamidation in a series of synthetic peptides. They predicted the half-time of 52 sets of asparagine (N) and glutamine (Q) containing peptides, based upon the variation in the flanking residues (i.e. X-N-Y or X-Q-Y), and using this data they then interpolated the remaining 348 combinations.

It is important to note that asparagine and glutamine deamidate at vastly different rates, with glutamine being considerably slower. For example, the average asparagine deamidation half-time is less than 100 days; whereas the mean half-time for glutamine deamidation is estimated to be more than 60 times longer (Robinson and Robinson, 2001). The differences in the rate of deamidation is a consequence of a difference in the primary reaction pathway. Asparagine is primarily deamidated in flexible regions (e.g. α-crystalline) (Kossiakoff, 1988) and denatured polypeptides such as gelatin by cyclic succinimide (Asu) formation. However, it is argued that this is suppressed within the rigid collagen backbone (van Duin and Collins, 1998). In contrast, glutamine deamidation occurs primarily by sidechain hydrolysis (Robinson and Robinson, 2004). The higher rate of aspartic acid (Asp) deamidation and attendant racemization at flexibile sites means that the increase in the non biological isomer D-Asp is sufficiently fast for the D/L Asp ratio in dentine to estimate the age of individuals (Sirin *et al.*, 2018). However, because aspartic acid and asparagine (together, commonly termed Asx) racemization is so dependent upon higher order structure it is less useful as a long-term geochronometer. Consequently glutamine deamidation has become more widely used as a means to assess fossil age (Leo *et al.*, 2011; van Doorn *et al.*, 2012; Wilson, van Doorn and Collins, 2012; Orlando *et al.*, 2013; F. Welker *et al.*, 2015; Hill *et al.*, 2015). These approaches have either used bulk measurements or from single peptides. However, the extent of glutamine deamidation has been found to be variable in macroscopically degraded samples (Simpson *et al*., 2016) and relatively pristine samples, even from within the same bone (Simpson 2015), and it has also been proposed to be a measure of preservational quality, rather than geochronological age due to its variability over time (Schroeter and Cleland 2016)

Here, we present a novel method to authenticate the antiquity of ancient proteins based upon relative rates of deamidation of glutamine residues. We test this method against a series of known age samples, including modern collagen and bone proteomes, as well as paleontological subfossil

proteomes. Finally, we demonstrate the utility of this approach in distinguishing between contaminant and authentic ancient proteins within complex metaproteomic samples: Neolithic dental calculus. As part of a wider investigation into the dental calculus of three Neolithic skeletons from Whitehawk Camp (2700 BCE), we identify chicken egg proteins alongside proteins consistent with the oral microbiome. The earliest evidence of chickens in Europe is in the late Bronze Age (Perry-Gal *et al.*, 2015), consequently the consumption of chicken eggs in the British Neolithic would be a remarkable discovery. The discovery of the consumption of chicken eggs, some 3000 years before the earliest zooarchaeological record of chicken bone in Britain (Maltby *et al.*, 2018) would represent a molecular milestone. It would significantly predate the earliest report of chicken in Germany, again detected based upon the proteomic investigation of an Iron Age pot from Heuneburg (an early Iron Age (750–400 BCE) hillfort in southwest Germany) (Wiktorowicz *et al.*, 2017). In order to examine this case more closely we developed a novel authentication method based upon relative rates of deamidation of glutamine residues.

3.2. Materials and Methods

3.2.1 Materials

3.2.1.1 Model dataset

In order to evaluate the results of measures of deamidation, we assessed deamidation within a series of known age samples (Table 3.1) first published by Welker *et al*. (2015) No experimental or analytical replicates were performed.

Sample	**Age**	**Source**	Element	**Museum accession *(where available)***
Aardvark *(Orycteropus afer)*	Modern	American Museum of Natural History	Not specified	AMNH 51910
Anteater *(Cyclopes didactylus)*	Modern	American Museum of Natural History	Not specified	AMNH 99199
Anteater *(Cyclopes didactylus)*	Modern	American Museum of Natural History	Not specified	AMNH 99199
Hippopotamus *(Hippopotamus amphibius)*	Modern	Zoological Museum, Natural History Museum of Denmark	Bone	
South American tapir *(Tapir terrestris)*	Modern	Zoological Museum, Natural History Museum of Denmark	Bone	
Toxodon	12,040 CAL YBP	Arroyo Tapalqué , Buenos Aires	Tibia	MACN Pv 17710
Toxodon	11,900 CAL YBP	Tapalqué, Buenos Aires	Mandible	MLP 44-XII-29-5
Macrauchenia	Lujanian (ca. 800,000-11,000 YBP)	near Monte Hermoso, Buenos Aires	Cervical vertebrae	MACN Pv 18952
Macrauchenia	Pleistocene (ca. 2,580,000 to 11,700 YBP)	Río Pilcomayo, Formosa	Thoracic vertebrae	MLP 96-V-10-19
Mylodon darwinii	Pleistocene (ca. 2,580,000 to 11,700 YBP)	Cueva del Milodon	Not specified	MLP 94-VIII-10-32

Table 3.1: *Model dataset used for result comparison. Data previously published by Welker (2015), and represents a wide timescale. Raw MS/MS and PEAKS search files available on ProteomeXchange with identifier PXD001411.*

3.2.1.2 Whitehawk Camp dental calculus

Whitehawk Camp is the remains of a Neolithic causewayed enclosure, thought to be inhabited around 2700 BCE. It was excavated in 1932 by Cecil Curwen, which led to the recovery of three individuals, currently curated at the Brighton and Hove Museum (Museum accession R3699/128). Skeleton I was selected for proteomics analysis of dental calculus. Skeleton I (SK1) is an adult female, (25-30 years at time of death) with a significant degree of dental wear: the incisors were badly worn and highly polished. It has therefore been suggested that she might have used her teeth as tools (Cecil Curwen, 1934).

Figure 3.1: *Image of the Whitehawk Camp Skeleton I. One sample of calculus was sampled from the upper right 1st molar; note the glue used to hold the teeth into place (indicated by an arrow). Image credit: Paola Ponce*

3.2.2 Methods

3.2.2.1 Proteomics GASP protocol

The Gel Aided Sample Preparation Protocol (GASP), as described in (Fischer and Kessler, 2015), was carried out on a portion of the dental calculus of SK1, including a blank extract to act as a negative control.

Dental calculus was removed from the upper right first molar of SK1 using a sterile pick and 12.4 mg of calculus was placed in a 2.0 mL Surelock Eppendorf tube. To decontaminate the sample, the dental calculus was rotated in 500 µL of EDTA solution (4% SDS, 0.5 M EDTA) for 5 minutes, centrifuged (13K rpm) for 1 minute, and the solution discarded. These steps were repeated two more times. The calculus was then ground with a micro pestle and rotated for three days in 1 mL of 0.5 M EDTA until fully demineralised. The sample was centrifuged at 13,000 rpm for 2 minutes and 950 µL of supernatant was removed. Proteins were extracted from the pellet and remaining supernatant following the Gel-Aided Sample Preparation Protocol (GASP) based on Fischer and Kessler (2015) and modified for mineralised samples according to Hendy *et al.* (2018c). Briefly, the pellet and supernatant were incubated with 5 µL of SDS (20%), 45 µL of B-PER (Bacterial Protein Extraction Reagent, Thermo Fisher) and 50 µL of DTT (1 M) at room temperature. Proto-Gel (100 µL at 30%, National Diagnostics) was added with 8 µL of tetramethylethylenediamine (TEMED), and 8 µL of 10% ammonium persulfate (APS) to polymerize the gel. The polymerized gel was shredded and the gel pieces were fixed through the addition of methanol/water/acetic acid solution (50/40/10). A series of washing and drying steps using acetonitrile were then performed to exchange buffers. Proteins were digested in 200 µL of ammonium bicarbonate (0.05 M) and trypsin (5 µL of 0.5 µg/µL) 37 °C overnight. Digested peptides were extracted from the gel through a series of washes with acetonitrile and 5% formic acid solution, and desalted using Millipore Zip-Tips prior to MS/MS analysis. Tryptic peptides were analysed at the Mass Spectrometry Laboratories of the Target Discovery Institute at the University of Oxford on a Q-Exactive tandem mass spectrometer according to previously published specifications (C. Warinner *et al.*, 2014). Q-Exactive analysis was performed after UPLC separation on an EASY-Spray column (50 cm × 75 µm ID, PepMap RSLC C18, 2 µm) connected to a Dionex Ultimate 3000 nUPLC (all Thermo Scientific).

3.2.2.2 Data analysis - Dental Calculus SK1

The raw spectral data, generated from the dental calculus of SK1, was converted to Mascot generic format (mgf) using Proteowizard MSConvert (version 3.0.4743) using the 200 most intense peaks in each MS/MS spectrum. MS/MS ion database searching was performed on Mascot (Matrix ScienceTM, version 2.4.01), against UniProt (release 2017_06) and the ancient Human Oral Microbiome database previously published in Warinner *et al.* (2014). The settings are as follows: the peptide tolerance was 10 ppm, MS/MS ion tolerance was 0.07 Da, with a tryptic search strategy with up to one missed cleavage, and a fixed modification of propionamide (C). Variable modifications included acetyl (protein N-term), deamidation of asparagine and glutamine, methionine oxidation, propionamide N-term and lysine propionamide. Searches were performed against a decoy database to estimate protein false discovery rates, which were adjusted to less than 5%. Proteins were identified with a minimum of two different peptides and a default significance threshold of $p<0.05$. Identification of proteins associated with the oral microbiome, human host, and putative dietary proteins were undertaken following the methods described in Hendy *et al.* (2018). In addition to a suite of proteins associated with oral microbes and human saliva, this search yielded multiple peptides deriving from putative dietary proteins, including chicken ovalbumin, lysozyme, and vitellogenin. The LC-MS/MS raw data and Mascot search results are available in the public database on the MassIVE repository (ID: MSV000084284; http://massive.ucsd.edu). File description and run order are presented in Supplementary Table 3.1)

Subsequently, the dental calculus raw files were searched using MaxQuant version 1.6.2.6a against an in-house database of chicken egg proteins and human endogenous proteins commonly found in dental calculus studies (Appendix A, File 1). In order to ensure high-quality matches, the minimum score for modified and unmodified peptides was set to 60. The search was semi-tryptic, with a fixed modification of propionamide (C). Variable modifications included acetyl (protein N-term), deamidation of asparagine and glutamine, methionine oxidation, propionamide N-term, lysine propionamide, Gln and Glu to pyro-Glu, and hydroxylation of proline. All other parameters were set to MaxQuant's defaults, which included an error tolerance of 4.5 ppm. MaxQuant's output, in particular the evidence.txt and peptides.txt files, were used as input for deamiDATE.

3.2.2.3 Data analysis - Ancient collagen model dataset

Raw files were obtained via ProteomeXchange from Welker (2015). MaxQuant 1.6.2.6a was used to search against the relevant collagen type 1 sequences for each organism, as found in UniProt. The search was semi-specific with enzymes relevant to the sample preparation , with a fixed modification of carbamidomethyl (C). In the original study, all samples apart from the aardvark were prepared using both trypsin and elastase in order to increase sequence coverage and improve protein identifications - an approach that was also used to retrieve the oldest peptide sequence ever recovered (Demarchi *et al*., 2016). This multi-protease approach was later found to increase the size of the proteome retrieved from ancient bone (Lanigan *et al*., 2020). Variable modifications were set for acetyl (protein N-term), deamidation of asparagine and glutamine, methionine oxidation, Gln and Glu to pyro-Glu, and hydroxylation of proline. As this search's purpose was to identify collagen type 1, a search for contaminants was not performed. The minimum score for modified and unmodified peptides was set to 60. All other parameters were set to MaxQuant's defaults, including an error tolerance of 4.5 ppm. As with the dental calculus data, the output of MaxQuant was then used as input for the algorithm.

3.2.2.4 deamiDATE 1.0

The deamiDATE program is written in Python 2.7 and is tested on Windows, MacOs, and ubuntu. For an example evidence file of 2503 lines, and a peptide file of 962 lines, deamiDATE took 0.597 seconds to run to completion (on a computer with an Intel i7-6700 processor running at 3.40 GHz with 32GB of RAM, running Ubuntu 16.04.5). The source code and running instructions are available on GitHub https://github.com/abigailramsoe/deamiDATE. As deamiDATE is available as a stand-alone program, it allows researchers with very little computational knowledge to use it. This ensures that all have an equal chance to apply such tools to their datasets, and improves the efficacy of interdisciplinary research. The inputs required are output files from MaxQuant (a free, open-source software), and the program contains spreadsheets with the half-times of deamidation reactions from Robinson and Robinson (2004).

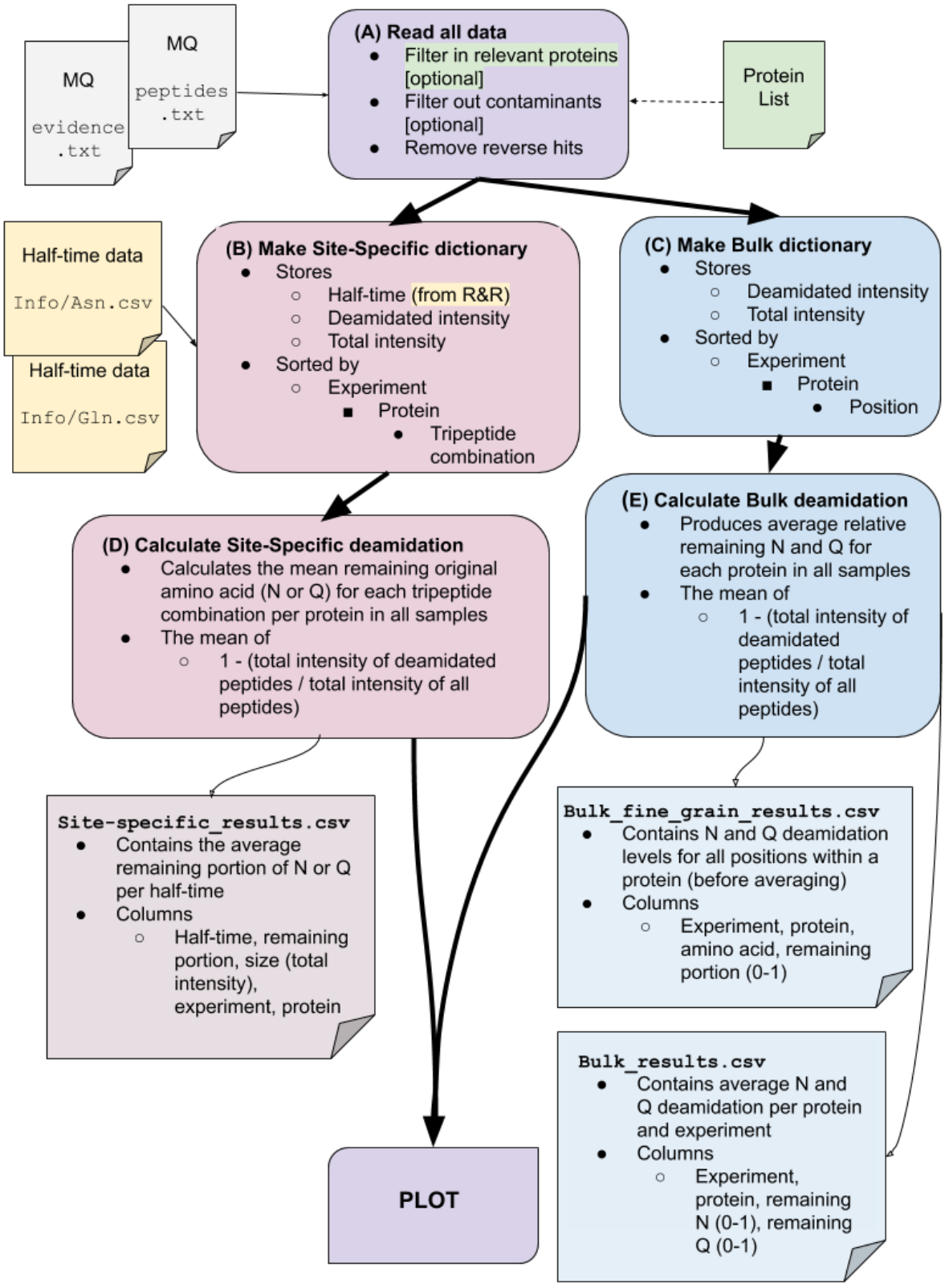


Figure 3.2*: A schematic overview demonstrating the deamiDATE 1.0 pipeline. In Step B, R&R refers to the half-time data reported in Robinson and Robinson (2004). This was downloaded into spreadsheet form and saved as Asn.csv and Gln.csv*

3.2.2.4.1 Data input

As shown in Figure 3.2, step A, deamiDATE reads the MaxQuant output files evidence.txt and peptides.txt. Here, it also optionally filters out contaminants (as defined by MaxQuant's contamination database) and reverse hits. If a protein list which contains unique protein IDs is provided, it is at this stage that the results are filtered to use these IDs. This step results in a multi-level dictionary structure, with data separated by experiment and protein.

3.2.2.4.2 Bulk deamidation

Bulk deamidation is calculated per sample, and per protein in each sample, using the same approach as presented in Mackie (2018). Each asparagine and glutamine position (for example, 103 Asn) is assigned a relative remaining intensity based on the intensity of the peptides it appears in in its deamidated form, divided by the total intensity of all peptides it appears in (both deamidated and not deamidated), as shown as steps C and E in Figure 3.2. The bulk deamidation results are then plotted, and the results are printed into new .csv files.

3.2.2.4.3 Site-specific deamidation

For each sample and protein, a new dictionary is created that stores each occurance of an asparagine (N) or glutamine (Q), and its surrounding residues. For example, in the peptide GLPNGTQAGLP, the two resulting dictionary keys are "PNG" and "TQA". This is depicted in Figure 3.2, Step B.

For each key, two variables are stored:

1) The sequence-dependent deamidation rate ("half-time") of the Q or N residue, dependent on its neighbouring residues, as estimated by Robinson *et al.* (2004).
2) A list containing the relative intensity of the deamidating residue. However, unlike bulk deamidation, where the deamidating residue is grouped with residues that appear in the same protein at the same position, for site-specific deamidation, deamidating residues are grouped with others that share the same neighboring amino acids.

Deamidation is then calculated (Figure 3.2, Step D) in the same way as in the bulk deamidation method, using the summed and averaged intensities. Like the bulk method, deamidation results are then plotted and saved to .csv format.

This leads to a large dataset where each unique X-N-Y and X-Q-Y combination has a wealth of data on its relative deamidation, which, when combined with their half-time, can be used to provide a more in-depth overview of the extent of protein deamidation in a particular molecule, relative to its age.

3.2.2.5 Statistics

Statistical analyses were performed using the 'scipy' package for Python (Jones *et al.*, 2001). Comparisons of bulk deamidation levels across datasets for both asparagine and glutamine were made using the Mann Whitney U test. For site-specific deamidation, the Wilcoxon signed rank test was used, as values for deamidation levels are paired by half-time. As there were so few half-times shared by both groups in the SK1 deamidation results (n=13), a tolerance was added to allow comparison between similar half-times (within 5%) (Supplementary Folder 3.1).

3.3 Results

3.3.1 Model dataset

3.3.1.1 Bulk measures of deamidation

We assessed the total extent of deamidation (i.e., bulk deamidation) in the model dataset, which has samples of a known age. Figure 3.3 displays the relative amount of deamidation in each of the two sample groups - modern and ancient - as presented previously (Mackie *et al.*, 2018; Mays *et al.*, 2018), calculated as above, with no relation to the half-time of the deamidation positions. The modern samples displayed relatively undamaged proteins. On average 60% of asparagine and 90% of glutamine residues in the modern samples (the aardvark, anteater, hippo and tapir collagen type 1) were undamaged. In contrast, the extent of deamidation in the ancient samples (*Toxodon*, *Macrauchenia* and *Mylodon* collagen type 1) were more advanced, with the *Macrauchenia* samples showing extremely high levels of deamidation. They display an average of only 20% undamaged asparagines and 47% undamaged glutamines. Mann Whitney U test confirmed that both asparagine and glutamine were significantly less likely to be undamaged in the ancient collagen type 1, compared to modern collagen type 1 (p=1.4065e-15, U=3088.5 and p=1.6206e-26, U=4498.0 for asparagine and glutamine, respectively).

Although the differences in extent of deamidation between modern and ancient collagens are statistically significant, there is considerable variability. Most notably, the *Mylodon* sample seems to fall within the range of the modern samples. This could be caused by the exceptional preservation of the *Mylodon* fossil in the dry, protected Cueva del Milodón; or by contamination with modern collagen. Deamidation alone cannot differentiate between these two possibilities.

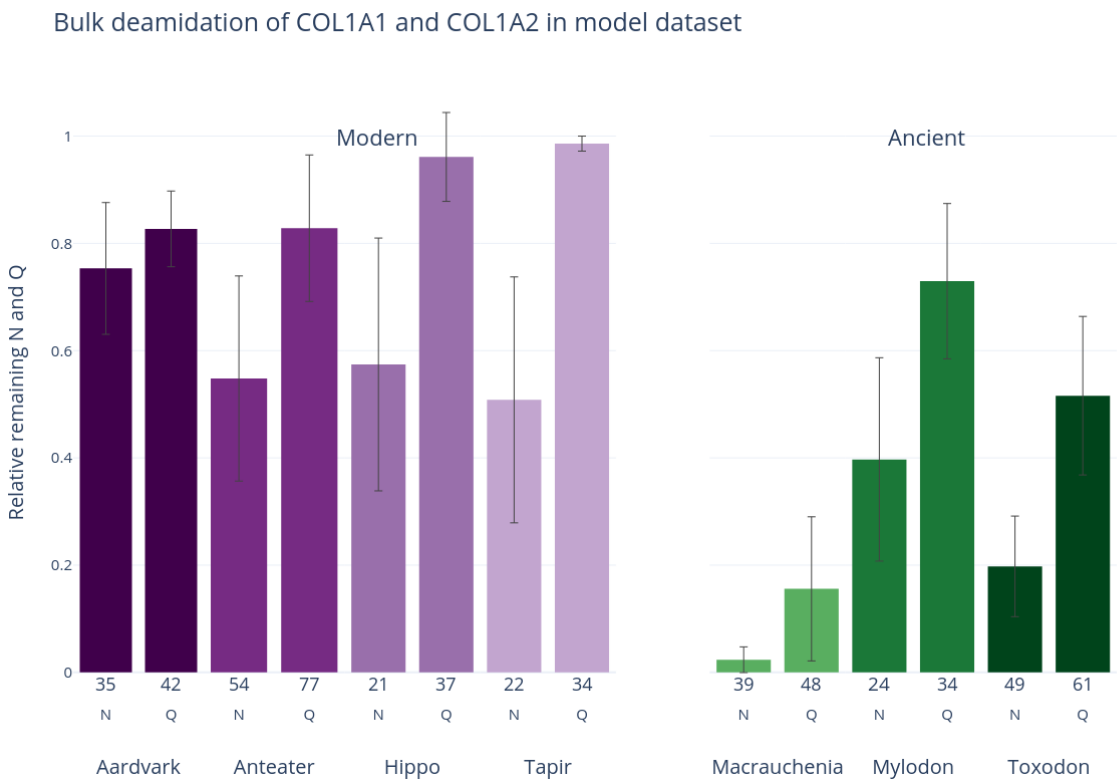


Figure 3.3: *The relative bulk level of non-deamidated N and Q per sample in the model dataset. The y-axis shows the relative "un-deamidated"portion(e.g. 1 would mean no deamidation, 0 is complete deamidation). The number of N or Q residues identified in collagens in samples is shown below each bar. The error bars represent standard deviation.*

3.3.1.2 Site-specific deamidation

When site-specific deamidation is applied to the known age dataset, as shown in Figure 3.4, it is clear that the modern samples only show high levels of deamidation at very low half-times, and sites with high half-time (>5000 days) remain mostly non-deamidated, though the variation is high. However, most of the ancient samples show high levels of deamidation, even at the highest half-time sites - the *Macrauchenia* has almost total deamidation at these sites, and the *Toxodon* sample shows similar patterns, with very high levels of deamidation. However, *Mylodon* results follow a different trend - there is far less deamidation, and these levels intermingle with those of the modern samples.

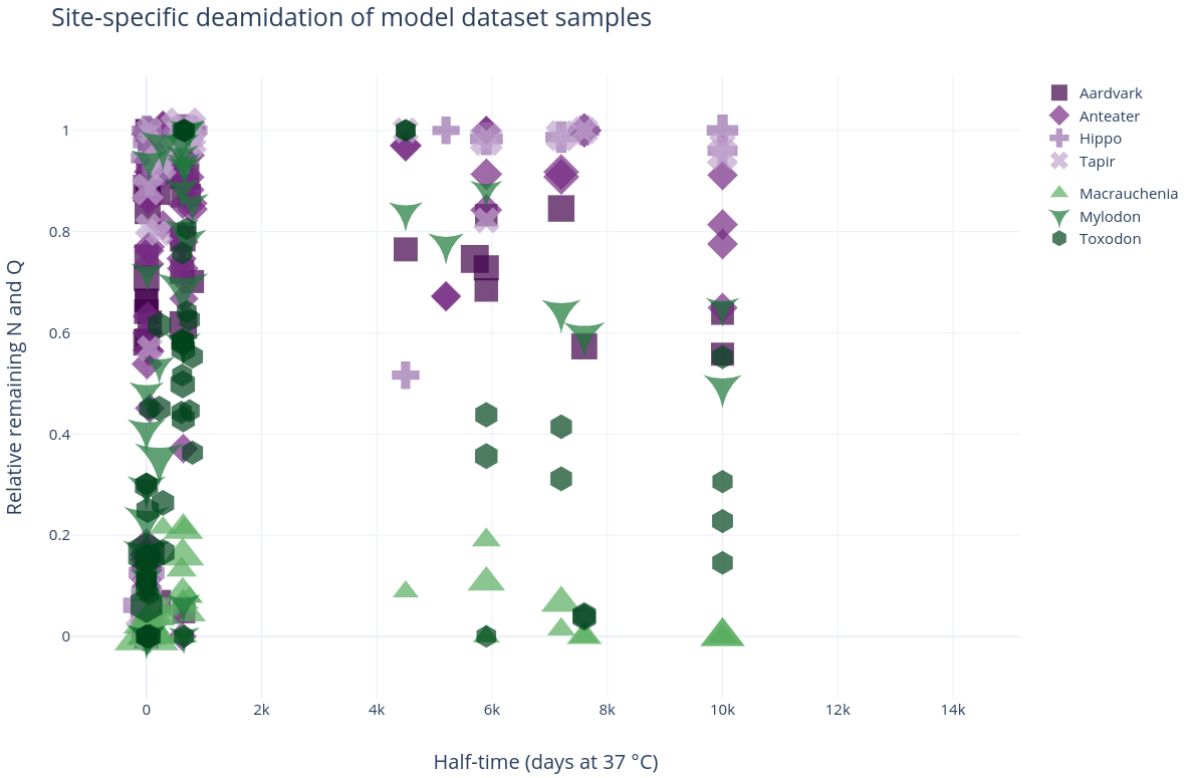


Figure 3.4: *Site-specific deamidation applied to the model dataset. Each point represents a specific three residue combination containing a deamidating site in the middle. The half-time of each combination is shown on the x-axis, and the relative proportion of non-deamidated residues is shown on the y-axis, with 1 representing no deamidation. The size of the point represents that site's intensity, relative to the total intensity of all peptides in the same sample.*

The Wilcoxon signed rank test on this data was performed by selecting deamidating sites (both asparagine and glutamine) from ancient and modern samples with the same half-time (n=184). These groups were shown to be significantly different (Z=761.0 ,p=1.6465e-23)

3.3.2 Whitehawk Camp calculus

3.3.2.1 Protein identification

A range of human- and chicken-originating proteins were identified from the dental calculus sampled from SK1 (Whitehawk Camp), as shown in Table 3.2. Seven egg-origin proteins were identified, with a total of 31 peptides and an average of 9.71% protein coverage. As many of these peptides were unique to chicken (Appendix A, Figure 1), and none were unique to any other species, it is highly likely that the egg proteins originate from chicken. Twelve total human proteins were identified, with an average of 13.88% sequence coverage over 62 total peptides.

Protein	UniProt ID	Sequence coverage	Peptide count
Egg			
Lysozyme C (Chicken)	P00698	25.20%	3
Ovalbumin	P01012	16.80%	5
Ovalbumin-related protein Y	P01014	3.10%	1
Ovotransferrin	P02789	9.20%	5
Vitellogenin-1	P87498	3.20%	4
Vitellogenin-2	P02845	8.20%	12
Vitellogenin-3	Q91025	2.30%	1
		Average 9.71%	**Total** 31
Human			
Alpha-1-antitrypsin	P01009	8.10%	3
Antithrombin-III	P01008	7.10%	3
Cathepsin G	P08311	25.10%	8
Lysozyme C (Human)	P61626	4.00%	8
Immunoglobulin kappa constant	P01834	15.00%	1
Lactotransferrin	P02788	16.20%	9
Myeloperoxidase	P05164	24.80%	17
Neutrophil defensin 1	P59665	22.30%	7
Neutrophil elastase	P08246	11.20%	3
Protein S100-A8	P05109	11.80%	1
Protein S100-A9	P06702	10.50%	1
Protein S100-A9	P06702	10.50%	1
		Average 13.88%	**Total** 62

Table 3.2: *All human and egg-origin proteins identified in SK1.*

3.3.2.2 Bulk deamidation

The same bulk method for quantifying deamidation was applied to the proteins identified in SK1 from Whitehawk Camp, as shown in Figure 3.5. The egg proteins display very low deamidation, with an average of just 8% N and 10% Q deamidation. The human proteins show a more advanced degree of deamidation, with 22% of N, and 25% of Q displaying deamidation. A Mann Whitney U test was performed on the deamidation of both asparagine and glutamine in egg and human origin proteins. Both were found to have significantly different levels of deamidation across protein groups ($p=0.0174$, $U=761.0$ and $p=0.0011$, $U=142$ for asparagine and glutamine, respectively).

It is not clear why both the chicken and human proteins display a higher level of glutamine deamidation than asparagine, as, as discussed above, asparagine deamidation takes place at a faster rate. As the variability is high (as shown by the error bars), this could be due to the presence of outliers. This also illustrates one of the key weaknesses of bulk deamidation calculations as a strategy.

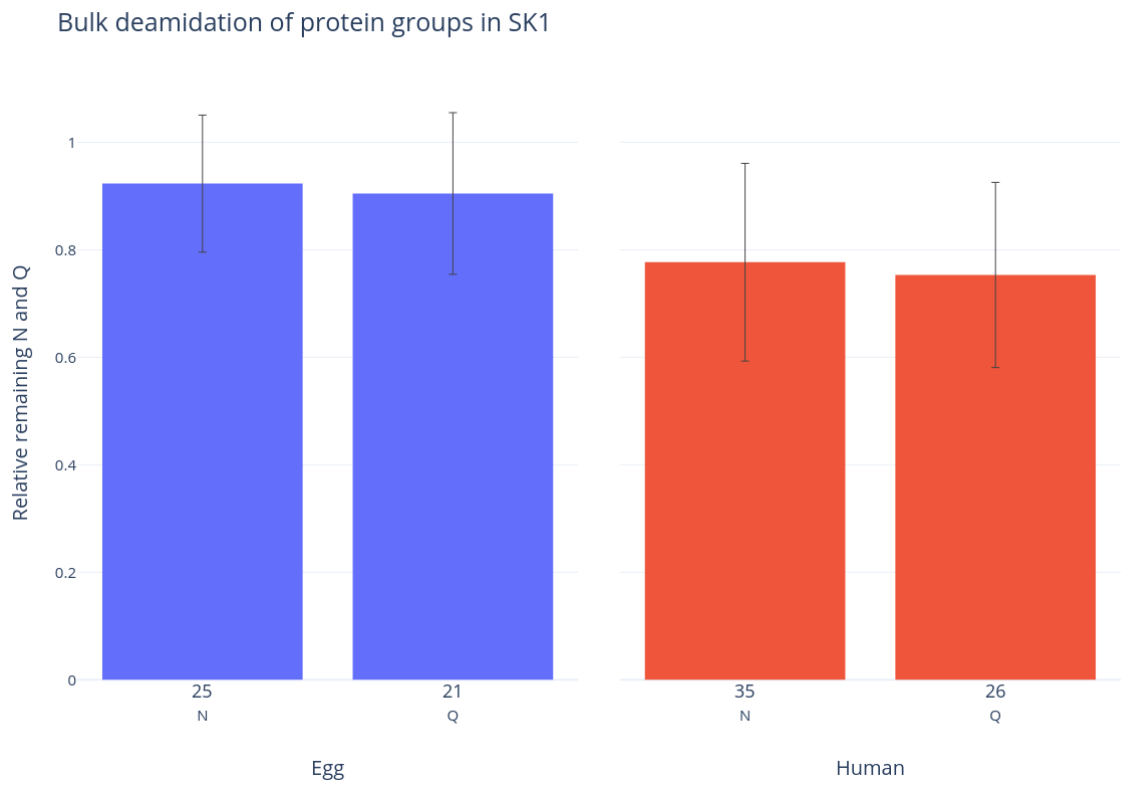


Figure 3.5: *The relative level of non-deamidated N and Q per protein group in SK1. The y-axis shows the relative "un-deamidated"portion, (e.g. 1 would mean no deamidation, 0 is complete deamidation). The number of N and Q residues in each protein group is shown below the bars.The error bars represent standard deviation.*

3.3.2.3 Site-specific deamidation

Figure 3.6 displays the half-time method of visualising deamidation, as applied to the two protein groups from SK1. The human endogenous proteins display varying levels of deamidation, with some partial and total deamidation at high half-time sites. Lastly, the chicken egg proteins for the most part only display deamidation at very low half-time sites (such as the blue point at ~0.18 on the y-axis). There are two exceptions to this - tripeptide combinations at 5200 and 6200 days are both totally deamidated. However, the fact that these are the only sites slower than 61.5 days that display any deamidation suggests that the egg proteins are of relatively recent origin, rather than representing ancient proteins.

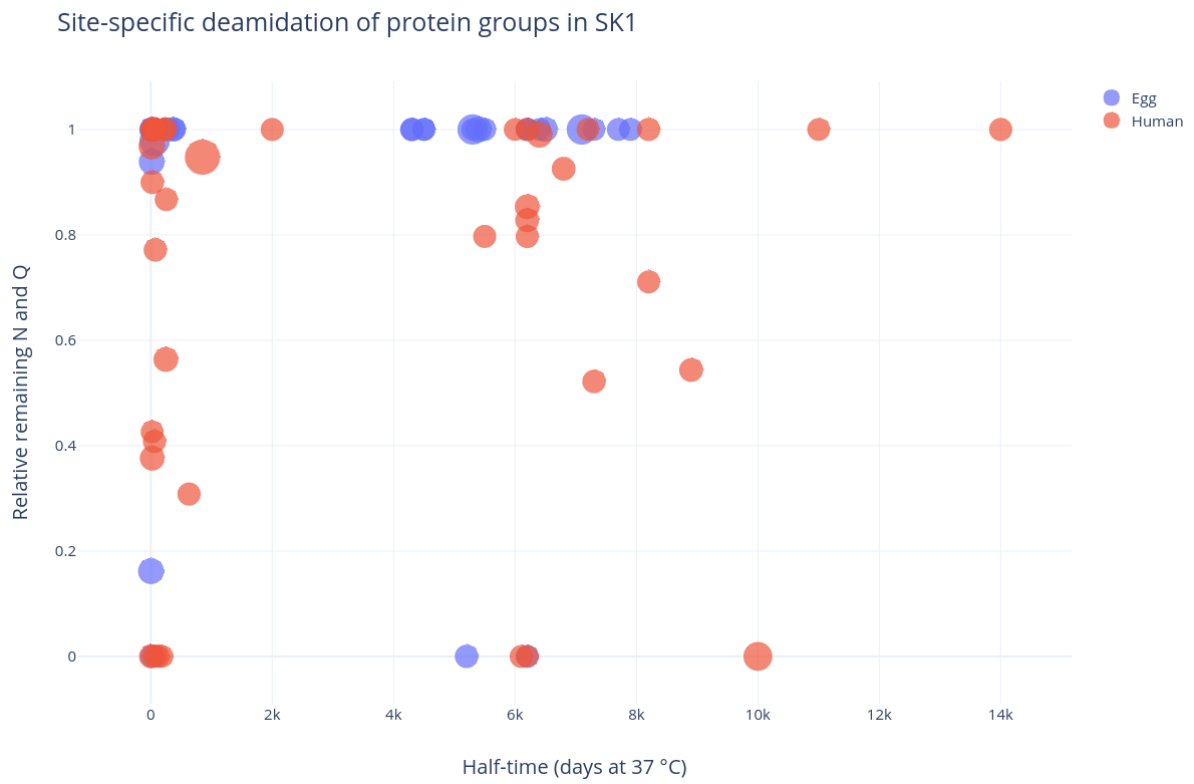


Figure 3.6: *Site-specific deamidation of the two protein groups from SK1. Each point represents a specific three residue combination containing a deamidating site in the middle. Half-time is shown on the x-axis, and the relative proportion of non-deamidated residues is shown on the y-axis, with 1 representing no deamidation. The size of the point represents that site's intensity, relative to the total intensity of all peptides in the same sample.*

A Wilcoxon signed rank test was performed by selecting deamidating sites (both asparagine and glutamine) from each of the protein groups. In order to allow greater comparison of paired values, a tolerance of 5% was added so that sites with similar half-time can be treated as pairs. This analysis showed significant differences between human and egg origin proteins (Z=724.0, p=7.8119e-05, n=111).

3.4. Discussion

Bulk deamidation has been recommended as a method of authentication for ancient proteins (Hendy *et al.*, 2018c), and indeed, can generally differentiate between the modern and ancient bone collagen samples as depicted in Figure 3.3. However, when applied to the proteins identified in the dental calculus of SK1 (Figure 3.5), these bulk measures cannot conclusively differentiate between true endogenous proteins and potential contaminants. This challenge is overcome by site-specific measures of deamidation. Figure 3.4 displays the site-specific deamidation method as applied to the model datasets. It is shown that the modern samples only exhibit full deamidation at the very low half-time sites, and although higher half-time sites are present in the sample, they are not completely deamidated. This correlates with Figure 3.3, which shows some glutamine deamidation, and around 25% asparagine deamidation. The ancient samples, *Macrauchenia* and *Toxodon*, however, display high levels of deamidation even in sites with high half-time, implying that they are both truly ancient samples. The *Mylodon* (from Patagonia) has more intermediate values reflecting its lower thermal age. Such conclusions are, however, only possible with additional information and analyses on the samples, and it is therefore impossible to reach this conclusion using deamidation alone.

In regards to chicken being present in Neolithic Britain, Figure 3.6 shows that the deamidation present in the chicken egg proteins is almost exclusively contained to low half-time sites, implying the deamidation was induced by protein extraction procedures, or occurred relatively recently. In contrast, the endogenous human proteins from the sample show higher levels of deamidation, even at high half-time sites, showing that the deamidation would have happened over a longer timescale, and the human proteins are more likely to be endogenous to the SK1 dental calculus. Considering that proteins entrapped in dental calculus have been proposed as a method for tracking ancient diet (Hendy *et al.*, 2018a), there is a need for a robust method by which to differentiate between authentic ancient dietary inclusions and potential contaminants resulting from handling and curation of skeletal material as well as laboratory contaminants. For example, bovine caseins are routinely used in western blot analysis and egg-derived proteins such as lysozyme and ovalbumin are often included in cell lysis buffers or used as molecular weight markers; both are listed as common laboratory contaminants in modern studies. In the case of the Whitehawk sample, the absence of ovalbumin and caseins in the blank extracts, alongside the observation of glue and consolidant materials on the teeth (see Figure 3.1) suggest that the modern proteins are not derived from laboratory reagents, but from conservational treatment of

the skull following excavation in the 1930s. Animal based glues have been used by curators throughout the last two hundred years in order to preserve the integrity of hard tissue morphologies. The detection of non-authentic biomolecules derived from glues has been recognised previously in ancient DNA investigations (Nicholson *et al.*, 2002) and is known to be used by conservators (Lopez-Polin, 2012).

In this study, site-specific deamidation could identify the chicken proteins within the archaeological dental calculus sample as originating from modern contamination, rather than from ancient dietary sources. As such, deamiDATE 1.0 provides a useful tool for any palaeoproteomic analysis that reports ancient egg proteins, or collagen (most commonly type 3 for hide glue, type 1 for bone glue) as they are both common in historic curatorial glues (Dallongeville *et al.*, 2011; Dooley *et al.*, 2013; Bleicher *et al.*, 2015). DeamiDATE 1.0 also presents a method by which researchers reporting ancient dietary proteins can differentiate between ancient sources of casein, bovine serum albumin, lysozyme, and ovalbumin and those derived from common laboratory reagents (Mackie *et al.*, 2017; Hendy *et al.*, 2018a; Mays *et al.*, 2018; Charlton *et al.*, 2019). This study confirmed that the extent of deamidation is linked to the age of the sample for collagen derived from animal bone, and for human proteins derived from dental calculus - as generally only very ancient samples display advanced deamidation in high half-time sites. Nevertheless, more systematic analysis is required to quantify the extent of deamidation within other proteomes, and preserved within other substrates. For example, although site-specific deamidation has not yet been systematically applied to proteins preserved on ancient ceramics, a recent study of Endmesolithic ceramics from Germany suggests that examining the extent of glutamine deamidation may be useful for distinguishing between endogenous and contaminant proteins on ceramic vessels (Shevchenko *et al.*, 2018). Systematic study of other archaeological substrates, including mummified tissues (Corthals *et al.*, 2012; Hendy *et al.*, 2016), keratinous products such as baleen, horn or feather (Solazzo *et al.*, 2013), and preserved food remains (Shevchenko *et al.*, 2014; Colonese *et al.*, 2017; Hendy *et al.*, 2018b), are required to assess the extent to which deamidation may be used to authenticate ancient proteins preserved across a range of materials and depositional contexts.

Furthermore, this approach relies on an abundance of peptide identifications - more high quality spectral matches results in a higher confidence of authenticity. However, in samples with only a few peptides authenticating a given protein, many of which may not even contain an asparagine or glutamine residue, levels of deamidation are unlikely to be a powerful measure of authenticity.

3.5. Conclusion

In this study, we developed a novel software tool, deamiDATE 1.0, to authenticate ancient proteins using a measure of site-specific deamidation. We demonstrate the effectiveness of this approach by assessing time-dependent damage in bone collagen from modern, archaeological and extinct taxa. We further apply this method to identify modern contaminant chicken proteins in Neolithic dental calculus. Currently, deamiDATE 1.0 is designed to quantify a single type of post-translational modification: deamidation. Deamidation rates will be influenced not only by the age of the sample, but by protein structure (primary, secondary, tertiary & quaternary structures), by environmental factors (pH, hydration, presence of catalysts, etc.), as well as by pre-depositional human activities (i.e. cooking, excarnation, mummification). In future studies, it will be useful to attempt to assess the deamidation rates directly from archaeological and fossil samples. For example, in the case of proteins derived from animal bones, we may hypothesize that cooking, boiling or roasting may accelerate deamidation rates. Likewise, cooking may also potentially increase deamidation rates for food proteins adsorbed on to pottery residues, or entrapped within dental calculus.

Systematic analysis of proteins recovered from a range of archaeological substrates, geographic areas, archaeological contexts, and temporal periods will be required to elucidate the relationship between protein origin, structure, human activities, depositional environments and deamidation rates for ancient proteins. In assessing such samples, other PTMs such as oxidation may also prove useful in some samples, to assess alternative patterns of protein diagenesis, such as deterioration of objects in museums and galleries.

Chapter 4: Milking it? Assessing the degradation of ancient milk proteomes

Abigail Ramsøe[1,2*], Mia Crispin[1], Meaghan Mackie[3,4], Krista McGrath[1,5], Roman Fischer[6], Beatrice Demarchi[7], Matthew Collins[3,8], Jessica Hendy[1], Camilla Speller[1,9,]

Corresponding authors: Abigail Ramsøe, Camilla Speller

Affiliations:

1. BioArCh, Department of Archaeology, University of York, York, UK
2. Department of Earth Sciences, Natural History Museum, London, UK
3. The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
4. The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark
5. Department of Prehistory and Institute of Environmental Science and Technology (ICTA), Universitat Autònoma de Barcelona, Bellaterra, Spain
6. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, UK
7. Department of Life Sciences and Systems Biology, University of Turin, Italy
8. McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK
9. Department of Anthropology, University of British Columbia, Vancouver, Canada

Relevance: This chapter comprises a comparative analysis of three experimental milk datasets and 274 previously published samples containing archaeological milk. Its focus is whether milk proteomes can be reliably authenticated using deamidation alone.

Ramsøe, A., Crispin, M., Mackie, M. *et al*. Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. Sci Rep 11, 7795 (2021). https://doi.org/10.1038/s41598-021-87125-x

Abstract: The origins, prevalence and nature of dairying have been long debated by archaeologists. Within the last decade, new advances in high-resolution mass spectrometry have allowed for the direct detection of milk proteins from archaeological remains, including ceramics, dental calculus, and preserved dairy products. However, proteins recovered from archaeological

remains are susceptible to pre- and post-excavation and laboratory contamination, a particular concern for ancient dairying studies as milk proteins are potential laboratory contaminants. Here, we examine how site-specific rates of deamidation can be used to elucidate patterns of peptide degradation, and authenticate ancient milk proteins. First, we characterize site-specific deamidation patterns in modern milk products and experimental samples, confirming that deamidation occurs primarily at low half-time sites. We then compare this to previously published ancient proteomic data from six studies reporting ancient milk peptides. We confirm that the extent of deamidation , on average, is more advanced in beta-lactoglobulin recovered from ancient dental calculus and pottery residues. Nevertheless, deamidation displayed a high degree of variability, making it challenging to authenticate samples with relatively few milk peptides. We demonstrate that site-specific deamidation is a useful tool for identifying modern contamination but highlight the need for multiple lines of evidence to authenticate ancient protein data.

4.1 Introduction

The adoption of animal milk into the human diet has long intrigued archaeologists as it represents a vital source of nutrition for past and contemporary populations (Craig *et al*., 2000; Dunne *et al*., 2012). As well as for food, milk is also utilized as a medium in cultural heritage objects, for example as a component of mortars and paint binders (Vinciguerra *et al*., 2019). It is unsurprising, therefore, that bioarchaeologists have sought out new molecular tools for identifying trace residues of dairy products preserved on or within material culture or skeletal remains. Whilst much of the previous work on the biomolecular detection of dairying has focused on the identification of milk fats (Craig *et al.*, 2005; Evershed *et al.*, 2008; Dunne *et al.*, 2012, 2019; Cramp *et al.*, 2014; Spiteri *et al.*, 2016), the detection of ancient milk proteins is a growing approach (Warinner *et al.*, 2014; Hendy *et al.*, 2018; Jeong *et al.*, 2018; Charlton *et al.*, 2019; Geber *et al.*, 2019). Although the survival of food proteins in archaeological artefacts and human remains is less well-studied than their lipid counterparts, proteins can often be more taxonomically specific, enabling the identification of specific taxa utilized and consumed— a 'zooarchaeology by proxy' approach. Ancient milk proteins have been detected from a suite of archaeological artefacts, human remains, and objects of cultural heritage. These include well-preserved whole cheese-like remains from Early Bronze Age China (Yang *et al.*, 2014), as well as from residues adhering to archaeological artefacts (Buckley and Melton, 2013; Xie *et al.*, 2016; Hendy *et al.*, 2018). Milk proteins have also been extracted from ancient dental calculus in Europe (Warinner *et al.*, 2014; Hendy *et al.*, 2018; Charlton *et al.*, 2019; Geber *et al.*, 2019) and Central Asia (Jeong *et al.*, 2018), providing direct evidence of dairy consumption as opposed to the evidence of dairy processing which is obtained from ceramic residues (Hendy *et al.*, 2018) or well-preserved remains (Yang *et al.*, 2014). Recently, dog-specific milk proteins were detected in the bones of a neonate dog, indicating the power of this methodology for detecting milk consumption from skeletal remains alone (Tsutaya *et al.*, 2019). Milk proteins have also been detected in ancient mortars, binders and paints (Lluveras-Tenorio *et al.*, 2017) providing valuable insight into past manufacturing and informing on contemporary conservation practices.

Archaeological artefacts are susceptible to contamination from modern (and ancient) biological materials and laboratory reagents (Barton, 2007; Wall and Kim, 2007). The challenge of contamination has been explored thoroughly in ancient DNA studies, where contamination from modern sources of DNA or PCR-product carry-over has been well documented (Gilbert *et al.*, 2005). Contamination with modern protein sequencers can occur at any point from burial, excavation and curation through to extraction and analysis by mass spectrometry (Hendy *et al.*,

2018c). For example, post-excavation handling or conservation treatments with plasters or glues that can contain milk or laboratory contamination by milk powders (e.g., blotting agents) can easily introduce a modern milk proteome to ancient artefacts or skeletal remains. Additionally, genuinely archaeological milk proteins are often detected at low abundance, which makes it difficult to differentiate them from sporadic post-excavation contamination.

As proteins degrade over time, damage identified in recovered proteins has been used to argue for or against their authenticity (Buckley *et al.*, 2008). Recently, the software tool deamiDATE (Ramsøe *et al.*, 2020) has been developed in order to assess patterns of asparagine and glutamine deamidation in proteins in relation to the expected rate of the deamidation reaction (Robinson *et al.*, 2004). The rates calculated by Robinson and Robinson (Robinson and Robinson, 2001) and used by deamiDATE refer to short artificial peptides under controlled pH and temperature conditions (37°C, pH 7.4). We note that these conditions are not those typically experienced by archaeological specimens, but they do provide a baseline for calculations. DeamiDATE is based on the idea that high levels of deamidation in peptides that are expected to take a long time to deamidate under these standard conditions should only be seen in genuinely old proteins; whereas deamidation events that have a theoretical fast rate are less reliable as a measure of authentication. Here, we use deamiDATE to investigate to what extent the amount of asparagine and glutamine deamidation can be used as a measure of authenticity for ancient milk proteins. First, we apply deamiDATE to characterise the extent of deamidation in two modern experimental datasets: laboratory milk powder and archaeological artefacts that have been intentionally exposed to potential laboratory contamination, in order to identify deamidation signatures of potential contaminating sources. Next, we compare these values to those displayed by ancient proteomic data, specifically 274 samples from previously published studies reporting the detection of milk proteins (casein, beta lactoglobulin) in archaeological dental calculus (Warinner *et al.*, 2014; Hendy *et al.*, 2018a; Jeong *et al.*, 2018; Mays *et al.*, 2018; Charlton *et al.*, 2019) and ceramic artefacts (Hendy *et al.*, 2018c). This reveals that deamidation displayed a high degree of variability, making it challenging to authenticate samples with relatively few milk peptides, although in general site-specific deamidation is more advanced in beta lactoglobulin recovered from ancient dental calculus and pottery residues.

4.2 Materials and methods

4.2.1 Experimental datasets

Two experimental milk datasets were used to characterize deamidation in potential contamination sources. Firstly, proteins from skim milk powder were extracted and analysed with the aim of understanding the deamidation patterns in milk-based laboratory reagents. We then undertook an experiment to intentionally expose archaeological samples to potential laboratory contamination to examine if patterns of deamidation differed meaningfully from putative endogenous ancient milk proteins.

4.2.1.1 Skim Milk Powder

We extracted and analyzed proteins from Sigma Aldrich skim milk powder [Cat. no. 70166] to: a) investigate the damage caused by the evaporation processes and b) to quantify deamidation patterns of milk powder that is used as a lab reagent, in order to improve our chances of detecting it as a contaminant in archaeological samples.

The protein extraction was based on the protocol published by Jersie-Christensen *et al. (2018)* for ancient dental calculus analysis, without the demineralisation step. Approximately 500 µg of skim milk powder was suspended in guanidine buffer (2 M guanidine hydrochloride solution, 20 mM chloroacetamide, 10 mM tris (2-carboxyethyl)phosphine) with 100 mM Tris) and ammonium hydroxide was used to adjust the pH to 7.5–8.5. The sample was heated at 10 min at 99°C to denature the proteins, and then subsequently cooled for 10 minutes. The sample was then digested at 37°C for 1 h with 0.2 µg of rLysC (Promega) under agitation. The sample was then diluted to a final concentration of 0.6 M guanidine hydrochloride using 25 mM Tris in 10% acetonitrile (ACN). The sample was then digested overnight with 0.8 µg of trypsin (Promega). To inactivate the trypsin 10% trifluoroacetic acid (TFA) was added until the pH was <2. The peptides were washed and collected on in-house made C18 StageTips and stored in the freezer until mass spectrometry analysis. Samples were eluted with 20 µL of 40% ACN followed by 10 µL of 60% ACN directly into a 96 well plate. Samples were evaporated in a SpeedVac Concentrator (Thermo Fisher Scientific) until ~3 µL was left and 5 µL of 0.1% TFA, 5% ACN was added.

The sample was then analysed using an EASY-nLC 1200 system connected to a Q-Exactive HF (Thermo Scientific) mass spectrometer at the Novo Nordisk Center for Protein Research, University of Copenhagen, according to previously published parameters (Mackie *et al.*, 2018).

4.2.1.2 Whey Protein Heat Treatment

In the original manuscript, we intended to use a previously published dataset that had investigated changes in whey proteins during heating (Xiong *et al*., 2020). However, after exploring the dataset and finding some very intriguing patterns, it was made clear to us that the denatured milk proteins were discarded in the original study - resulting in very few deamidated peptides. This meant that this dataset did not reflect the true variation in damage in heated milk, and therefore was unsuitable for this study. We would recommend the careful and thorough examination of third-party datasets before use in meta-analyses such as this chapter.

4.2.1.3 Contamination Experiment

We undertook an experiment to quantify potential milk contamination that can result from exposure of archaeological samples within a modern laboratory setting, though no specific methodology was undertaken in order to ensure they became contaminated. We selected five archaeological samples for shotgun proteomic analysis, including three human dental calculus samples; one ostrich (*Struthio kakesiensis*) eggshell and one limescale deposit from an historic sewer pipe (Table 4.2). These samples were selected since: 1) they represent a range of archaeological or paleontological substrates previously demonstrated to preserve ancient proteins; and 2) based on their origin and provenience, the majority are unlikely to preserve endogenous milk proteins. The dental calculus samples are from individuals buried in Rupert's Valley, St. Helena and date to the mid 19th century (Pearson *et al*., 2011). Previous proteomic analyses of dental calculus from the same archeological context (and extracted within a dedicated ancient proteins laboratory) did not produce any evidence of milk proteins (Warinner *et al.*, 2014). The ostrich eggshell was recovered from the Laetoli site in Tanzania and dates to ~3.85 to 4.3 Ma (Harrison & Msuya 2005); as above the eggshell, previously analyzed in Demarchi *et al.* (2016), displayed no evidence of milk proteins or analogues. The fifth sample represented a

limescale deposit on an early 19th century water pipe from Hungate, a 'slum' part of York (Rowntree, 1901), where milk products are unlikely to have been present in large quantities.

Sample Code	Weight (mg)	Archaeological site [identifier]	Age
Dental Calculus A	50	Rupert's Valley, St Helena. SK358 (8786)	Mid-19th century
Dental Calculus B	81.9	Rupert's Valley, St Helena. SK520 (8836)	Mid-19th century
Dental Calculus C	39.8	Rupert's Valley, St Helena. SK245 (8753)	Mid-19th century
Limescale	102.1	Hungate, York. Subsampled from S Bend drain.	Early-19th century
Egg Shell	89.8	Lot 13898, Kakesio 1-6, Lower Laetoli Beds, Tanzania	~3.85 to 4.3 Ma
Blank	N/A	Extraction blank control	

Table 4.1: *Contamination experiment sample extracted and analyzed in this study*

The samples were weighed out as whole pieces and were crushed into powder in a 2.0 ml Eppendorf tube using a sterile micropestle for each sample. 1800 µL of ethylenediaminetetraacetic acid (EDTA) 0.5 M pH 8.0 was added to each sample. The samples were then parafilmed and incubated at room temperature on a rotator for 5 days. Protein extraction was undertaken at the Discovery Proteomics Facility at the Target Discovery Institute (TDI) in Oxford, within their protein extraction laboratory. The laboratory follows standard procedures for molecular biology laboratories: protective equipment was worn (i.e., gloves, lab coats) but other contamination controls recommended for degraded samples were omitted (i.e., aerosol resistant pipette tips, or extraction hoods were not used).

Samples were vortexed for 30 seconds and then centrifuged for 5 minutes; proteins were extracted from the pellet and from 100 µL of supernatant using a gel-aided sample preparation (GASP) method (Fischer and Kessler, 2015). To each sample, 100 µL of Pierce IP Lysis Buffer (Thermo Fisher Scientific; Catalogue Number 87787) was added, before the addition of 100 µL of acrylamide (30%). The samples were vortexed for 30 seconds and left for 5-10 minutes. 2 µL of TEMED (tetramethylethylenediamine) followed by 2 µL of 10% APS (ammonium persulfate) was

added to samples, and left for 10-20 minutes until a gel had formed. Once set, the gel was shredded through a filter membrane and fixed through rotation with 10% acetic acid, 50% methanol, 40% water. The solution was centrifuged, and the supernatant discarded. 1 mL of acetonitrile was added to dehydrate the gel pieces. A series of washing and drying steps using acetonitrile were then performed to exchange buffers, following the method outlined in Hendy *et al.* (2018a). Finally, the samples were digested overnight at 37°C in 250 µL of ammonium bicarbonate (0.05M) and 1 µg trypsin. The next morning, samples were centrifuged for 1 minute. 250 µL of acetonitrile was added and the samples were placed on a shaker for 5 minutes. The supernatant was removed and retained in a new tube, and 250 µl of 5% formic acid was added to the gel pieces for 5 minutes. The supernatant was removed and added to the first fraction, and 100 µL of acetonitrile was added to the gel pieces, and shaken for 5 minutes. The supernatant was taken off and combined with the first and second fractions before being desalted using Millipore Zip-Tips prior to MS/MS analysis.

All samples were analysed using a Thermo Fisher Scientific n-LC Q-Exactive tandem mass spectrometer at the Discovery Proteomics Facility, Target Discovery Institute, Oxford according to previously published specifications (Warinner *et al.*, 2014).

4.2.2 Ancient milk

We re-analyzed six recently published datasets reporting proteomic evidence of ancient milk, including five dental calculus studies (Warinner *et al.*, 2014; Hendy *et al.*, 2018a; Jeong *et al.*, 2018; Mays *et al.*, 2018; Charlton *et al.*, 2019) and one proteomic analysis of ceramic vessels (Hendy *et al.*, 2018). Although a number of publications have recently reported the recovery of milk proteins from ancient samples (Kuckova *et al.*, 2009; Hong *et al.*, 2012; Buckley and Melton, 2013; Shevchenko *et al.*, 2014; Yang *et al.*, 2014; Rao *et al.*, 2015; Villa *et al.*, 2015; Xie *et al.*, 2016; Greco *et al.*, 2018), few studies to date have made their raw data available for re-analysis through public databases. These six datasets were selected as they span a wide timescale and geographical area, and most importantly, raw data and sample identifiers were made publically available through ProteomeXchange. Where possible, ages were assigned to samples using published radiocarbon dates. In the absence of direct dating, age was assigned to be in the middle of the reported archaeological period or date ranges (see Appendix B, Table 1).

Dataset *(Accession)*	**Substrate**	**Number of samples**	**Age (range)**	**Site/s**	**Radiocarbon dates available?**
Warinner *et al.,* 2014 *(PXD001357, PXD001359, PXD001360, PXD001361, and PXD001362)*	Dental calculus	122	Neolithic - present day 4000 - 0 YBP	UK, Denmark, Norway, Germany, Hungary, Italy, Switzerland, Armenia, Russia, St Helena, Greenland	No
Mays *et al.,* 2018 *(MSV000081706)*	Dental calculus	12 samples from 3 individuals	Neolithic - Bronze Age 4000 - 700 YBP	UK	Yes
Jeong *et al.,* 2018 *(PXD008217)*	Dental calculus	12 samples from 9 individuals	Bronze Age 2600 - 700 YBP	Mongolia	Yes
Hendy *et al.,* 2018a *(PXD009603)*	Dental calculus	112	Iron Age - present day 800 - 0 YBP	UK (USA - modern samples only)	No
Hendy *et al.,* 2018b *(PXD008647)*	Ceramic vessels	18 samples from 10 vessels (33 including replicates)	Neolithic 4000 - 2200 YBP	Çatalhöyük, Turkey	Yes - indirectly from context
Charlton *et al.,* 2019 *(PXD012893)*	Dental calculus	10	Neolithic 4000 - 2200 YBP	UK	Yes - indirectly from context

Table 4.2: *Archaeological datasets reanalysed in this study*

4.2.3 Bioinformatic analysis

4.2.3.1 MaxQuant

Proteomic data analysis followed the same bioinformatic methods for all datasets. Raw files from all datasets were searched using MaxQuant 1.6.2.6a against a database of caseins (including alpha-S1-, alpha-S2-, beta-, and kappa-casein) and beta-lactoglobulin (BLG) from horse, goat, sheep, and cow, using a semi-tryptic search strategy (Appendix B, File 1). The minimum score for modified and unmodified peptides was set to 60. For previously published archaeological datasets extracted using a filter-aided sample preparation method (FASP), the fixed modification was set to carbamidomethyl (C), and the variable modifications hydroxyproline, Glu and Gln to pyro-Glu, deamidation (NQ), acetyl (N-term) and oxidation (M). For the published datasets, and archaeological samples extracted with the GASP method, the fixed modification was propionamide (C), and the variable modifications included all those in the FASP setup, plus propionamide N-term and propionamide (K). As there is considerable overlap between the database used and MaxQuant's contaminant database, the search for contaminants was turned off. All other parameters were set to MaxQuant's defaults, including a false discovery rate (FDR) of 1%.

In order to allow for robust comparison between datasets, the experimental datasets (milk powder and the contaminated artefacts) were run using the same settings as the FASP-extracted archaeological data.

4.2.3.2 deamiDATE 1.0

We then ran deamiDATE 1.0 (Ramsøe *et al.*, 2020) on the MaxQuant output files in order to quantify the levels of bulk and site-specific deamidation. Deamidation refers to the process in which asparagine and glutamine lose an amide group, and thereby are converted into aspartic acid and glutamic acid, respectively. Due to differences in the pathways involved, glutamine deamidates at a much slower rate than asparagine (with median half-times of 6100 and 60 days respectively), and as such, it has been proposed as a method for determining relative archaeological age (Leo *et al.*, 2011; van Doorn *et al.*, 2012; Wilson, van Doorn and Collins, 2012;

Orlando *et al.*, 2013; Hill *et al.*, 2015; Welker *et al.*, 2015). In theory, the higher the extent of glutamine deamidation, the more degraded (i.e. "older") the sample. DeamiDATE calculates the deamidation of proteins within a sample, including site-specific deamidation —i.e., differences in the extent of deamidation in glutamine and asparagine based on the presence of specific neighbouring acids (Robinson *et al.*, 2004).

This site-specific approach is more nuanced than reporting the bulk levels of deamidation in a sample, as it can differentiate between rapid deamidation events (e.g., those that can occur within days of deposition, or during the protein extraction process —and therefore have low predictive power as to estimate the age or or authenticity of a sample) and high half-time deamidation events (e.g., those that occur more slowly and should therefore only be observed in genuinely ancient proteins). In order to avoid spurious protein identifications, only proteins with a minimum of two supporting peptides were included in downstream analyses, as per Hendy *et al.* (2018c).

4.3 Results

4.3.1 Modern milk

4.3.1.1 Skim milk powder

As expected, proteins extracted from skim milk powder resulted in a large number of both casein and BLG peptides (n=295 and n=72, respectively). These results are also in line with the reported ratio of casein-to-whey in raw and powdered milk of 80:20 (Lara-Villoslada, Olivares and Xaus, 2005). As milk is subjected to extreme heat in the dehydration process, and heating has a denaturing effect on protein structure, we expected to find increased levels of deamidation. In contrast to our expectations, for both BLG and caseins, the median values display very low levels of deamidation (above 95% undamaged peptides in all cases, Figure 4.1; Appendix B, Table 2). For both proteins, there is slightly more asparagine deamidation than glutamine, consistent with expectations of facile asparagine deamidation and with the difference in calculated activation energies (92-100 kJ/mol for Asn and 131 kJ/mol or 134 kJ/mol for either direct hydrolysis of cyclization-mediated glutamine deamidation (Stratton *et al.*, 2001)). There is also greater variability in asparagine deamidation in BLG compared to casein (excluding outliers).

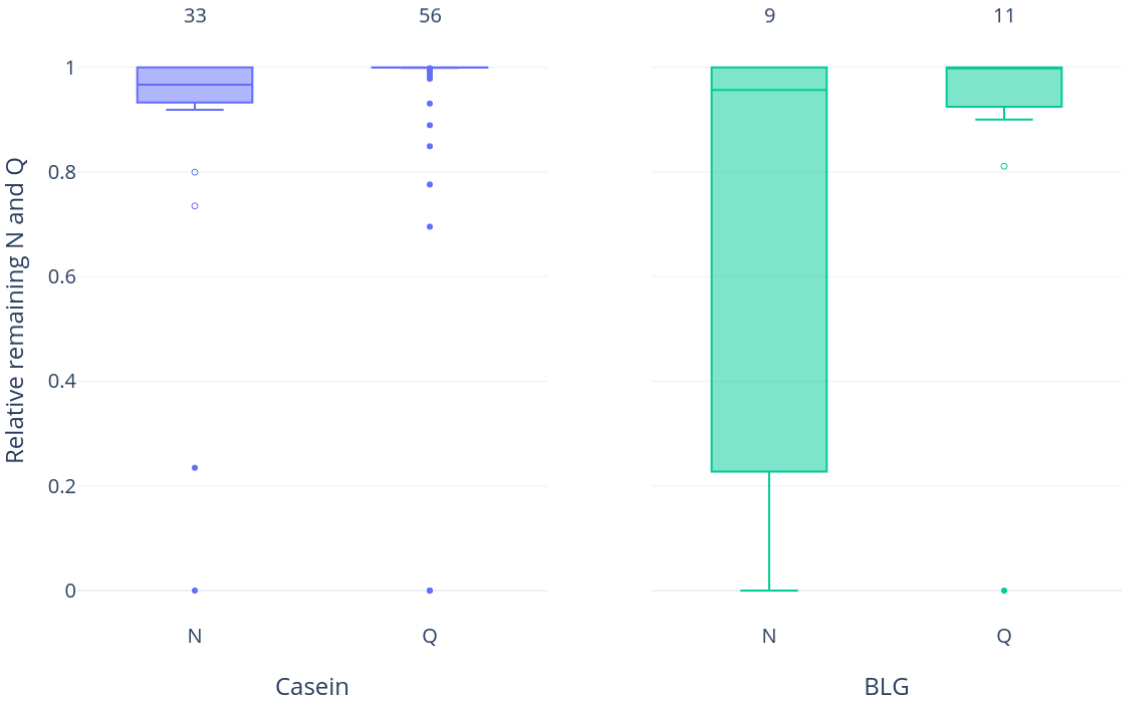


Figure 4.1: *Deamidation of casein and BLG proteins in milk powder samples. Numbers at the top of the bars show the number of asparagine or glutamine represented by this result. The y-axis represents the relative remaining amount of the deamidating amino acid - therefore high values mean less deamidation; whereas low values imply higher levels of deamidation. Outliers are shown as points, while suspected outliers are shown as hollow points.*

We also examined patterns of site-specific deamidation in the skim milk powder (Figure 4.2; Appendix B, Table 3). BLG displays only high-levels of deamidation in low half-time sites, i.e., sites which take on average a few days (in the standard conditions used for the theoretical rate calculations) to deamidate. The casein results are very similar, although fewer deamidated low half-time sites were detected, and surprisingly shows some sites with very high half-time (above 5000 days) that are deamidated. This implies that casein and BLG react differently to heat, and therefore also possibly to the diagenetic effects of time. It is important to point out that temperature alters the overall order of occurrence of the various diagenetic reactions, so much so that high-temperature laboratory experiments are notoriously unable to mimic the combined effect of time and normal burial temperatures on the diagenesis of proteins (Collins *et al.*, 2000). The rapid exposure to extreme heat is therefore likely to produce effects that are significantly different from those observed in archaeological specimens.

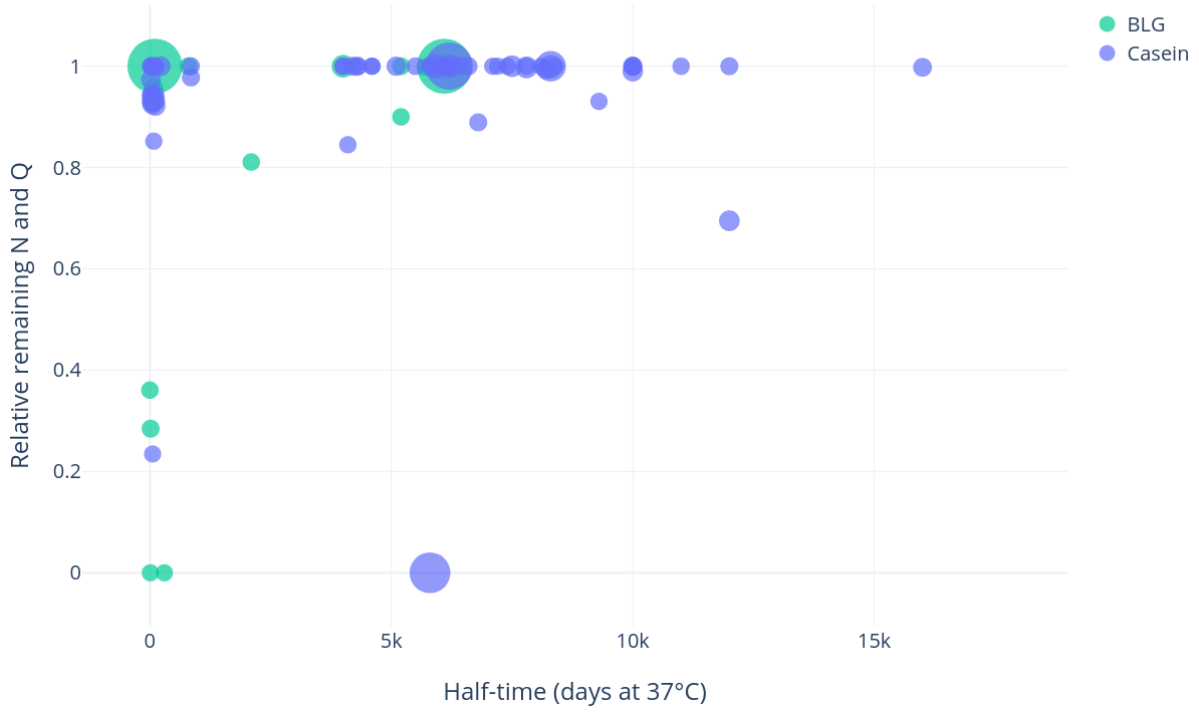


Figure 4.2: *Site-specific deamidation of milk powder proteins. The size of the points is relative to the intensity of the peptide identified. Half-times as estimated by Robinson and Robinson (2001)*

4.3.1.2 Contamination experiment

We extracted proteins from archaeological materials within a modern protein laboratory setting, first to determine whether traces of milk proteins from the laboratory environment could be detected through mass spectrometry, and second, to characterize deamidation in any contaminating milk peptides. Caseins were confidently detected (with two or more unique peptides) in all five samples in the contamination experiment, as well as the extraction blank control (five peptides). BLG was detected in two out of the five samples, as well as the extraction blank control, but only a single peptide was identified in each sample (each representing a different amino acid sequence).

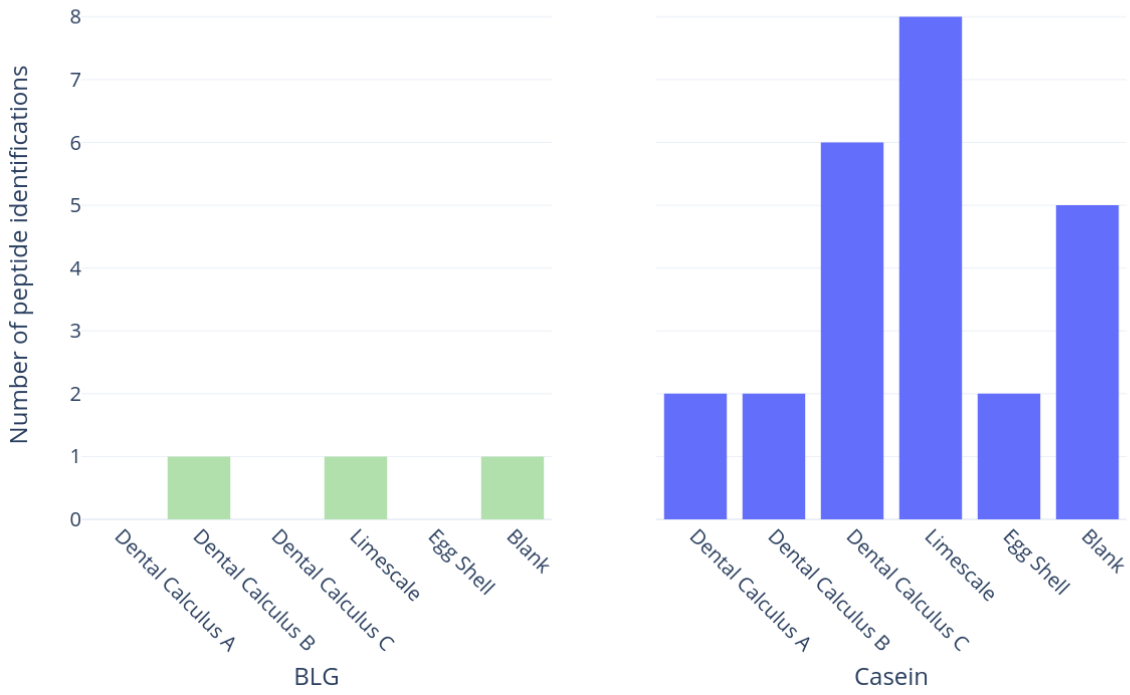


Figure 4.3: *Number of BLG and casein peptides identified in experimentally contaminated samples (extracted and analyzed in this study) according to MaxQuant's 'razor and unique peptides' output.*

Most of the experimental samples displayed no deamidation within the casein peptides, which is consistent with expectations of contamination from modern laboratory reagents, such as skim milk powder (presented above). There is a noteworthy amount of asparagine deamidation in two of the dental calculus samples, though this could have been a result of laboratory-induced deamidation (Appendix B, Table 6, see discussion below). The limescale sample exhibits the highest levels of deamidation, with an average of 20% remaining asparagine and 57% remaining glutamine, though a wide range of glutamine deamidation.

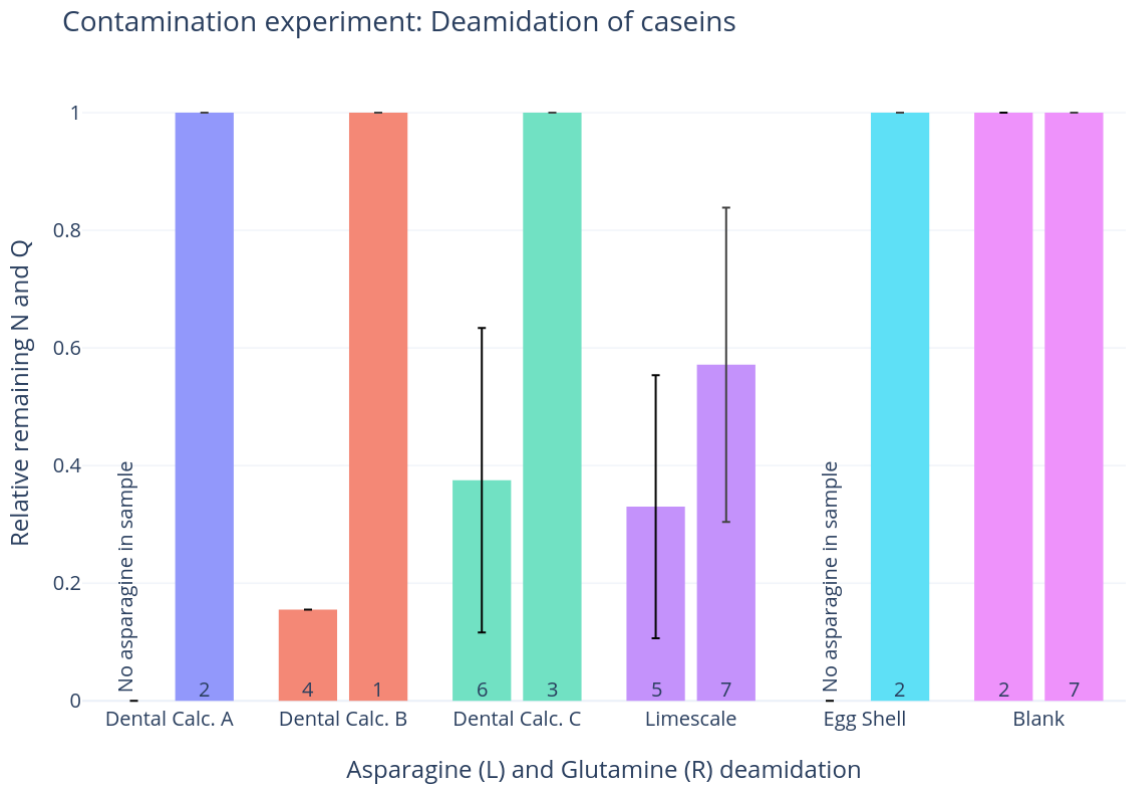


Figure 4.4: *Deamidation in contamination experiment samples. The number of residues involved in this calculation is shown in each bar. Error bars represent the standard deviation. There were no casein peptides containing asparagine in the Dental Calculus A and Egg Shell samples.*

Figure 4.5 displays the site-specific deamidation of caseins within the contamination experiment, and indicates that the only casein deamidation events in the dental calculus samples are at very low half-times, so could likely be a result of sample preparation and processing, and are unlikely to be a signature of ancient milk. The limescale results are more nuanced. It shows some high half-time deamidation, which is theoretically indicative of genuine archaeological proteins. However, these peptides with high half-time deamidation have very low intensity (shown by the size of the scatter points), and there seem to be an equal amount of high half-time deamidation events in limescale that show no deamidation, and these points have higher intensity. Lastly, the limescale casein peptides are either totally deamidated or totally intact (e.g. they are at 0 or 1) - there are no cases when the same peptide is found in both states, which would lead to a deamidation value of above 0 and below 1. Limescale deposits have been proposed to provide stable environments for protein preservation through mineral binding (Hendy *et al.*, 2018), though this would imply that the casein found in the limescale exhibits deamidation due to its genuine antiquity - which, due to its provenance, is unlikely. It is also possible that modern laboratory-origin casein bound to the limescale, and that deamidation was induced in sample preparation.

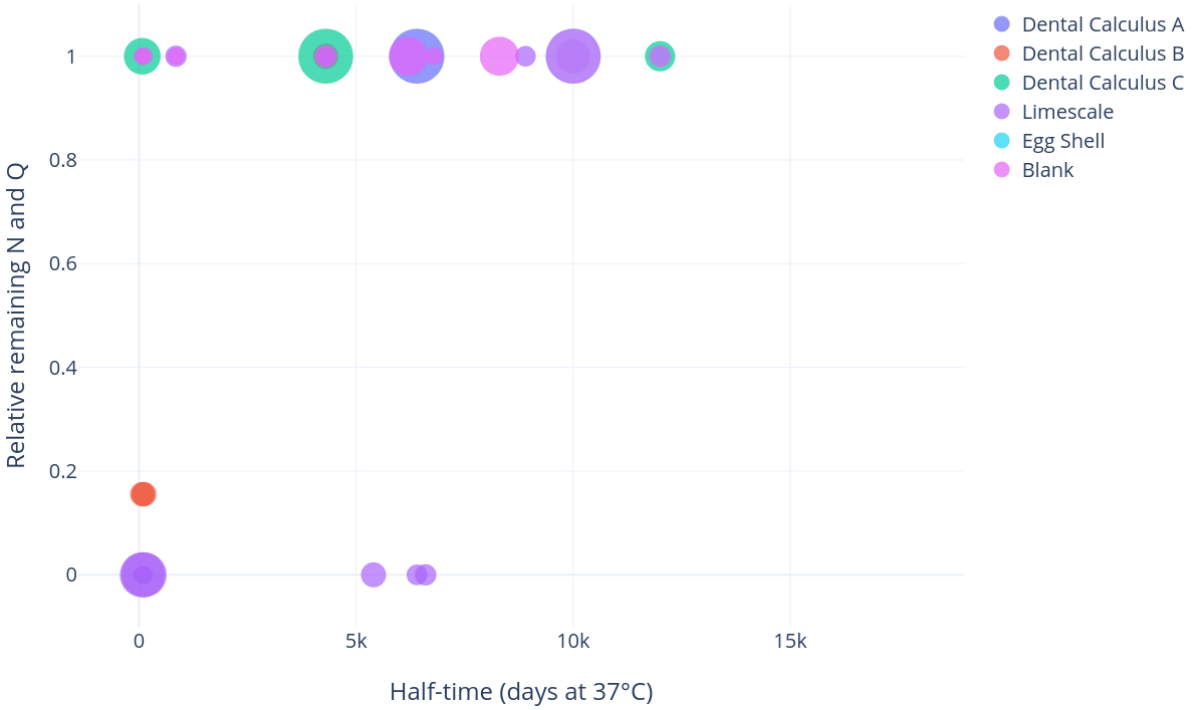


Figure 4.5: *Site-specific deamidation in contamination experiment*

4.3.2 Ancient milk

We characterized deamination patterns in six previously published datasets with putative ancient milk proteins obtained from dental calculus and ceramic vessels. All datasets contained samples that produced confident identifications of BLG and/or caseins (i.e., two or more unique peptides). As shown in Figure 4.6, when data from all samples are grouped, more BLG peptides are detected compared to caseins. In general, caseins are detected relatively infrequently in ancient dental calculus; the only samples to produce abundant casein peptides were modern dental calculus samples (Hendy *et al.,* 2018a) (see Appendix B, Table 6) and ancient ceramic pottery residues from Catalhöyük (Hendy *et al.,*2018b).

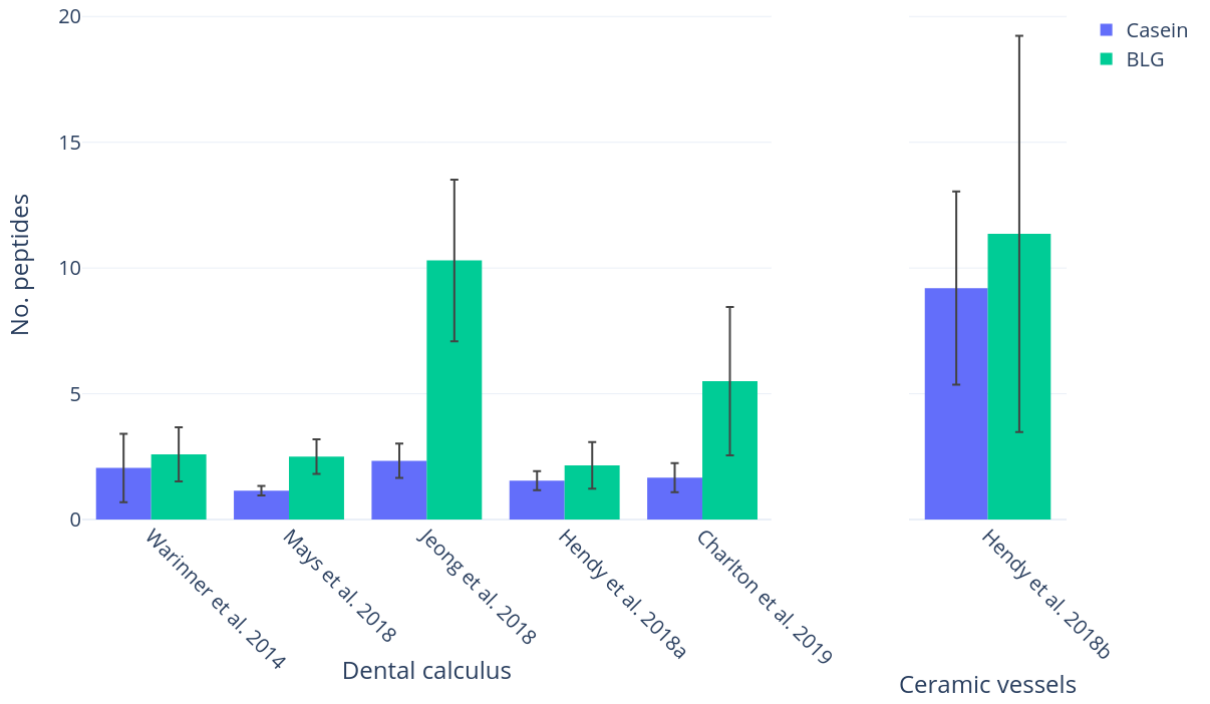


Figure 4.6: *BLG and casein identifications in reanalysed papers. The average number of unique BLG and casein peptides is shown per sample for samples that have at least one peptide. The error bars represent standard deviation.*

We characterized levels of glutamine deamidation in each sample, as shown in Figure 4.7 (average asparagine deamidation is displayed in Appendix B, Figure 1). There is no clear pattern of increasing glutamine deamidation in samples of increasing age for neither BLG nor casein, and variability is fairly high. This could indicate a mix of modern and ancient origins of the peptides recovered or simply the difficulty in tracking the behaviour of multiple amino acids in complex protein structures embedded in widely different mineral matrices, undergoing degradation in different environments.

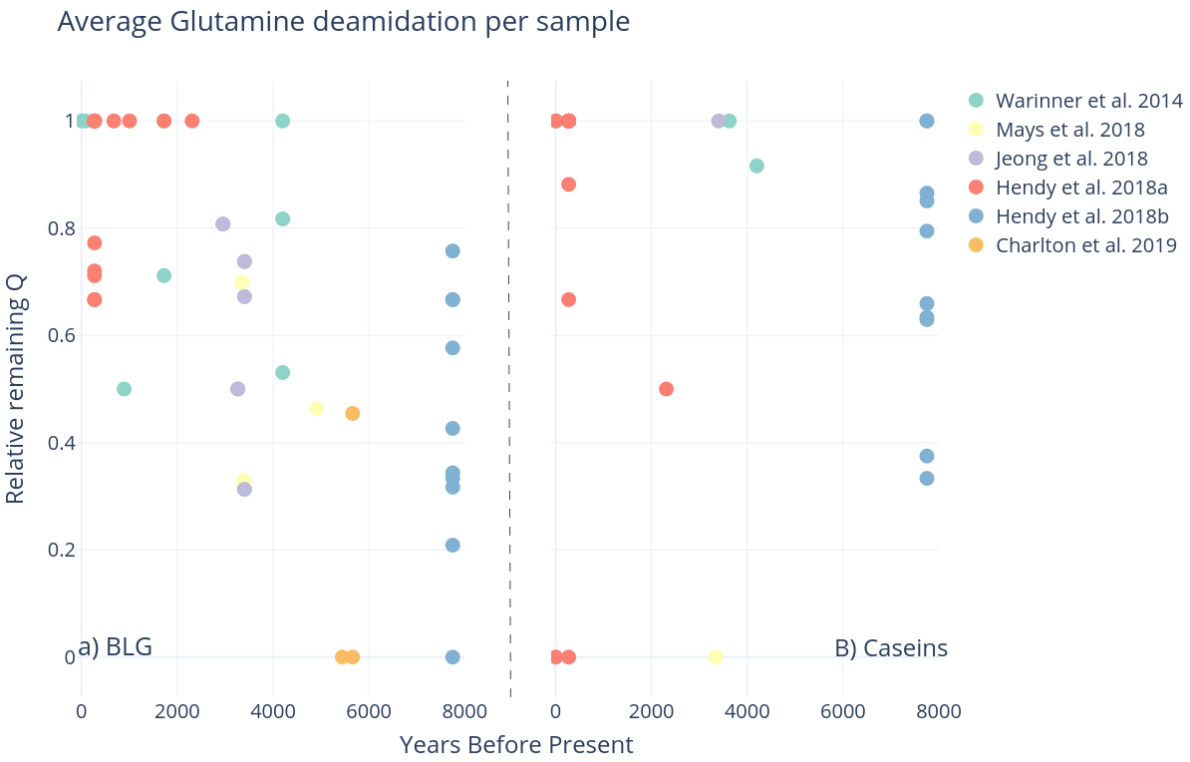


Figure 4.7: Average glutamine deamidation per sample, A) BLG, B) Caseins

We characterized site-specific deamidation for both BLG and caseins in the ancient samples. We expected that high levels of high half-time deamidation would be detected in the most ancient samples — this is not the case, however. Figure 4.8a suggests that BLG deamidation varies greatly, and does not seem to correlate with the age of the sample. Indeed, at the highest half-times detected, three peptides show complete deamidation, yet they are all <1000 year old samples from Hendy *et al.,* (2018a). This highlights how complex post-depositional histories can affect the reliability of this type of analysis if the data are taken at face value.

Deamidation of caseins, shown in Figure 4.8b demonstrates even more variable patterns. Identified peptides across all samples seem to either be completely deamidated, or completely intact. The obvious exception to this is the ceramic vessel residues analysed by Hendy *et al.,* (2018b), which are the only samples to consistently display more intermediate levels of deamidation (discussed below).

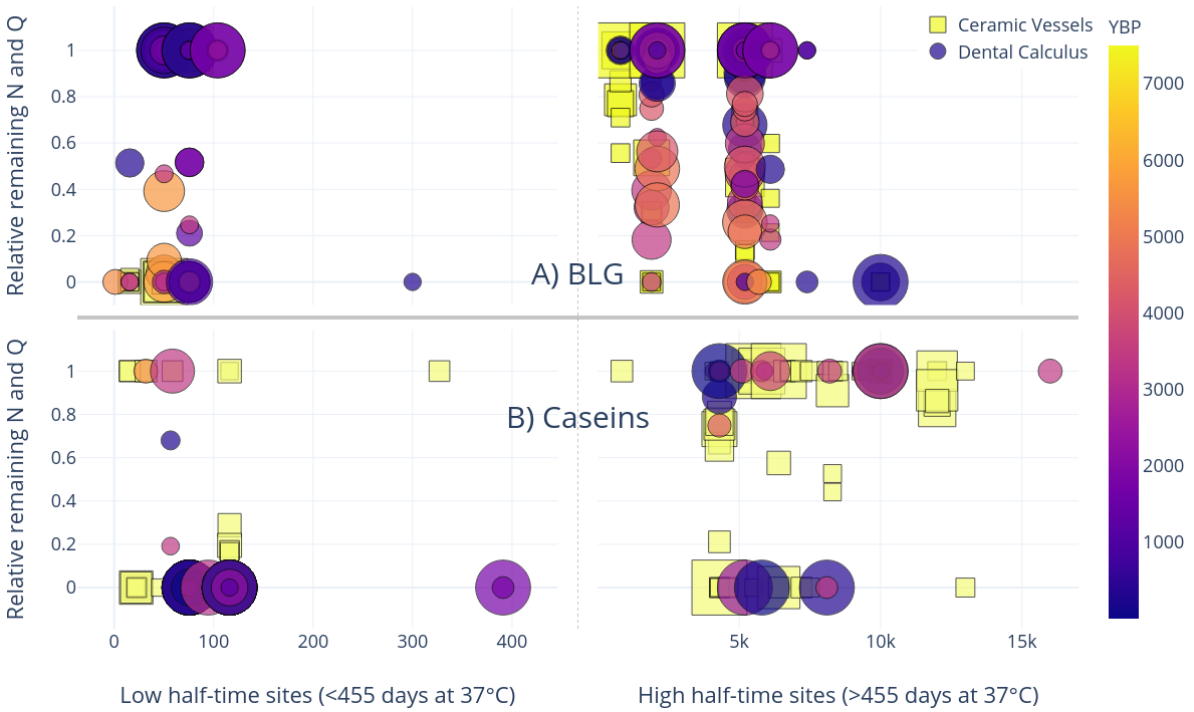


Figure 4.8: *Site-specific deamidation of A) BLG and B) Caseins at low and high half-time sites. Age of the sample in years is shown in the colour of the points. Dental calculus samples are represented by circles, while ceramic vessel origin proteins are shown as squares.*

4.4 Discussion

This study sought to quantify the degradation of milk proteins in archaeological data, reanalysing 274 archaeological samples from 6 papers; along with two experimental datasets. The main aim was to explore the degree to which deamidation could reliably differentiate between modern and ancient milk proteomes.

4.4.1 Modern milk displays limited damage compared to ancient milk

Though the mechanisms involved are complex, and as yet poorly understood, it is clear that modern milk sources generally display less damage than putative ancient milk proteins. As expected, asparagine deamidation occurs relatively rapidly in both casein and BLG, as is shown in the milk powder analysis (Figure 4.1) and the contamination experiment (Figure 4.4). Samples in all three experiments show higher asparagine deamidation than glutamine, although asparagine deamidation is still not very advanced. Very little glutamine deamidation is observed in the experimental samples - even when the sample has undergone extreme heating, such as in the milk powder. While glutamine deamidation is useful as a crude measure of overall degradation, we stress that the extent of variability observed in this and other studies (due to the extreme variability of substrates and environments, as well as the intrinsic nature of the molecules), would prevent a systematic use of glutamine deamidation alone for authentication purposes.

Glutamine deamidation of both casein and BLG seems to somewhat increase over time, as shown in Figure 4.7, however this is far from a clear pattern, and the picture is even fuzzier in the case of asparagine deamidation (see Appendix B, Figure 1). More accurate dating of the archaeological samples may result in clearer trends, as 78% of sample ages are inferred based on broad archaeological periods alone - however, as most of the variation takes place on the y-axis (the extent of deamidation), this fine-tuning of points along the x-axis is unlikely to be conclusive. Furthermore, as the degradation of biomolecules is also affected by a sample's thermal history, the integration of thermal-ages could further improve the correlation of degradation and age (Mackie *et al.*, 2017). Thermal ages take the temperature of the depositional environment into account, and assign a new "age" for a sample, corrected to a constant 10 °C (Demarchi *et al.*, 2016). Therefore, samples of the same age from cold environments have a "younger" thermal age than those from hot environments, whose diagenesis would have happened at increased rates due to

the temperature. An important caveat is that these calculations are affected by a number of errors, due to both the difficulty in estimating the kinetic parameters of the reactions in archaeological samples and the uncertainty related to palaeotemperature reconstructions. Furthermore, the concept of thermal age implies that all other environmental factors affecting diagenesis rates (microbial attack, pH, effect of metals and solutes) are irrelevant - a condition which is only verified in very few biomaterials (avian eggshell, dental enamel) and that does not occur in dental calculus or ceramic containers. However, future studies of open-access palaeoproteomics datasets from a range of archaeological ages and sites, as well as a better understanding of the pathways of diagenesis and mechanisms of degradation may help assessing these issues.

4.4.2 Milk proteins are difficult to authenticate

A key concern in palaeoproteomic analyses is differentiating endogenous, ancient proteins from those originating from post-excavation and/or laboratory sources. Previous publications have outlined best practices for reducing potential contamination in ancient proteins studies, including (but not limited to) the use of a dedicated laboratory for ancient protein extractions, sterile equipment and pure reagents, non-proteinaceous protective equipment (e.g., no latex, silk, leather, wool), extraction blank controls, experimental replicates, and injection blanks between samples during LC-MS/MS (Hendy *et al.*, 2018c). Here, we demonstrate the importance of extracting ancient proteins within a dedicated laboratory, using appropriate controls to limit contamination from reagents and cross-sample handling. We were able to pick up low levels of casein and BLG contamination, merely by exposing archaeological substrates to a laboratory setting in which modern milk products (e.g., skim milk powder, caseins), are routinely used. Both BLG and caseins were detected in the extraction blank, demonstrating the utility of these controls in monitoring for and detecting potential contaminants within the extraction process. It is important to note that the milk contamination detected in this study is unlikely to be the results of instrument carry-over (due to the lack of milk proteins in the injection blanks), but instead likely occurred at the extraction stage. Previously, samples from two of the archaeological sites (i.e., dental calculus samples from St. Helena and the ostrich eggshell from Leatoli) were extracted in a dedicated ancient proteins laboratory, and analyzed on the same LC-MS/MS platforms at the Target Discovery Institute; milk proteins were not detected in these samples. The ostrich eggshell sample was also exposed to high concentration bleach (NaOCl, 12 % w/v) for more than 100 hours - a key step in palaeoproteomics for biomineralised samples, as it ensures the isolation of endogenous (intracrystalline) proteins (see Demarchi *et al.,* 2016). This experiment further

reinforces the importance of authenticating ancient proteins that are commonly present in laboratory reagents, such as milk, blood, and eggs (Hendy *et al.*, 2018a), and the increased burden of proof to demonstrate their authenticity.

Interestingly, caseins were more common in all of the experimental samples compared with BLG, and in all cases were identified with two or more peptides. One of the strategies for reducing the impact of false positive results in mass spectrometry is to only accept protein identifications based on the presence of two or more distinct peptides (Carr *et al.*, 2004; Wilkins *et al.*, 2006). According to this criterion, none of the BLG identifications within the experimental samples would have been deemed confident, as none of the samples displayed more than one peptide assigned to BLG. In contrast, the casein identifications would have met this 'two-peptide rule'. Thus, while multiple distinct peptides may act as one measure of authenticity, it is insufficient to rule out all potential contaminating sources. Another strategy may be to compare the ratio of BLG to casein peptides in ancient samples. Previous analyses of archaeological dental calculus have recovered BLG in higher ratios than caseins (Figure 4.6), and confident casein identifications were primarily recovered from modern dental calculus samples, or skeletons of relatively recent antiquity (e.g., post-medieval). Therefore, in dental calculus studies, a relatively high ratio of BLG compared to caseins, may be further evidence of an endogenous origin; nevertheless, the reasons for this difference in BLG and casein survival are not currently understood.

4.4.3 Database selection significantly influences peptide identifications

It is important to note that differences in peptide identifications may be due to search strategy. When performing searches using an extremely specific database (for example, of ten different milk proteins), there is an increased chance of milk proteins being identified, as the software attempts to maximise the number of matches. When performing an exploratory search, a well-established practice is to perform a search across a wide, yet relevant, database, such as Uniprot (UniProt Consortium, 2015) and/or the Human Oral Microbiome database (HOMD) (Chen *et al.*) to reduce the potential for false positive matches. For example, in their analysis of dietary proteins entrapped in dental calculus, Hendy *et al.* (2018a) noted that multiple putative 'dietary' proteins were identified when searching dental calculus proteomes against Uniprot alone, but that these were identified as false positives after HOMD was also included within the search strategy. Once the presence of the proteins of interest has been validated, searches using smaller databases may be useful for further analyses.

Furthermore, in order to confirm whether recovered peptides that purport to be from a certain protein are indeed unique to that sequence, a BLASTp search (Altschul *et al.*, 1990) should be carried out (see for example the impact of this approach in testing the authenticity of dinosaur peptides in Buckley *et al.,* 2017). Often peptides are not unique to one protein, and, if searched against a wider database, would have matched to several proteins. If a protein identification is made up entirely of peptides that are not unique to that protein, especially if they are shared by common laboratory contaminants, they should be discarded.

4.4.4 Deamidation - how much is enough?

In this study, we also examine the extent to which site-specific deamidation can act as a measure of authenticity for ancient milk proteins, specifically BLG and caseins. This meta-analysis of 274 samples ranging in age from the Neolithic to the present day demonstrates that a moderate overall trend of deamidation is visible in milk proteins over time (Figure 4.7), but the real question is whether this can translate to individual studies, which will likely have a smaller number of samples and a limited time-span. Although the rule of thumb that more deamidation events at high half-time sites imply genuinely ancient proteins is valid for some samples, conclusions like this become increasingly frail as the sample size and the number of peptides recovered decreases. Even proteins which meet the 'two-peptide' rule may lack sufficient data to assess deamidation. For example, if two unique casein peptides are detected in a sample, only one of them may have a glutamine residue. This glutamine residue is either deamidated, or is not. Is a single deamidated residue sufficient to demonstrate authenticity? Likely not, especially considering the sporadic deamidation that was identified in modern milk proteins. DeamiDATE's site-specific deamidation calculation may provide some nuanced information. However, if the glutamine residue in question has a very high half-time, and is totally deamidated, it is more likely the source protein is ancient; and if it has a very low half-time, and is not deamidated, it could be that the protein is a modern contaminant. This type of sporadic, low-level recovery of milk proteins is particularly common in dental calculus studies, where on average in the meta-analysis conducted here, in individuals with any evidence of BLG, on average 3 BLG peptides are recovered per individual. However, only an even smaller subset of these peptides will contain glutamine or asparagine. Only by assessing dozens of individual peptides may clear deamidation patterns emerge. These challenges in authenticating individual peptides and proteins demonstrate the necessity for relatively large and robust datasets in order to apply deamidation as a measure of antiquity. However, the generation of such a dataset may not be possible in all ancient protein

studies due to a lack of surviving peptides. Even when restricting analysis of site-specific deamidation to samples containing more than ten unique BLG peptides, obvious patterns of older samples being more deamidated do not appear (Appendix B, Figure 2).

Hendy *et al.*'s (2018b) study on ancient proteins from ceramic residues at Çatalhöyük demonstrates expected variation of deamidation levels. There are a number of factors which support the authenticity of recovered proteins in this study. First, the authors followed all the laboratory guidelines for the analysis of ancient proteins (Hendy *et al.*, 2018c) sampled from three different parts of the ceramic vessels; the study made use of extraction replicates, which were analysed at different facilities, and which produced consistent results. Moreover, these ceramic vessels produced an average of 13 unique casein peptides and 12 unique BLG peptides per vessel, and 5 different ruminant milk proteins, including whey proteins (BLG), caseins (alpha-S1, alpha-S2, beta-, kappa-casein) as well as a protein regulating the secretion of milk fat droplets. In contrast to the dental calculus studies which typically display the sporadic presence of relatively few BLG peptides, the Çatalhöyük vessels display evidence of a complete milk proteome. As the vessels were excavated from the West Mound of Çatalhöyük, they are assumed to have originated from 7100 BC to 5700 BC, and represent the oldest samples in this meta-analysis. Nevertheless, a large variation of deamidation levels is found in the peptides identified in all the ceramic vessel samples, ranging from 24% to 100% glutamine deamidation in BLG, and 0% to 67% in caseins. This challenges the expectation that thermally old samples should always be completely deamidated.

Deamidation rates may also vary based on archaeological substrate. Here, we examined deamidation levels in previously published datasets from dental calculus (Warinner *et al.*, 2014; Hendy *et al.*, 2018a; Jeong *et al.*, 2018; Mays *et al.*, 2018) and ceramic vessels (Hendy *et al.*, 2018b). Although both substrates display deamidation of milk proteins, they also illuminate clear evidence of differential preservation. In the case of deamidation of casein (Figure 4.8), the ceramic vessels (shown in yellow diamonds) are the only samples that display nuanced levels of deamidation - all other samples have either complete deamidation or none. Although some caseins detected in archaeological substrates may be spurious, the above evidence points to preservation of genuinely ancient caseins in at least some ceramic vessel residues.

As mineralised residues adhering to ceramic vessels and calculus represent different binding substrates, it would follow that the preservation conditions would be different. Furthermore, ceramic vessels would often have contained whole milk (Bleasdale *et al.*, 2021); while dairy proteins found in dental calculus would represent the consumption of either processed or raw

milk - that is, that the input proteins for these two substrates could be different. Further open-access analysis into ancient milk products, such as cheese and kefir, could shed light on how processed and raw milk proteins preserve in the absence of a mineralised substrate. This would allow a much larger abundance of data, and a full range of proteins and peptides, as the milk proteins would not be competing for sequencing against the environmental or oral microbiome proteomes. This would allow researchers to investigate further how protein structure and peptide position influences preservation and diagenesis - including deamidation, other post-translational modifications, and hydrolysis.

4.4.5 Multiple lines of evidence are necessary

Using characteristic signatures of taphonomic damage to authenticate ancient biomolecular data is not new, but faces several unique challenges in palaeoproteomics. Many readers will be familiar with MapDamage (Ginolhac *et al.*, 2011; Jónsson *et al.*, 2013), a commonly-implemented software program designed to quantify damage patterns in ancient DNA data based on the size distribution of recovered sequencing reads and the frequency of cytosine to thymine (C-to-T) misincorporations at 5' ends of sequences. The fundamental principle behind deamiDATE is the same, namely that we are trying to use molecular damage to authenticate ancient sequences. Nevertheless, there are notable differences between the two approaches as a consequence of the different structures of the two biopolymers, and the methods of analyses. Firstly, given that the extent of diagenetic (natural) hydrolysis is typically unknown in ancient samples, it is common to enzymatically cleave ancient proteins to obtain shorter sequences, while in aDNA chain scission is a wholly diagenetic phenomenon. Therefore, tryptic peptide length distributions cannot offer the same insight into preservation as DNA fragment size distributions. Non-tryptic peptides do however offer some insights into the extent of molecular fragmentation. Moreover, peptides may have been exposed on the surface or buried deep within a structure, yet the dielectric constant is known to impact rates of deamidation (Wakankar and Borchardt 2006). The double helix of DNA is largely a fixed molecular organization. There are 20 amino acid residues as opposed to four (five) nucleotide bases and consequently there is a much greater variation in structural organization of proteins. In ancient DNA there is periodicity in fragment length resulting from the wrapping of DNA around histones (Kistler *et al.*, 2017). It is known that deamidation rates in proteins vary as a consequence of entropic effects (e.g., secondary structure (Xie and Schowen, 1999)) and we would envisage deamidation rates will be more difficult to assess and are impacted by the extent of degradation/denaturation. Finally, position dependent nucleotide

misincorporation patterns may be assessed in every DNA sequence mapping to the genome of interest, and thus produce taphonomic data at millions or billions of distinct nucleotide sites. In contrast, deamidation can only be assessed at two of the 20 amino acid sites, limiting the number of datapoints available for analysis.

To increase the quantity of peptides available for site-specific analysis, we propose that the deamidation levels of proteins of interest, such as, in this case, dietary proteins, are compared to proteins that are likely to be endogenous to the substrate. For example, dietary proteins originating from dental calculus should be compared to endogenous proteins from the host. The similarity, or lack thereof, of the diagenetic status between the endogenous host proteins and the protein of interest could be used as further evidence towards genuinity. This approach has been informally applied in past analyses (Mackie *et al.*, 2017; Wilkin *et al.*, 2020), but should be adopted by future ancient protein studies.

Information about protein structure may also provide additional insight into protein preservation and authentication. For example, in this study, we explored to what extent the various regions of beta-lactoglobulin were represented in peptides recovered from modern and ancient milk (Appendix B, Figure 3). Our analysis demonstrated that while the beta-bridge region was well represented in skim milk powder this region was virtually absent in all of the ancient dental calculus and ceramic samples (Appendix B, Figure 4). Although the absence of beta-bridge peptides in ancient calculus and ceramic samples may reflect both differential preservation, protein aggregation and/or differential mineral binding, it suggests that, with further research, assessments of protein coverage may provide additional signatures for ancient protein authentication.

4.5 Conclusion

Some milk proteins discovered in archaeological artefacts show deamidation signatures distinct from modern milk samples, though this damage is highly variable, and a large sample size is required to elucidate such patterns. This study reinforces the importance of the use of dedicated facilities for the extraction of ancient proteins, free from modern contaminants that could obfuscate results of archaeological importance - such as milk powder. Lastly, site-specific deamidation represents an additional method to authenticate ancient milk proteins, but damage patterns may only be robustly assessed through the careful analysis of large-scale data. Thus, it is

increasingly important that raw data resulting from palaeoproteomics experiments is made open access, so that more meta-analyses may shed more light on the patterns of degradation in ancient proteomes and across various archaeological substrates.

Chapter 5: Collagen type I degradation in limed skins

Abigail Ramsøe[1,2], Sean Doherty[3], Matthew Collins[4,5]

Corresponding author: Abigail Ramsøe

Affiliations:

1. BioArCh, Department of Archaeology, University of York, York, UK
2. Department of Earth Sciences, Natural History Museum, London, UK
3. Department of Archaeology, University of Exeter, Exeter, UK
4. The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark
5. McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, UK

Relevance: This chapter assesses the degradation of collagen through a time series of skins that have been subjected to liming; with a focus on what drives collagen diagenesis.
This chapter represents the sum of the work I have done on this dataset, and I have plans to refine it for submission, most likely to Heritage Science.

Abstract: As the primary writing medium in Europe, North America (post-Columbian) and the Near East for over 1000 years, parchment (processed animal skin) is one of the most abundant resources available for the study of past societies and cultures. Tens of millions of parchment skins were manufactured, many of which survive today in archives, libraries and private collections. It is therefore important that they are conserved in such a way that the textual and biological information they record can be preserved. However, the degradation of collagen in parchment is poorly understood. Here, we elucidate the patterns of degradation in experimentally limed parchment in order to better understand the diagenetic process. It is observed that during liming the peptide count, coverage and deamidation patterns progress at a predictable rate. Furthermore, a hotspot of hydrolytic activity is discovered, and its mechanisms are explored. Lastly, it is shown that the skin removed from stillborn lambs is dramatically different from that of flayed older sheep, but that after liming, there is no detectable difference on the molecular level. This study represents the most in-depth biomolecular analysis of collagen parchment to date, and sheds light on the driving forces behind its degradation.

5.1 Introduction

Historical manuscripts represent an important source of information about past societies and cultures. Though the origins of parchment are under great debate (Johnson, 1970; Reed, 1975); the first mention of documents written on skins was in the Egyptian 4th Dynasty (c. 2613 to 2494 BCE), and parchment was widely used until the mass production of paper in the 19th century (Ryder, 1964). Such documents are often easily dated due to direct references in the text, or the date can be inferred from palaeographical details. The biological study of parchments (termed 'biocodicology' Fiddyment *et al.* 2019)) is a more recent field of study, and can give researchers insight into parchment production, dietary and husbandry practices, and handling and conservation of the object (Teasdale *et al.*, 2017; Fiddyment *et al.*, 2019; Demarchi *et al.*, 2020).

During parchment production, perishable animal skins (Figure 5.1) - typically sheep, goat, and cow (Reed, 1972) - are transformed into durable sheets historically used as a writing medium. Firstly, the skin is removed from the animal and subsequently placed into a highly alkaline lime solution. This is in order to: break the disulfide bonds in hair and epidermis, making it easier to remove them (Bienkiewicz, 1983; Covington, 2009); remove cutaneous lipids which makes parchment white and able to absorb inks (Koppenhoefer, no date); hydrolyse the amide groups attached to asparagine and glutamine, therefore lowering collagen's iso-electric point, facilitating the removal of non-collagenous proteins (Menderes *et al.*, 1999) and in the case of modern chrome tanned leather offering more sites for interaction with Cr salts. After liming, the skin is fleshed to remove fats, muscle, and hair follicles, leaving just the collagen-rich dermis layer. The skin is washed in order to remove lime introduced during the production process, stretched to prevent it becoming brittle, and lastly dried under tension (Doherty *et al*., 2021)

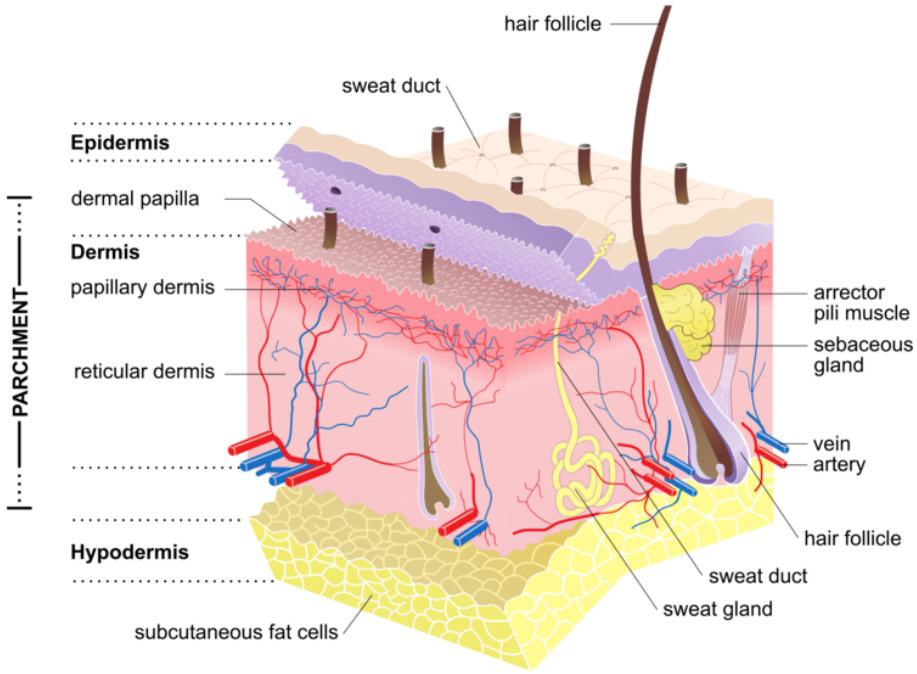


Figure 5.1: Structure of mammalian skin and the layers typically present, and removed, during parchment manufacture (S. Doherty, Wikimedia Commons)

The primary protein in parchment is collagen, of which the majority is type I, with a further 5-40% type III (decreasing with the animal's age and the depth of the dermis) (Axelsson *et al.*, 2016). Collagens are commonly found in ancient samples (van Doorn *et al.*, 2012) due to their unique stability caused by the triple-helical arrangement and intra- and inter-molecular crosslinks (Némethy and Scheraga, 1986; Wojtowicz *et al.*, 1999; San Antonio *et al.*, 2011). All collagens are composed of three chains. In collagen type I, there are two identical collagen I alpha 1 chains (coded by the gene *COL1A1*), and one collagen I alpha 2 chain *(COL1A2)*. These three polypeptides coil round each other in a triple helix structure. During synthesis, collagen chains are assembled as longer procollagens with non-helical propeptides at both the N- and C-terminal ends. When the procollagens are fully formed and are secreted from the cell, these propeptide ends are cleaved off, leaving the final protein with an entirely helical structure (Lodish *et al.*, 2000), also known as a tropocollagen.

The structure of collagen is determined by a unique amino acid composition. The presence of a glycine residue at every third position, gives collagen its characteristic helical structure, as only the side chain of the glycine is small enough to fit into the center of the helix. Proline and hydroxyproline (the product of the post-translational modification of proline) also support the triple helical structure, as their side chains point outwards. These tropocollagens pack themselves together in groups of fibrils. In collagen type I, these are staggered so that between

each tropocollagen there is a gap of around 67 nm or 234 residues. Therefore, every unique position along the tropocollagen fibril is either an "overlap zone", where there are five molecules, or a "gap zone", where there are only four (Piez and Trus, 1977), as shown in Figure 5.2.

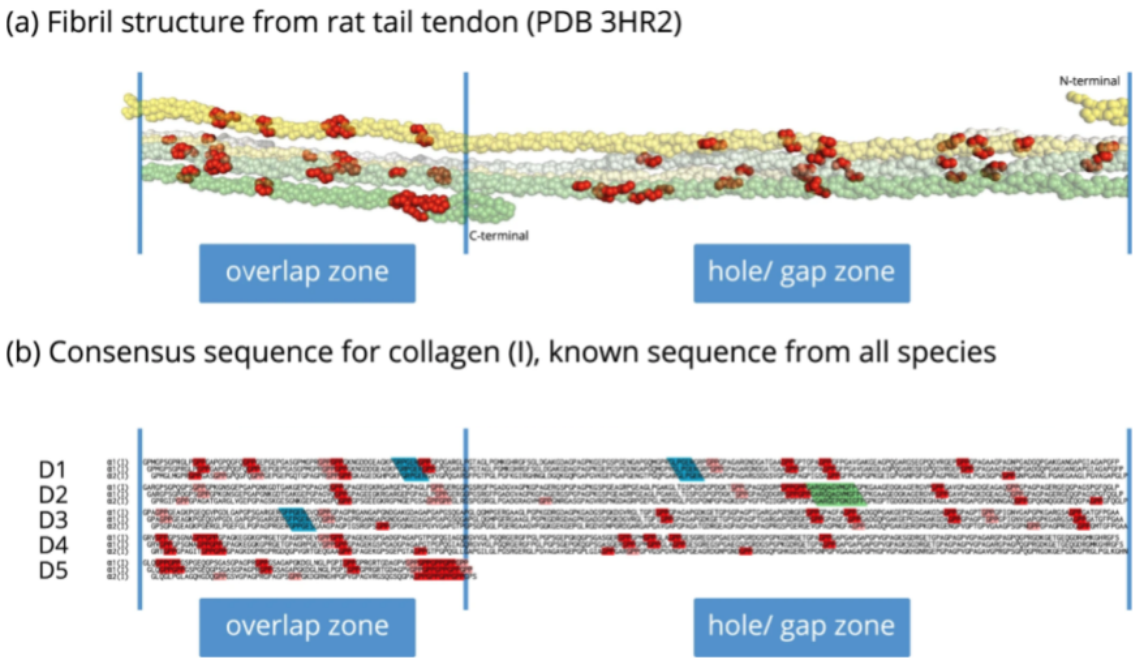


Figure 5.2: *The physical structure of collagen I, showing the overlap and gap zones. From Chow et al. (2018) (Figure 6). The coloured annotations are irrelevant to this study.*

Although *in-vivo* collagen degradation is well-understood, and there is a plethora of research into archaeological collagen in bone (Collins *et al.*, 1995; Smith *et al.*, 2007; Dobberstein *et al.*, 2009), the mechanisms of collagen degradation in parchments are unclear (Kennedy and Wess, 2003). Understanding the degradation and its causes will allow both the refinement of conservation techniques and the development of targeted analytical techniques. This chapter aims to better understand the diagenetic changes skins undergo during parchment production. We used LC-MS/MS to characterise the content of five experimentally produced parchments at different stages of the liming process in order to quantify the degradation of collagen and elucidate the patterns of degradation.

5.2 Materials and Methods

5.2.1 Experimental parchment production

Fresh skins were obtained from five sheep (*Ovis aries*), details of which are shown in Table 5.1. Skins were removed from the animals shortly after death, and subsequently salted and stored at 5°C prior to analysis. The skins were processed into parchment using traditional pre-19th century European techniques (Saxl, 1954; Ryder, 1964; Reed and Reed, 1972), without the use of additional chemical depilatory agents (e.g. sodium sulfide) or deliming acids (e.g. formic acid) (Saxl, 1954; Ryder, 1964; Reed and Reed, 1972).

Skin ID	**Breed**	**Sex**	**Colour**	**Age**	**Parchment Batch**	**LC-MS/MS Batch**
Skin 06	Shetland	Male	Fawn	10 months	2	1
Skin 08	Portland	Female	White	13 months	2	1
Skin 11	Shetland	Female	White	Stillborn	2	2
Skin 12	Shetland	Female	Brown	5 years	2	2
Skin 15	Shetland	Female	Brown	18 months	3	1

Table 5.1: *Sheep skins used for parchment production*

Each skin was washed and rehydrated in water at 8 °C (pH 7.5) for 48 hours. The water was replenished every 8 hours and any foreign materials adhering to the skins were removed. The unsplit skins were submerged in a 3.5% calcium hydroxide solution (pH 13.5) in 220 litre high-density polyethylene (HDPE) drums at room temperature for 6 to 18 days, and agitated three times per day to ensure an even exposure across the skin. The ability to remove the fibre was appraised daily, and deemed sufficiently limed when it could be removed at the root by hand with ease (Covington, 2009). Once de-haired, the epidermis and subcutaneous tissues were removed with a double-handled knife. The skins were subsequently returned to the lime for a further 12 hours, allowing for a uniform penetration of the solution, which may have been inhibited by hair or fat. The skin was then mechanically squeezed with the knife to force out as much liquid as possible (in a process known as 'scudding'). To neutralise the alkalinity, each skin was washed vigorously for 30 minutes in running water and then allowed to soak for 48 hours, with the water replaced every 8 hours. Each skin was then tightly stretched with ropes on a wooden frame, and allowed to dry under tension. While still wet and then again once dry, each side was shaved with

a sharp knife to remove further layers of the dermis and produce a clean even surface (Doherty *et al*., 2021)

5.2.2 Sampling

Samples were taken from skins prior to liming, and again after 5, 10, and 15 days of liming using the non-invasive eraser technique of Fiddyment *et al*., (2015). No replicates were performed. After cleaning the peptides on C18 ZipTips, the samples were eluted with 50% ACN and dried in a vacuum centrifuge for transport.

5.2.3 LC-MS/MS analysis

The cleaned and vacuum dried peptides were transferred to the Novo Nordisk Foundation Center for Protein Research, the University of Copenhagen for LC-MS/MS analysis on an EASY-nLC 1200 (Proxeon, Odense, Denmark) coupled to a Q-Exactive HF-X (Thermo Scientific, Bremen, Germany). The samples were resuspended in 50 µL of 80% acetonitrile (ACN) and 0.1% formic acid (FA), and 10 µL of each sample was placed in separate wells on a new 96-well plate. They were then vacuum centrifuged until near dry and resuspended with 10 µL of 0.1% TFA 5% ACN, and 5 µL of sample analysed by LC-MS/MS. Previously published LC-MS/MS parameters for palaeoproteomic samples were used (Mackie *et al.*, 2018), in short: MS1: 120k resolution, maximum injection time (IT) 25 ms, scan target 3E6. MS2: 60k resolution, top 10 mode, maximum IT 118 ms (108 ms in the case of Skin 11 and Skin 12), minimum scan target 3E3, normalised collision energy of 28, dynamic exclusion 20 s, and isolation window of 1.2 m/z.

5.2.4 Data analysis

Resulting raw files were searched using MaxQuant 1.6.2.6a (Cox and Mann, 2008) against a database containing only sheep collagen type 1 alpha-1 and alpha-2 (Appendix C, File 1), using a non-specific search strategy (i.e. no protease was selected). The variable modifications included Hydroxyproline, Glu and Gln to pyro-Glu, deamidation (NQ), acetyl (N-term) and oxidation (M). All other parameters were set to MaxQuant's defaults, including a false discovery rate (FDR) of 1%. A further search was performed against the sheep proteome (retrieved from UniProt 2020-08-06) with more stringent search parameters: semi-specific cleavage, a minimum score for modified and unmodified peptides of 60, and a contaminant search. All other parameters remained the same.

The hydrolysis of collagen in the quarter stagger structure was visualised with an in-house script (https://github.com/abigailramsoe/Parchment) written in Python 2.7 (Van Rossum and Drake, 1995), using the packages NumPy (Walt *et al.*, 2011), and matplotlib (Hunter, 2007). Peptide profiles were produced with Peptigram (Manguy *et al.*, 2017). Deamidation was calculated and characterised using deamiDATE 1.0 (Ramsøe *et al.*, 2020).

5.3 Results

5.3.1 Hydrolysis

Figure 5.3 displays the patterns and sites of hydrolysis in all the skins, grouped by time point. The COL1A1 and COL1A2 are arranged by D-period, in a similar fashion to Figure 5.2, in which there are five main rows, each containing two COL1A1 strands, and one COL1A2 strand, each offset by one residue. Hydrolysis events are shown as vertical bars, with the height based on the intensity of the cleavage. Amino acids are coloured using a modified version of the RasMol shapely colour definitions (Sayle and Milner-White 1995), as shown in Table 5.2.

Alanine	Glycine	Leucine	Serine	Valine
Threonine	Lysine	Aspartic acid	Isoleucine	Asparagine
Glutamic acid	Proline	Arginine	Phenylalanine	Glutamine
Tyrosine	Histidine	Methionine	Tryptophan	Cysteine

Table 5.2 *Amino acid colour scheme*

The samples from skins before liming (Figure 5.3.A) show the least hydrolysis, with most high-intensity cuts being at arginine and lysine (dark blue). There is also evidence of low-intensity hydrolysis occurring at residues adjacent to these high-intensity cuts, for example at row-group 3 at around position 140, and row-group 4 and position 181-205 (gene position 1069 - 1093). After five days of liming (Figure 5.3.B), more “random” (i.e. at non-tryptic sites, and not near any other hydrolysis sites) hydrolysis occurs, and the low-intensity hydrolysis identified in the fresh skins grows in intensity. After ten days of liming (FIgure 5.3.C) there are high-intensity hydrolysis sites fairly evenly distributed across the collagen chain, including a growing amount at the end of the COL1A2 molecule (row-group 5, positions 1-50), which may indicate further fraying. The region of high-intensity hydrolysis at row-group 4 around position 181-205 has continued to hydrolyse, and is clearly the most damaged part of the triple helical structure. After 15 days of liming (Figure 5.3.D) the hydrolysis in this region is so intense that it has exceeded the maximum bounds for the plot. Interestingly, the COL1A2 chain adjacent to this (row-group 4, sub-row 3 -

shown in the green boxes) does not exhibit particularly intense hydrolysis, which seems to indicate that although the COL1A1 strands are hugely frayed at this point, the triple helical structure is still intact. Additionally, the region directly above this high-intensity region (row-group 3, position 175-230) experiences only very limited hydrolysis, even at lysine residues (that should be cut by trypsin), and after 15 days of liming treatment.

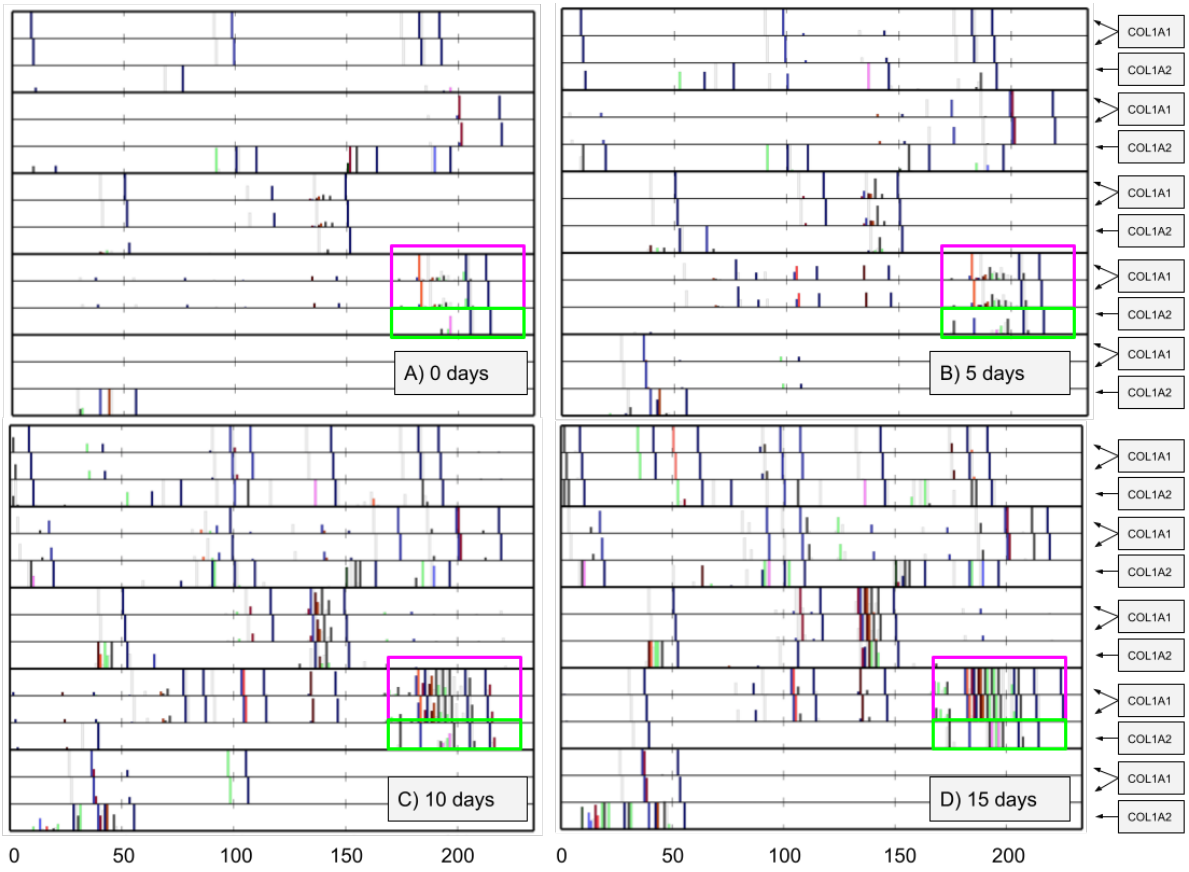


Figure 5.3: *Collagen hydrolysis in all skins before liming (A), and after 5 (B), 10 (C) and 15 (D) days of liming. In order to preserve the nuances in the lower-intensity hydrolysis sites, for all plots the y-axis is limited to a maximum of 100,000,000 intensity. The pink box shows the region of high intensity hydrolysis at gene position 1069 - 1093. The green box below this shows the corresponding region in COL1A2.*

5.3.1.1 Hydrolysis of the high-intensity region

This high intensity region (row-group 3, position 175-230 - positions 1069 to 1093 in the gene numbering system) of COL1A1 in Skin 6 is shown in Figure 5.4. As combining all skins may lead to patterns that are due to natural variation between skins rather than the liming process, Skin 6 was chosen as it is a representative skin, in the middle of the age range. The high-intensity cleavage region is bookended by two tryptic residues - lysine (K) at 1069 and arginine (R) at 1091 - and also contains an arginine at residue 1073. Before liming (Figure 5.4.A), all these three have been cleaved by trypsin and represent the highest intensity cleavages at this time point. Additionally, there is some "random" (i.e. non-tryptic) cleavage in the middle of this zone before liming, between 1078 P and 1079 A, and 1079 A and 1080 G.

After five days of liming (Figure 5.4.B), the intensity of the tryptic residues has increased. As also shown in Figure 5.5.A, many low intensity cleavage events occur near the sites that were already hydrolysed before liming, for example around 1073 R - the four residues immediately before and after begin to exhibit hydrolysis after five days. This is indicative of the tryptic cleavage being the catalyst for more hydrolysis due to structural damage. After ten days of liming (Figure 5.4.C and Figure 5.5.B) nearly every amino acid in the high-intensity hydrolysis region is hydrolysed, and after 15 days (Figure 5.4.D and Figure 5.5.C) the intensity of this hydrolysis increases, though no new peptides are found.

Within this region that experiences this intense hydrolysis, there are five pairs of amino acids that there is no hydrolysis between, even after 15 days of liming. These are shown in Table 5.3. Of these, only two pairs are ever found hydrolysed in the other skins. 1070 S - 1071 G and 1081 P - 1082 A are both found hydrolysed in skins 11 and 8. The rest are never hydrolysed in any skin nor liming stage. This, coupled with the fact that they lie within four residues of each other, points to a structural or physical reason for their apparent robustness.

N-terminal side	C-terminal side
1070 S	1071 G
1081 P	1082 A
1084 P	1085 I
1085 I	1086 G
1087 P	1088 V
1088 V	1089 G

Table 5.3: *Positions in the intense hydrolysis region that never exhibit cleavage in Skin 6.*

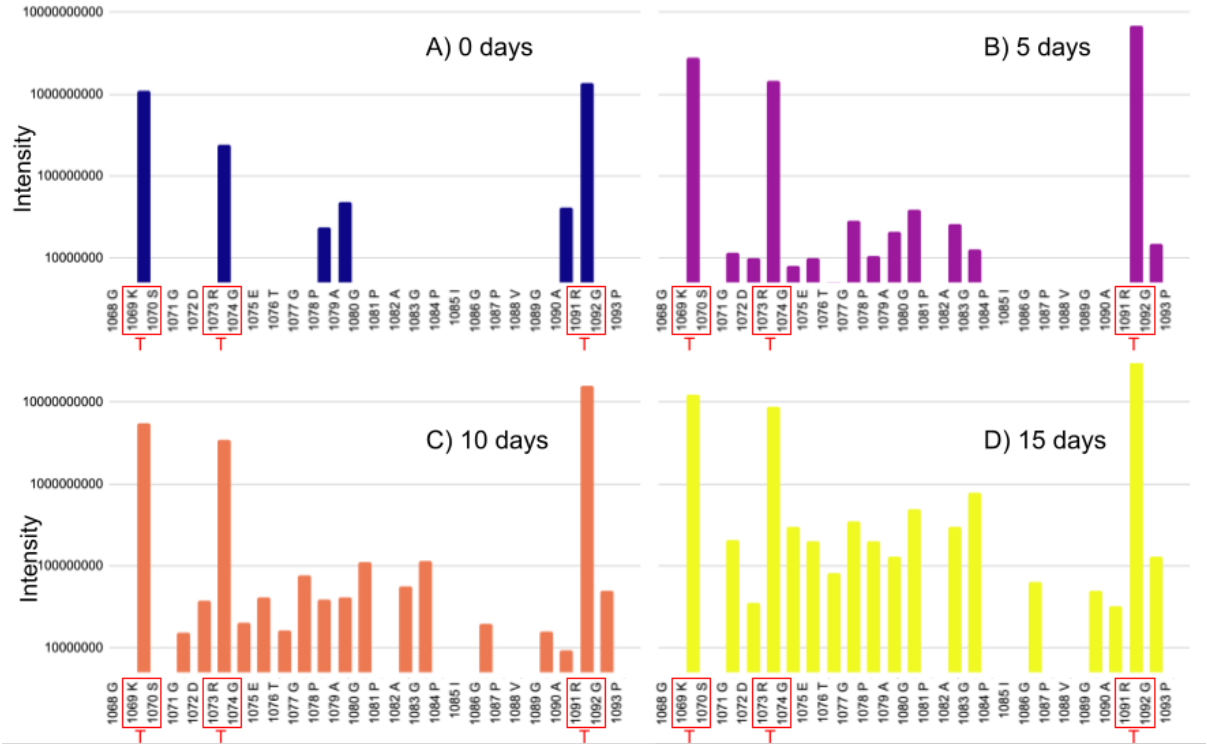


Figure 5.4: *Intensity of hydrolysis of high-intensity cleavage region in COL1A1 in Skin 6 at each time point. Bars represent cleavage between the two residues they are between on the x-axis. T represents hydrolytic events caused by the action of trypsin.*

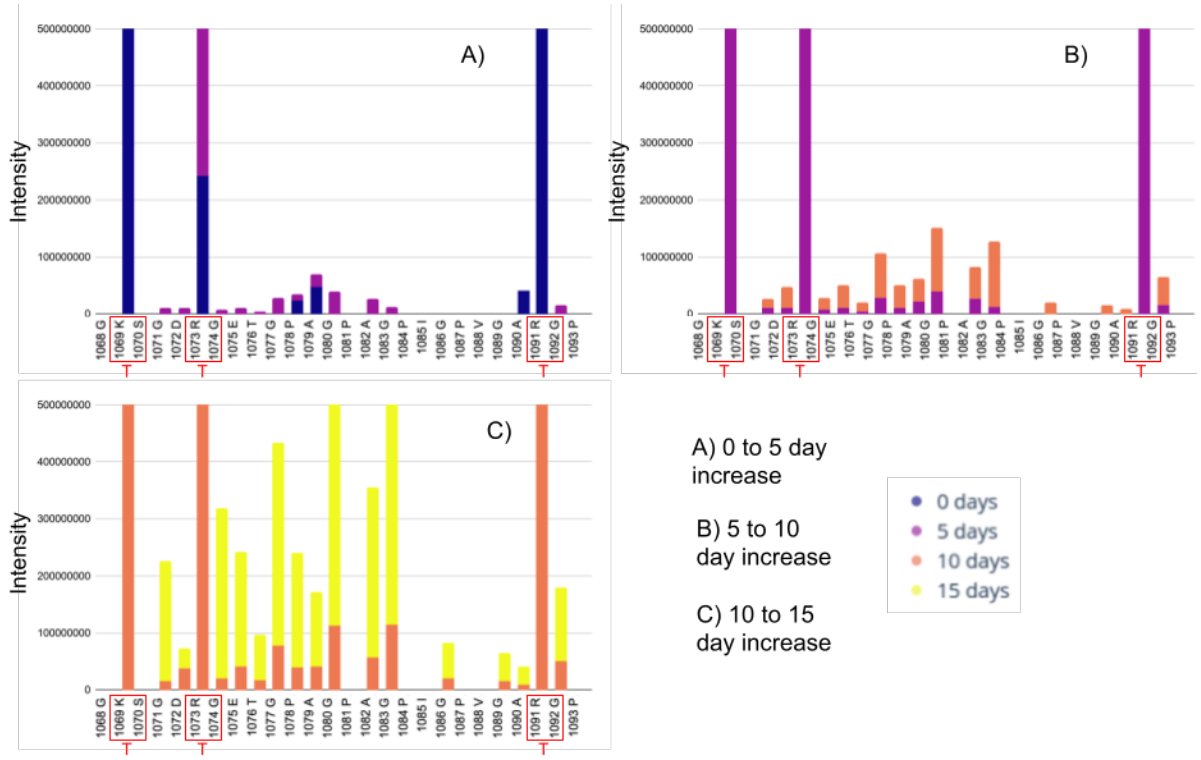


Figure 5.5: *The increase of intensity of hydrolysis of high-intensity cleavage region in COL1A1 in Skin 6 between each time point. T represents hydrolytic events caused by the action of trypsin.*

5.3.1.1.1 Skin 11 hydrolysis

Although this high-intensity cleavage region is present in all of the analysed skins, the way the hydrolysis develops can differ. Figure 5.6 shows the differences between Skin 6 and the stillborn sheep skin, Skin 11. The difference is already apparent before liming (Figure 5.6.A), Skin 11 exhibits lots of low intensity hydrolysis across this region, whereas Skin 6 (and indeed, the other skins) showed very little random cleavage before liming. After five days of liming (Figure 5.6.B), there are no non-tryptic cleavage events detected in Skin 11 (that meets the lower bound of the plot), which is in stark contrast to Skin 6 (n=11) and even Skin 11 before liming (n=13). After 10 days of liming (Figure 5.6.C), the Skin 11 has a similar pattern (though with somewhat lower intensity) of hydrolysis to Skin 6, and the cleavage between 1070 S and 1071 G is detected. Lastly, by 15 days of liming (Figure 5.6.D), the two skins exhibit strikingly similar patterns and intensity of hydrolysis, with the only easily detectable differences the cleavage between 1070 S and 1071 G, and the newly formed cleavage between 1081 P and 1082 A.

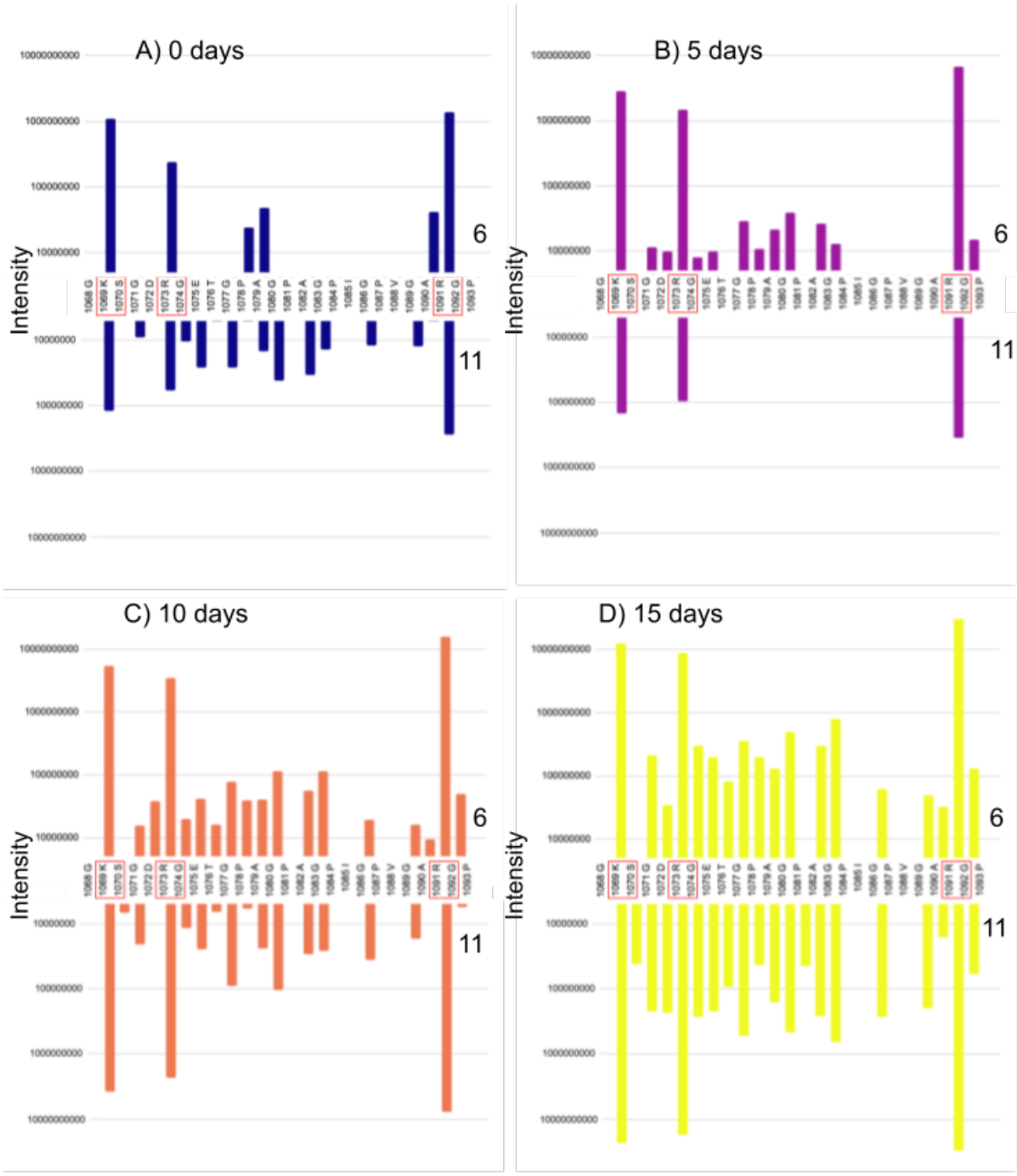


Figure 5.6: *Intensity of hydrolysis of the high-intensity hydrolysis region at each time point, with Skin 6 on the top y-axis, and Skin 11 on the bottom. The red boxes represent hydrolytic events caused by the action of trypsin.*

5.3.1.2 Hydrolysis per amino acid

The hydrolysis events per unique residue was investigated in order to determine whether some amino acids are more likely to be cleaved during liming. Figure 5.7 shows the intensity of hydrolysis in both COL1A1 (Figure 5.7.A) and COL1A2 (Figure 5.7.B) per time point per amino acid. In COL1A1, hydrolysis increases in most amino acids fairly uniformly at each time point. After 15 days of liming, the intensity of hydrolysis of some amino acids seems to increase dramatically - for example tyrosine and glutamine.

The most common semi-tryptic peptide arising from a cleavage of tyrosine is GISVPGPMGPSGPR, which starts at 181 in COL1A1 and has a T at position 180. Across skins 6, 8, 12 and 15, this peptide is never identified before liming, identified six times after five days of liming, 20 after ten days, and 77 times after 15 days. This reflects the type of exponential growth shown in Figure 5.7.A. As this peptide starts in the propeptide region of COL1A1 and ends in the triple helical region, it is most likely very susceptible to liming due to its location. Furthermore, as shown in Figures 5.4 and 5.5, there is a tyrosine in the middle of the high intensity hydrolysis region that only experiences hydrolysis after five or more days of liming.

For glutamine, there are only two peptides that show substantial hydrolysis. One of which is GPPGSAGTP*, which starts at 1148 and has a Q at 1147. The G that starts at 1148 gives rise to 11 different “versions'' of this peptide. The most common is the semi-tryptic peptide, GPPGSAGTPGK, which appears at all time points (0=2, 5=11, 10=15), but mostly after 15 days of liming (n=20). The other versions, of which all except GPPGSAGTP include a missed tryptic cleavage at the lysine, are only seen in significant quantities after 15 days of liming. Though this peptide is covered 20 times across the non-outlier skins after 15 days of liming, it only makes up 8% of the total intensity of peptides with a glutamine cleavage.

The peptide (which always ends at a lysine at position 991, and is preceded by a glutamine at 992) that makes up 91% of the total intensity of glutamine cleavage is shown in Table 5.4. The fully tryptic form of this peptide, shown in row 1, is most common, and the semi-tryptic forms, especially one with a N-terminal proline (row 3) are mostly found after high duration liming. However, because this peptide ends with a lysine, it is unlikely its high level of hydrolysis is due to the presence of the glutamine, rather than the action of trypsin on lysine. This peptide is notable as it is just 30 residues away from the collagenase cleavage site (and 960 - 961 (Chung *et al.*, 2004)), so therefore originates from a region of the collagen chain that is likely to be frayed.

Peptide	Amino acid before	0 days	5 days	10 days	15 days
GFPGLPGPSGEPGK	R	14	25	45	55
FPGLPGPSGEPGK	G	2	0	0	0
PGLPGPSGEPGK	F	0	2	11	22
GLPGPSGEPGK	P	0	1	0	1
LPGPSGEPGK	G	0	0	0	1
PGPSGEPGK	L	0	0	2	6

Table 5.4: *The peptide/s responsible for 91% of the glutamine hydrolysis in COL1A1*

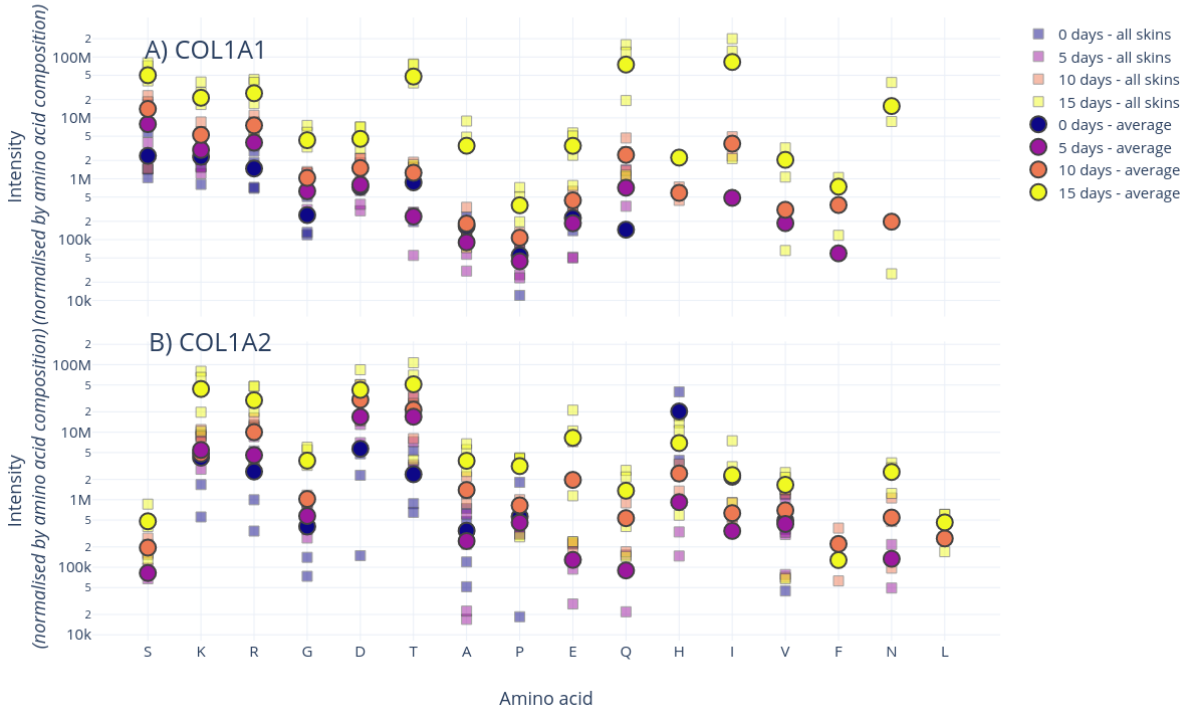


Figure 5.7: *Hydrolysis per amino acid in COL1A1 and COL1A2 in each time point. Amino acids are ordered by their initial average intensity in COL1A1 before liming. Intensities represent average intensities for all peptides in a certain skin and time point where the amino acid is at the N- or C-terminal end and the intensity is not 0, normalised by the starting amino acid composition of each protein. Skin 11 is not included in this plot.*

Like COL1A1, COL1A2 (Figure 5.7.B) exhibits a general pattern of intensity increasing in most amino acids as the duration of liming increases.

One of the most noticeable differences is that hydrolysis at serine is far less likely to occur in COL1A2, despite being the most hydrolysed amino acid in COL1A1 before liming, and both proteins having comparable amounts of serine (3.75% in COL1A1 and 3.03% in COL1A2). There are two primary reasons behind this serine hydrolysis: firstly, there is a serine at position 179 which appears in (or directly before or after) three peptides with a cumulative coverage of 134 across the non-outlier skins (excluding Skin 11) after 15 days of liming. Secondly, there is a serine at 1070 in the region of intense hydrolysis, which is directly preceded by a lysine at position 1069, which may be responsible for some of this hydrolysis assigned to serine.

However, in COL1A2 there are three occurrences of a serine preceding an arginine. These peptides are only detected very sporadically (two are detected once in total; one is detected thrice but in different skins), and are never assigned an intensity value by MaxQuant.

5.3.1.2.1 Skin 11 hydrolysis per amino acid

Hydrolysis per amino acid in Skin 11 was compared to Skin 6, a representative example, in order to ascertain whether Skin 11 exhibited different patterns during liming. Surprisingly, the overall pattern was generally the same, despite Skin 11's different pattern of hydrolysis in the high-intensity region in COL1A1.

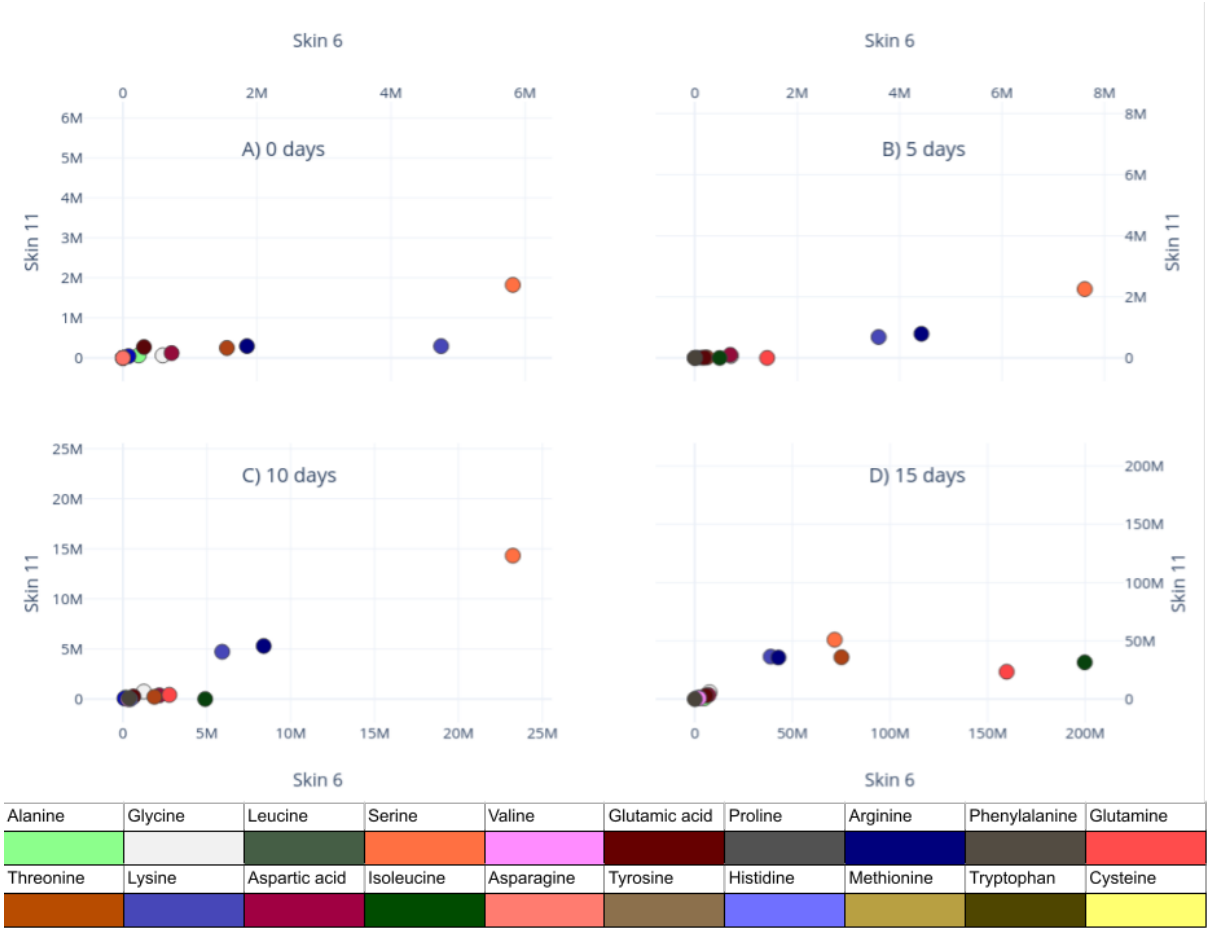


Figure 5.8: *Hydrolysis per amino acid in Skin 11 (y-axis) versus Skin 6 (x-axis).*

5.3.1.2.2 Gap vs overlap hydrolysis per amino acid

The differential hydrolysis was also investigated in the gap and overlap region. As shown in Figure 5.9, there is substantially higher hydrolysis for some amino acids in the gap region than the overlap before and after five days of liming.

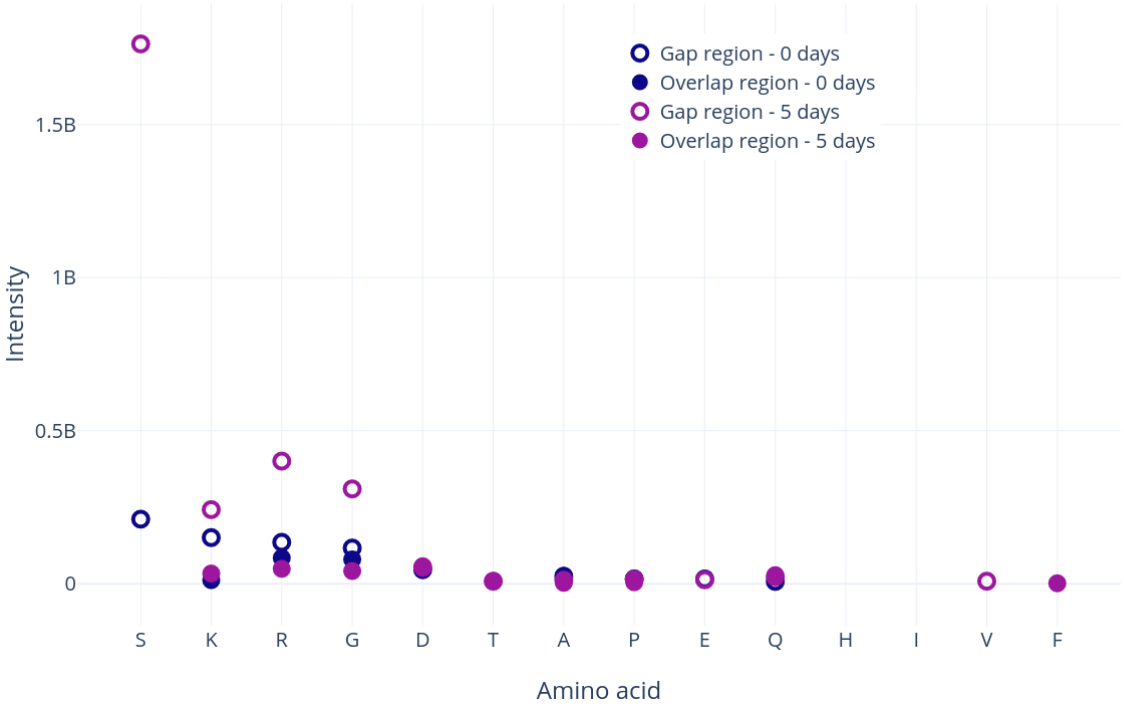


Figure 5.9: *Hydrolysis in the gap and overlap regions of COL1A1 before liming (blue) and after five days of liming (purple) in all skins apart from Skin 11. Gap amino acids are shown as hollow points, while the overlap are represented by filled points.*

This pattern continues at ten days of liming (Figure 5.10), though by 15 days there does not appear to be a substantial difference between the gap and overlap regions. Interestingly, serine only exhibits hydrolysis in the gap region - most of which is down to a tryptic cleavage of lysine at position 1069 (with serine at 1070) at the start of the high intensity hydrolysis region.

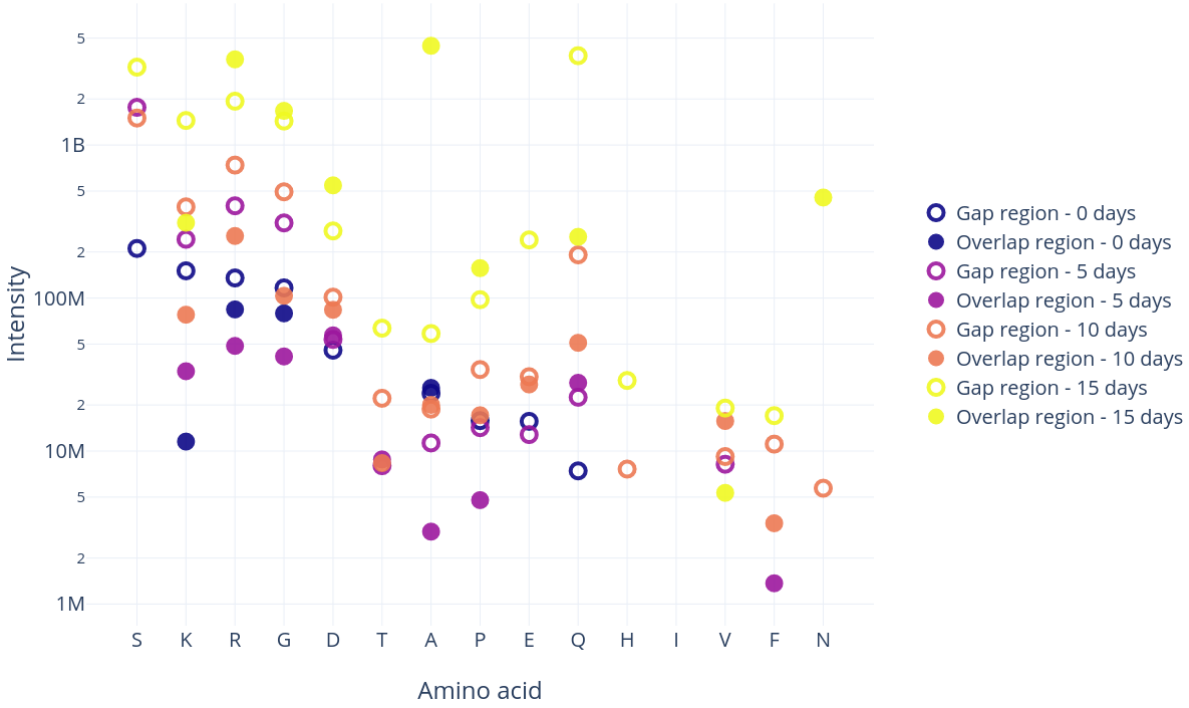


Figure 5.10: *Hydrolysis in the gap and overlap regions of COL1A1 before liming (blue), after five (purple), ten (orange), and 15 (yellow) days of liming in all skins apart from Skin 11. Gap amino acids are shown as hollow points, while the overlap are represented by filled points. The intensity is represented on a logarithmic scale.*

This pattern is less clear in Skin 11 (shown in Figure 5.11). While hydrolysis of the gap and overlap regions in COL1A1 in all the other skins (Figure 5.10) is generally similar within the same amino acid, in Skin 11 there is much less correlation. Across all time points, the gap region seems to exhibit much higher hydrolysis, and many amino acids in the overlap region (particularly arginine, glycine, and tyrosine) exhibit less hydrolysis in Skin 11 before liming than in the other skins. Serine hydrolysis in the gap region is still very high, while, unlike the other skins, there is a single serine residue in the overlap region that undergoes hydrolysis after ten days of liming.

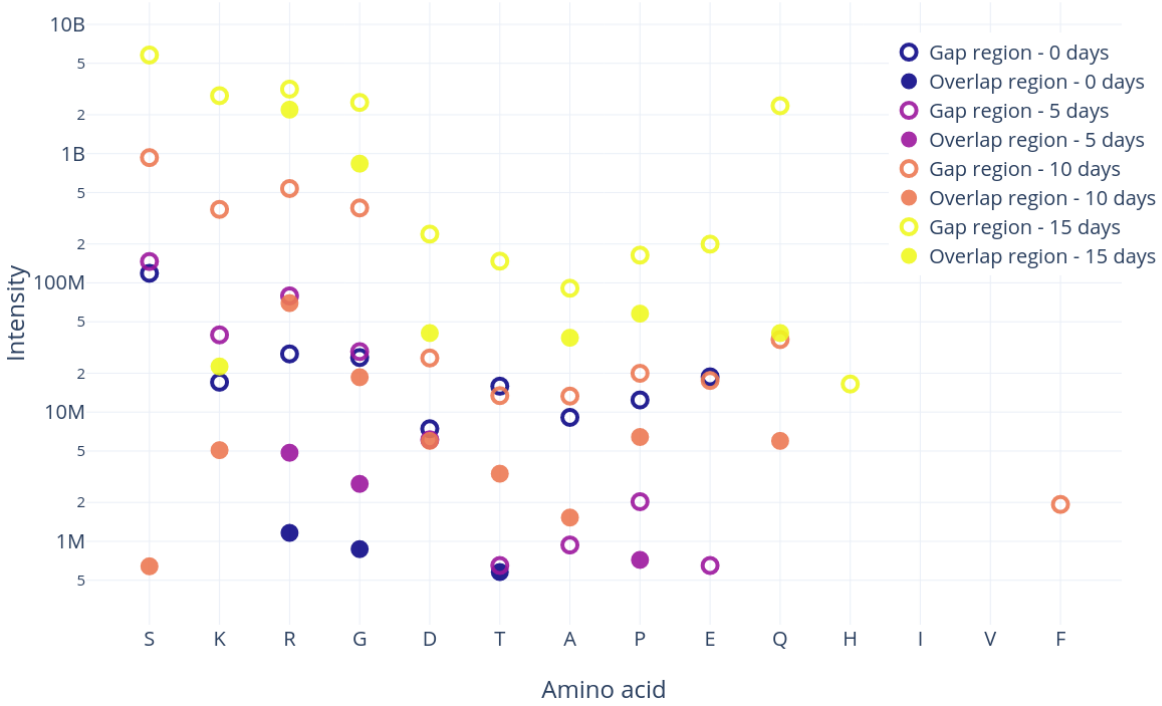


Figure 5.11: *Hydrolysis in the gap and overlap regions of COL1A1 before liming (blue), after five (purple), ten (orange), and 15 (yellow) days of liming in only Skin 11. Gap amino acids are shown as hollow points, while the overlap are represented by filled points. The intensity is represented on a logarithmic scale.*

5.3.1.2.3 High intensity hydrolysis region

Lastly, hydrolysis per amino acid was compared in the high intensity hydrolysis region at 1069 to 1093 of COL1A1 versus all other gap positions (Figure 5.12). Most residues fit the pattern that hydrolysis increases with duration of liming, even in the high intensity hydrolysis region. The only substantial difference is that serine is highly hydrolysed in the high intensity region, but exhibits much less hydrolysis in the rest of the gap regions. This is in agreement with the hypothesis that the hydrolysis between lysine and serine in this region is the driving force behind serine's seemingly high hydrolysis. Isoleucine and valine, though found in the high intensity hydrolysis region, they are not found cleaved in any skin.

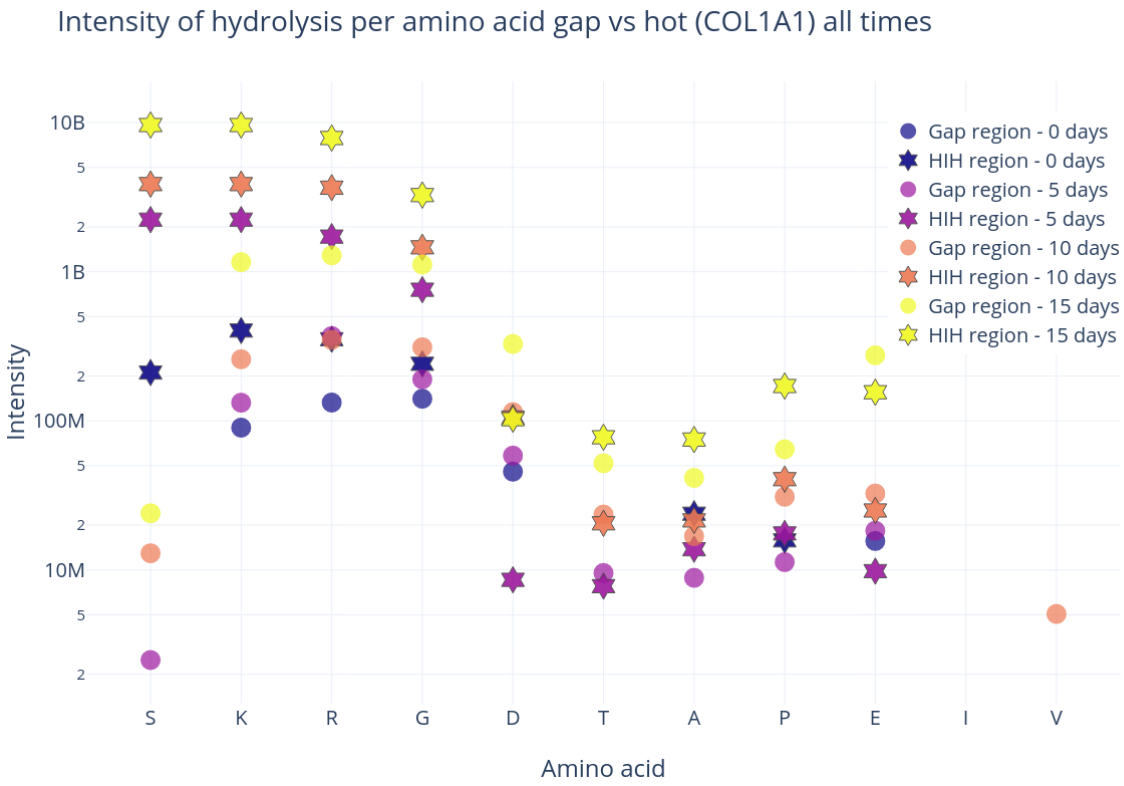


Figure 5.12: *Hydrolysis in the gap and overlap regions of COL1A1 before liming (blue), after five (purple), ten (orange), and 15 (yellow) days of liming in all skins apart from Skin 11. High-intensity region amino acids are represented by stars, while other gap region residues are filled points. The intensity is represented on a logarithmic scale.*

5.3.2 Peptide profiles

Peptide profiles in both COL1A1 and COL1A2 were generated by peptigram (Manguy *et al.*, 2017) for each skin at each time point in order to visualise the coverage and intensity of peptides recovered.

As shown in Figure 5.13, peptide coverage and intensity in COL1A1 increases fairly evenly across each skin and the duration of liming increases. Skin 11, the stillborn lamb, shows higher peptide counts before liming than it does after five days of liming. For all other skins, the width and depth of coverage increases with liming. The high intensity hydrolyses region identified earlier is easily visible at position ~1000 in all skins at all time points, as it has a high depth of coverage and increases in intensity as liming duration increases. The start of the triple helical region is also visible as a high depth region that increases in intensity, at around position 200.

COL1A2, in Figure 5.14, shows a similar trend of growing intensity and depth and width of coverage with liming. Like in COL1A1, Skin 11 shows an increased peptide count before liming than after five days. However unlike COL1A1, there does not appear to be regions with particularly high levels of intensity or depth. The only semi-consistent pattern is that there is normally high intensity at the start of the helical region at around position 100. This is expected, as high intensity peptides in COL1A1 are also detected at the start of the helical region. Moreover, depth does not seem to meaningfully increase with liming as it does in COL1A1. For example, in Skin 15, five, ten, and 15 days of liming exhibits strikingly similar bar heights.

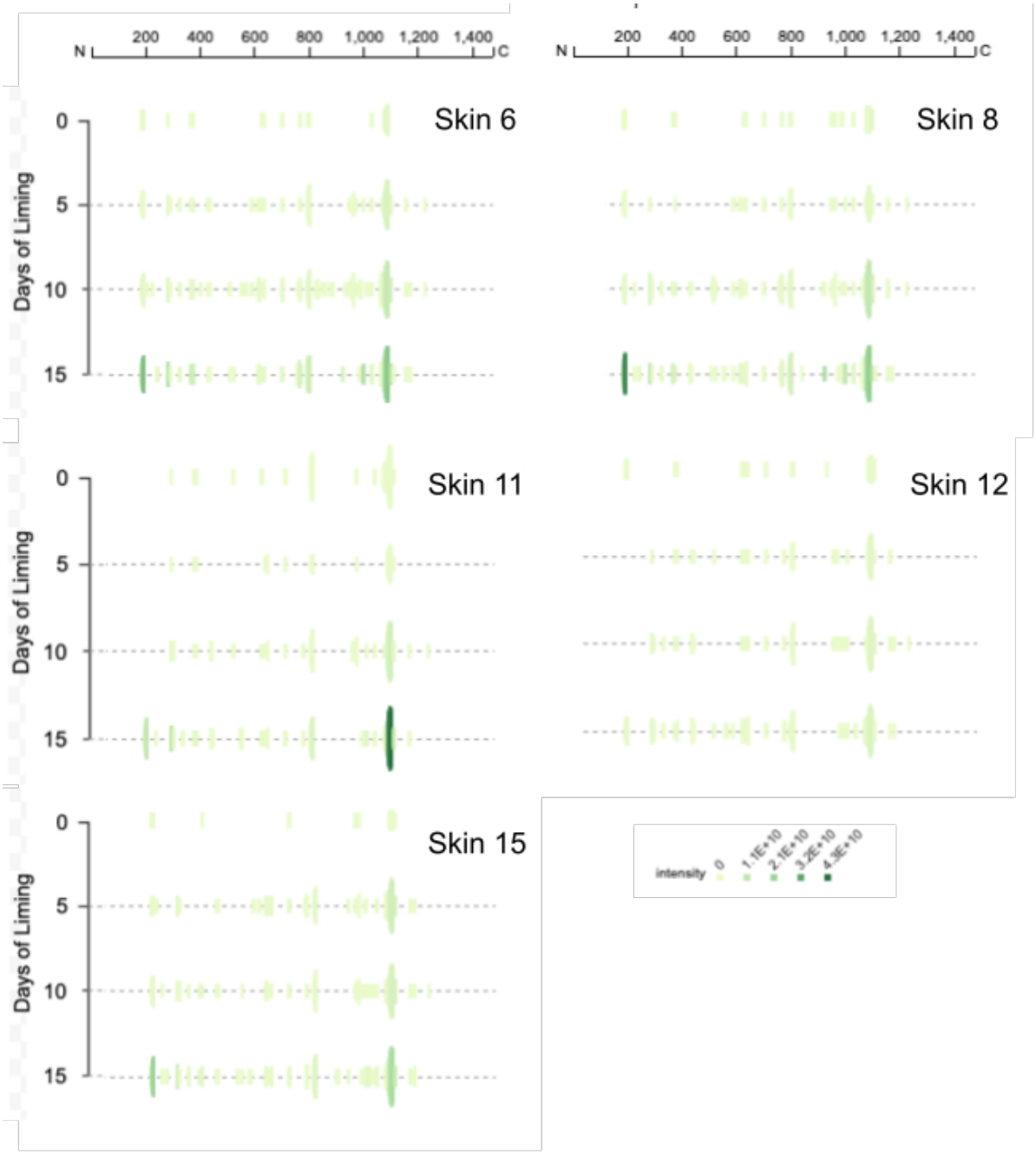


Figure 5.13: *Peptigram showing coverage and intensity of peptides in COL1A1. For each residue, a green bar is drawn if the position is covered by at least one peptide. The height of the bar is proportional to the count of the peptides overlapping the position; while the colour intensity is proportional to the summed ion intensities of peptides overlapping the position - briefly, the darker the green, the more intense the peptide.*

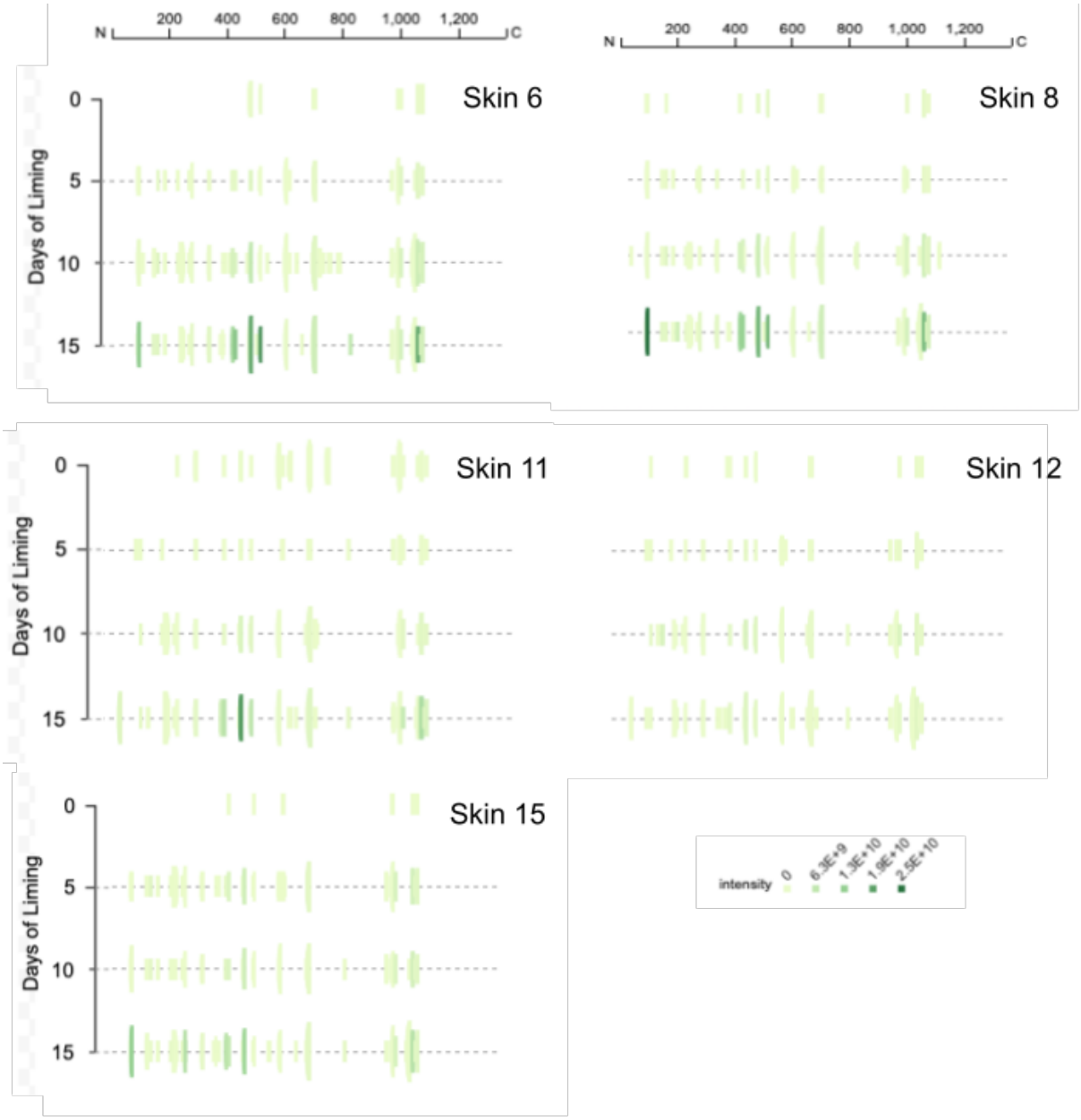


Figure 5.14: *Peptigram showing coverage and intensity of peptides in COL1A2. Note the different scale of the intensity colour between this figure and Figure 5.13. For each residue, a green bar is drawn if the position is covered by at least one peptide. The height of the bar is proportional to the count of the peptides overlapping the position; while the colour intensity is proportional to the summed ion intensities of peptides overlapping the position - briefly, the darker the green, the more intense the peptide.*

5.3.3 Peptide counts

As shown in Figure 5.15, the number of unique peptides increases linearly with days of liming. This correlation across all skins is moderate, with an R^2 of 0.632 and 0.776 for COL1A1 and COL1A2 respectively. Skin 11 can be seen as a clear outlier, with many more peptides before liming than the other skins. Indeed, with Skin 11 removed, the R^2 values for both proteins indicates a strong correlation: 0.901 for COL1A1 and 0.902 in COL1A2.

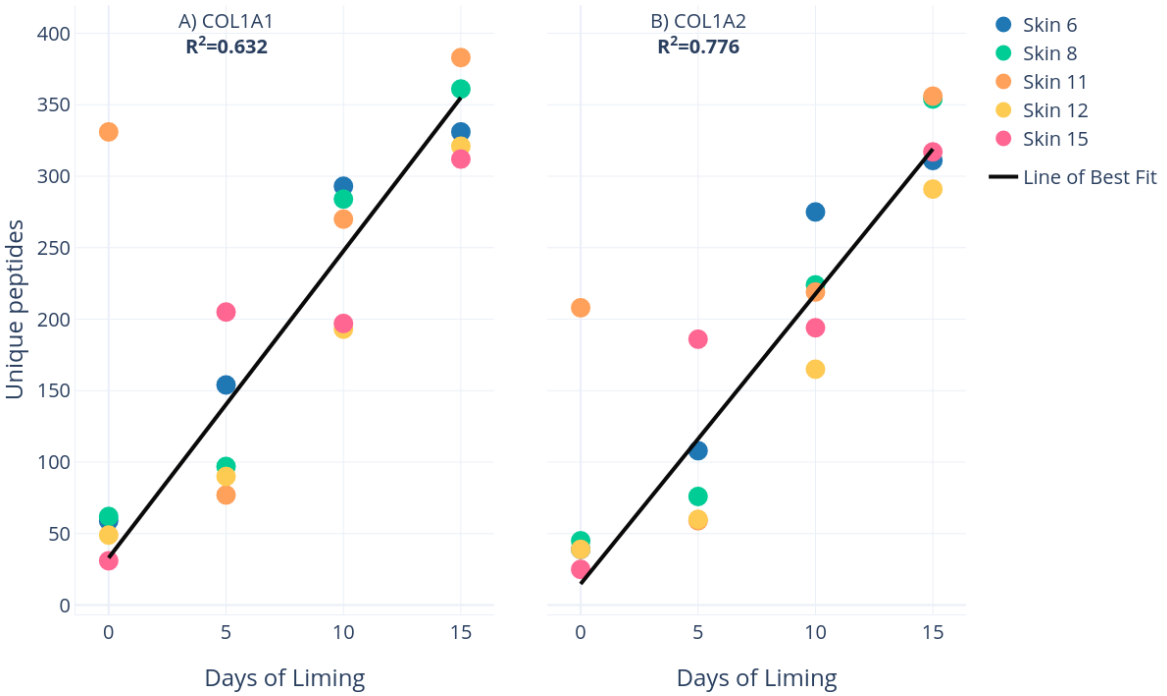


Figure 5.15: *Unique peptide count (from MaxQuant's 'unique + razor protein count' field versus days of liming for all skins.*

5.3.3.1 Gap vs overlap peptide count

There is very little difference between the relative peptide counts of the gap and overlap region as liming progresses (Figure 5.16). Even in Skin 11, where the peptide count does not follow the same linear trend as the other skins, there is no meaningful difference between the gap and overlap regions. Without Skin 11, the R^2 value for the other skins, in both the gap and overlap regions in COL1A1 and COL1A2 is very strong, as shown in Table 5.5. However, due to its abundance of unique peptides before liming, the R^2 values for Skin 11 show either no correlation (COL1A1) or an extremely weak correlation (COL1A2).

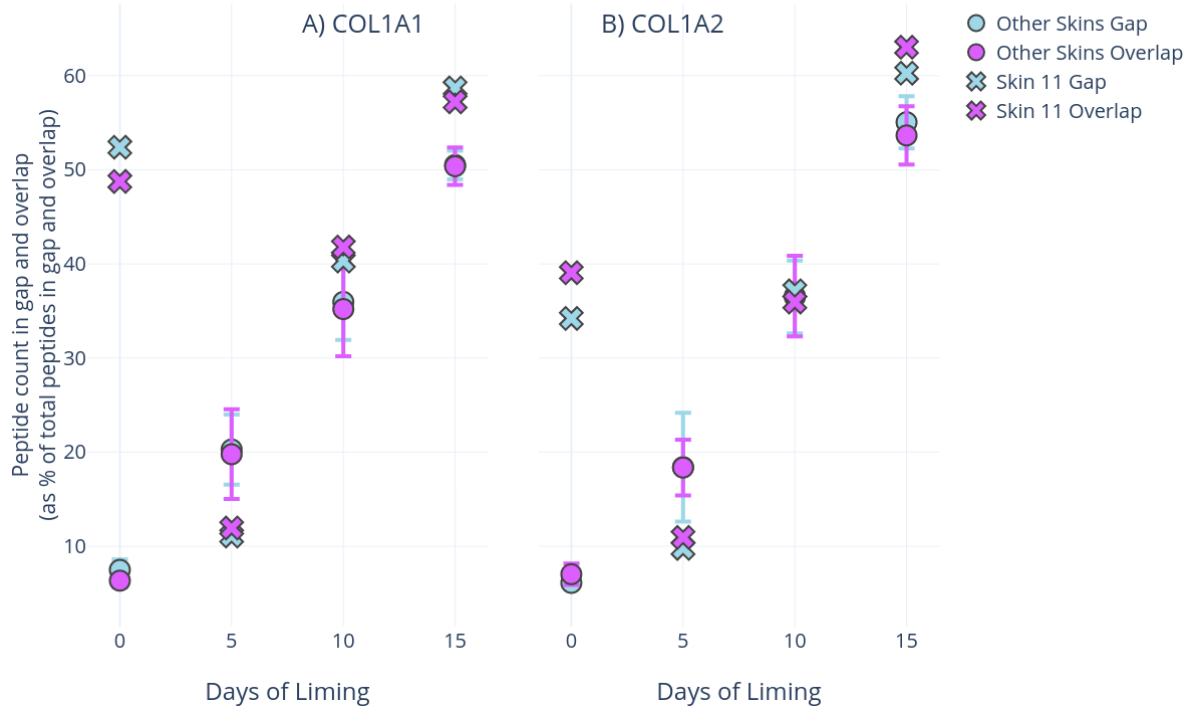


Figure 5.16: *Peptide count in the gap and overlap regions during liming. Peptide count is normalised between regions to correct for the difference in total length of the regions. Skin 11 is shown as a cross, while all other skins are shown as filled points. For the combined "other" skins, standard deviation is represented by error bars.*

	Gap		Overlap	
	Col1a1	**Col1a2**	**Col1a1**	**Col1a2**
Skin 11	0.087	0.435	0.0131	0.345
Other Skins	0.988	0.992	0.971	0.998

Table 5.5: *R^2 values for the unique peptide count in the gap and overlap regions of COL1A1 and COL1A2 for Skin 11 (row two) and all other skins combined (row three).*

5.3.3.2 High intensity hydrolysis region peptide count

As discussed above, and seen in Figures 5.16 and 5.17, the unique peptide count in the gap region increases linearly with days spent in lime. However, in the high intensity hydrolysis region (between 1069 and 1093 in COL1A1), this relationship is exponential (R^2=0.998). As expected, Skin 11 does not follow this trend, as it has an extremely high unique peptide count in the high intensity region even before liming.

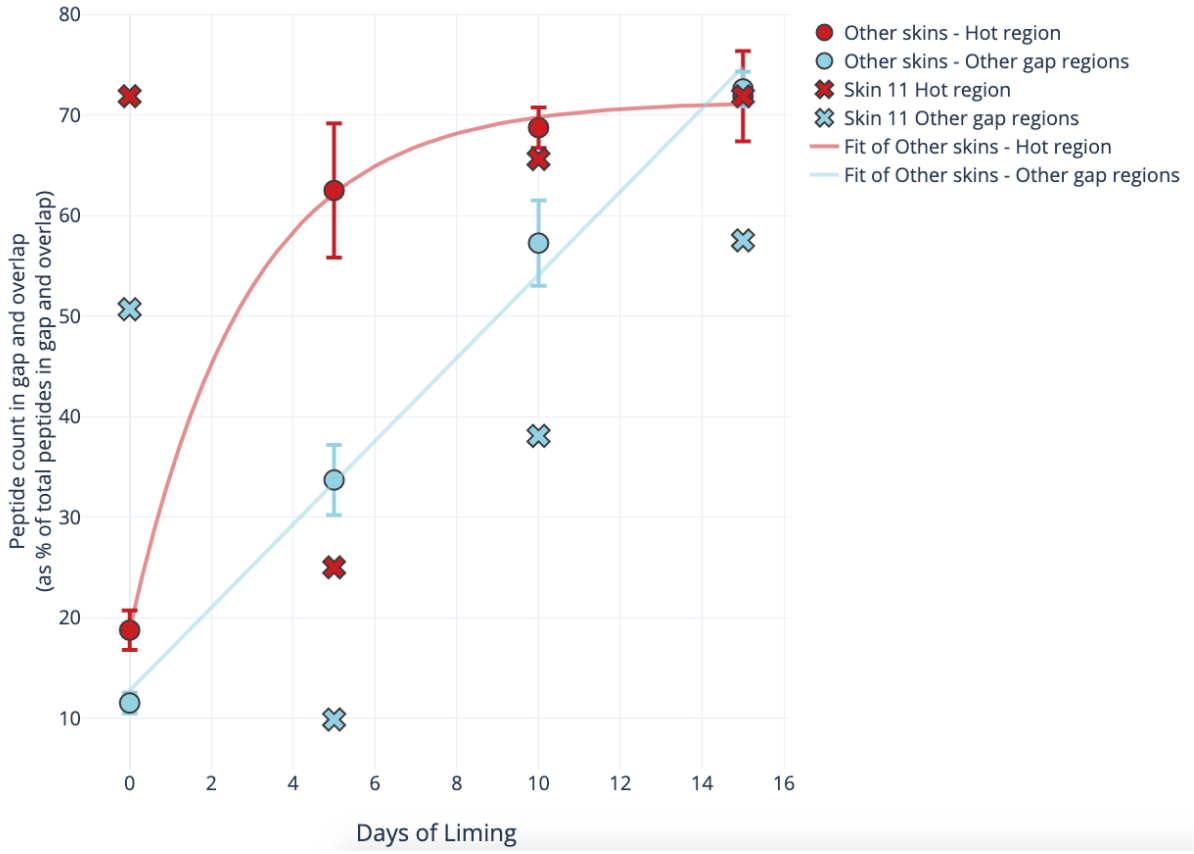


Figure 5.17: *Peptide count in the high intensity hydrolysis region of COL1A1 (1069-1093) and the other gap regions during liming. Peptide count is normalised between regions to correct for the difference in total length of the regions. Skin 11 is shown as a cross, while the average of all other skins are shown as filled points. Lines of best fit for the high intensity region are exponential, while the other gap regions are linear.*

5.3.4 Depth of coverage

The average depth of coverage across both COL1A1 and COL1A2 was calculated in order to investigate if there was differential coverage, either across the proteins or between the liming time points. As shown in Figure 5.18, coverage is very variable across both COL1A1 and COL1A2, with many regions of high coverage, and some areas of either very low, or no coverage. In general, the helical region (186 - 1183 in COL1A1 and 87 - 1103 in COL1A2) is well covered, with at least one peptide covering every position. Surprisingly, there is a fairly large area of no coverage just after the high intensity region of hydrolysis, spanning 44 residues, from position 1113 to 1147.

5.3.4.1 Intensity versus coverage

Another unexpected result is that the region of high intensity hydrolysis is not the most covered part of the protein. Since, as seen in Figure 5.17, it has a much higher unique peptide count than other gap regions, it would follow that it would be the most covered region. However, this is not the case, as shown in Figure 5.18.A, there are three “peaks” of higher average depth of coverage.

As seen in Figure 5.19, the relationship between the summed intensity of peptides at a position and the coverage at the position are not related.

Figure 5.18: *Stacked area plot of average depth of coverage per position in COL1A1 (A) and COL1A2. The high intensity hydrolysis region at 1069 to 1093 in COL1A1 is marked by an arrow. Note that there is more hydrolysis on the N-terminal side.*

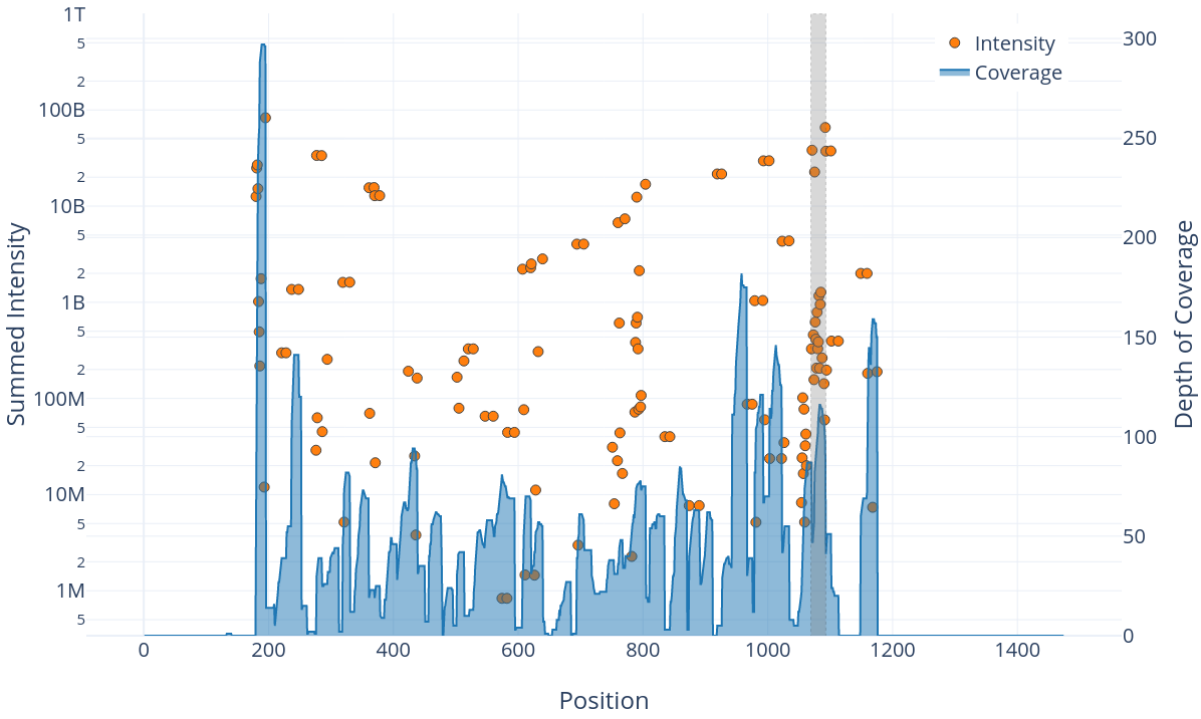


Figure 5.19: *Summed intensity of each position in all skins apart from Skin 11 (on a logarithmic scale) versus the summed depth of coverage at each position. The high intensity hydrolysis region is shown as a grey area.*

5.3.4.2 Skin 11 coverage

Skin 11 also has a notably different depth of coverage profile to the other skins. Figure 5.20 shows the depth of coverage of COL1A1 in Skin 11 before liming versus Skin 6. While the most deeply covered points in Skin 6 reach around 5x, many regions in Skin 11 reach 10x. Indeed, the average coverage for positions that are covered by at least one peptide in Skin 11 is 9.7x, while this is just 2.3x for Skin 6.

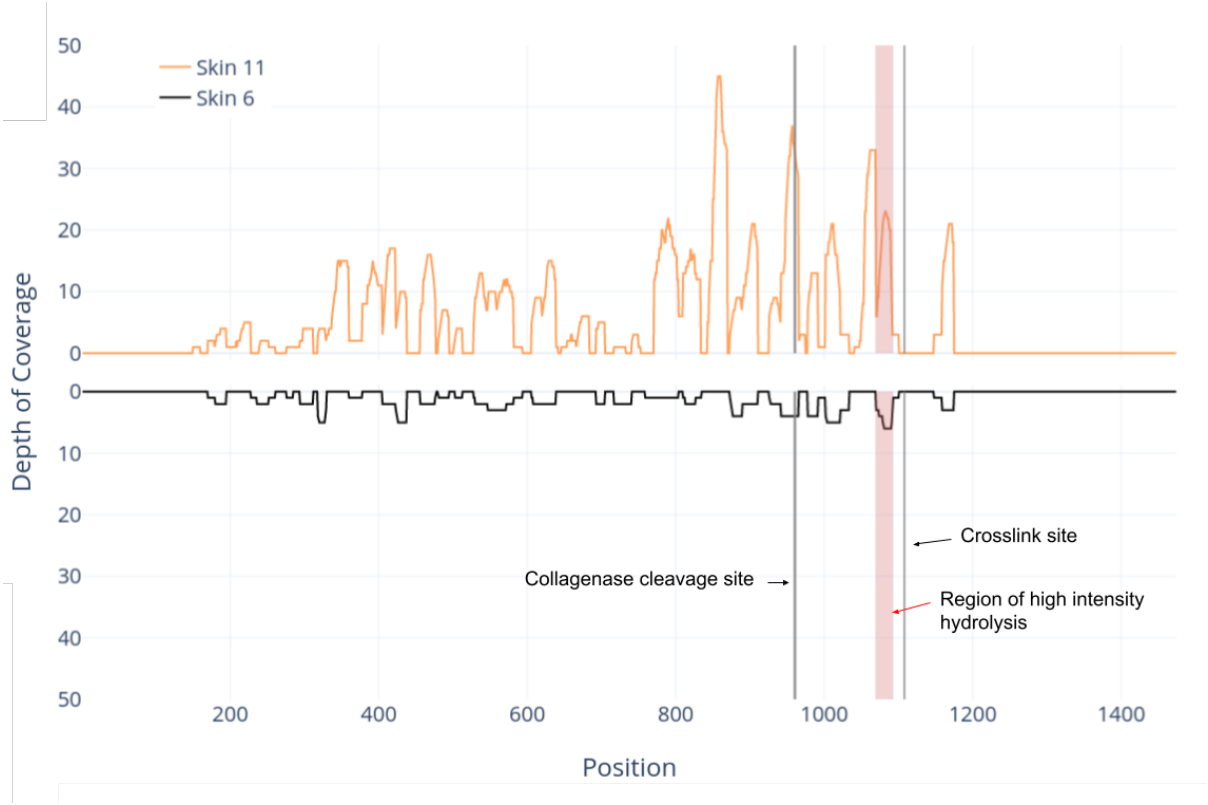


Figure 5.20: *Average depth of coverage for Skin 11 (orange) versus Skin 6(black) in COL1A1 before liming.*

5.3.4.3 Gap vs overlap depth

As with the peptide count (discussed in section 5.3.3) and shown in Figure 5.21, the depth of coverage increases linearly (R^2 values all above 0.970, as seen in Table 5.6) with days of liming in all skins apart from Skin 11, and there is no clear difference between the coverage of the gap and the overlap regions. However, due to its higher average depth of coverage before liming, Skin 11 does not have as significant a correlation as the other skins. Nevertheless, it does show a moderate correlation in COL1A1 (Figure 5.21.B), with R^2 values for the gap and overlap regions of 0.704 and 0.617 respectively. This is due to the coverage before liming being markedly lower than at the same time point in COL1A1, likely because of the relatively higher abundance of COL1A1.

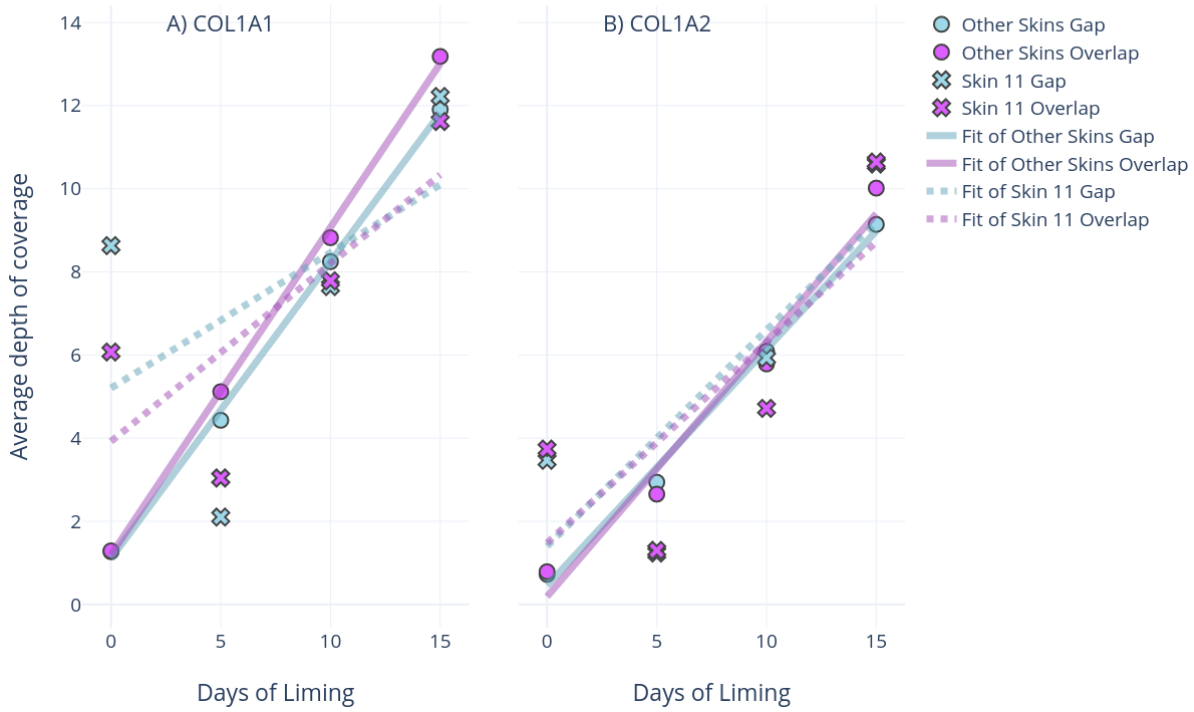


Figure 5.21: *Average depth of coverage in the gap and overlap regions during liming. Skin 11 is shown as a cross, while all other skins are shown as filled points.*

	Gap		Overlap	
	Col1a1	**Col1a2**	**Col1a1**	**Col1a2**
Skin 11	0.253	0.704	0.595	0.617
Other Skins	0.999	0.994	0.999	0.971

Table 5.6: *R^2 values for the average depth of coverage in the gap and overlap regions of COL1A1 and COL1A2 for Skin 11 (row two) and all other skins combined (row three).*

5.3.4.2 High intensity hydrolysis region coverage in Skin 11

As shown in Figure 5.18.A, although the region of high intensity hydrolysis in COL1A1 does correspond to a particularly deep portion of the protein, it is not the most deeply covered region at any time point. Figure 5.22 shows the difference between all skins and Skin 11 in this region. It is clear that Skin 11 has deeper coverage in this region before liming, with an average of 16x; compared to an average of 3x across the other skins. Unlike the other skins, the coverage does not increase substantially after five days of liming. Lastly, at ten and 15 days of liming all skins have a similar level of total depth of coverage, and increase of coverage in this region.

Figure 5.22: *Average depth of coverage in the high intensity hydrolysis region in COL1A1 (from 1069 - 1093) in A) Skin 11 and B) the average of all other skins at each liming time point. Frames C) and D) show that after liming the skins are extremely similar.*

5.3.5 Predicted melting point of collagen type 1

Periskov *et al* (2005) predicted the melting temperature of each sequence position in COL1A1 and COL1A2 from their sequence. The relationship between this melting point and hydrolysis was investigated to determine whether this could be the reason behind why some regions are more susceptible to hydrolysis. It could be expected that regions with low melting temperatures would be the most susceptible to hydrolysis (an example of which is shown by the trend line in Figure 5.23). In both proteins (Figure 5.23), most positions in the helical region have a very similar melting point, and as such there is not much differential hydrolysis, though in both proteins and across the range of temperatures it is clear to see that many 15 day liming points have the highest intensity, while the before liming samples have the lowest.

In COL1A1 (Figure 5.23.A), there seem to be three melting temperatures that only hydrolyse after ten or (more often) 15 days of liming. These correspond to 184 V, 185 P, and 187 P, which are the amino acids right at the start of the triple helical region (which starts at 186 G). In COL1A2 (Figure 5.23.B), there is only one point with a clearly lower temperature than the rest, and it seems to hydrolyse relatively evenly across all time points and skins. This position represents a K at position 187.

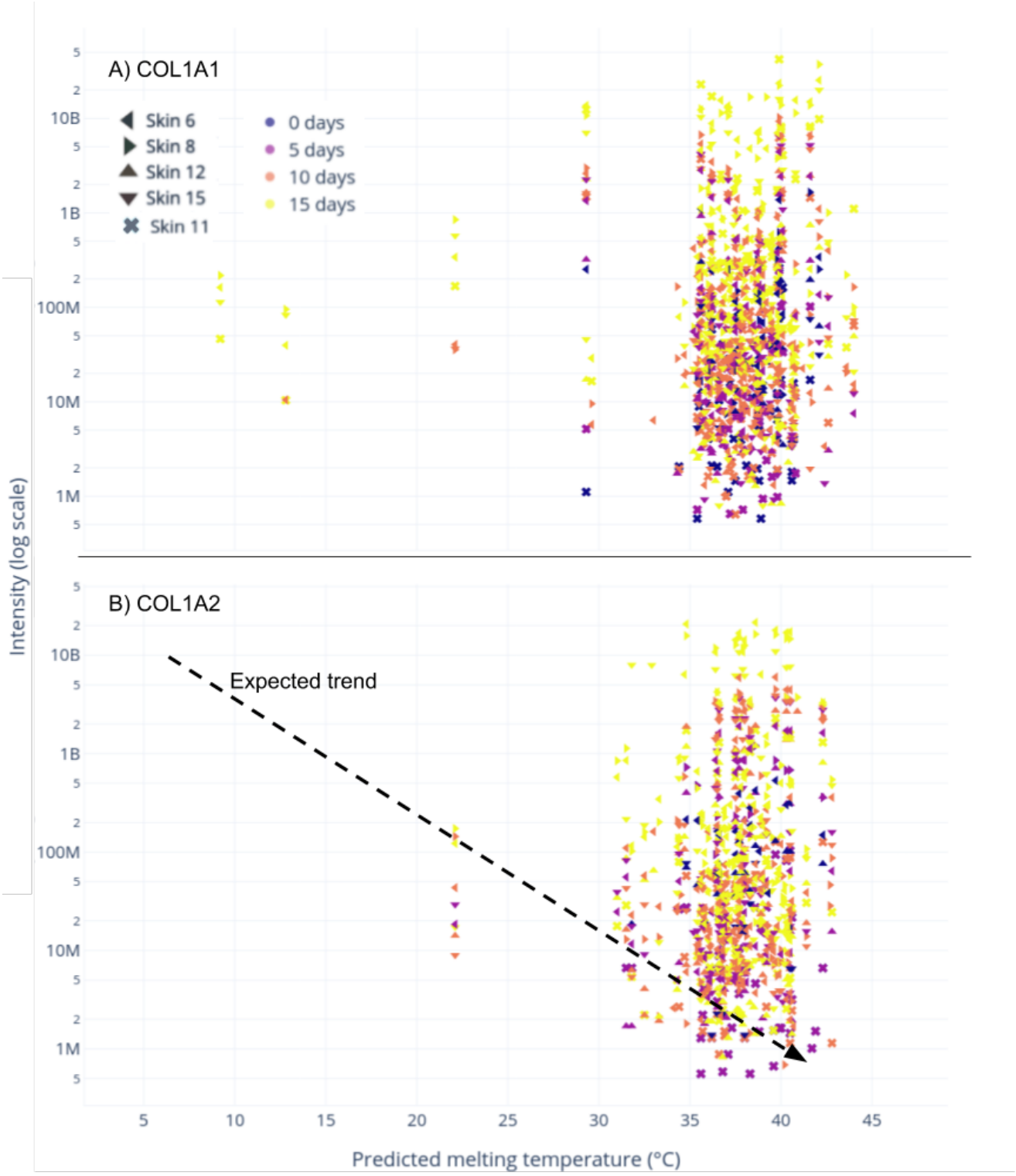


Figure 5.23: *Hydrolysis versus melting temperature of the cleaved amino acid in A) COL1A1 and B) COL1A2.*

5.3.6.1 Regional melting points

The melting points across the different regions of COL1A1 were plotted in Figure 5.24 in order to ascertain whether there is a substantial difference in melting temperature between the regions. Most of the temperatures are very similar across the three regions, though the overlap region has more outliers, due to it encompassing the start and end of the triple helical part of the chain. The high intensity hydrolysis region has no such outliers, and has a slightly higher mean melting point than the other regions, though all are within the margin of error.

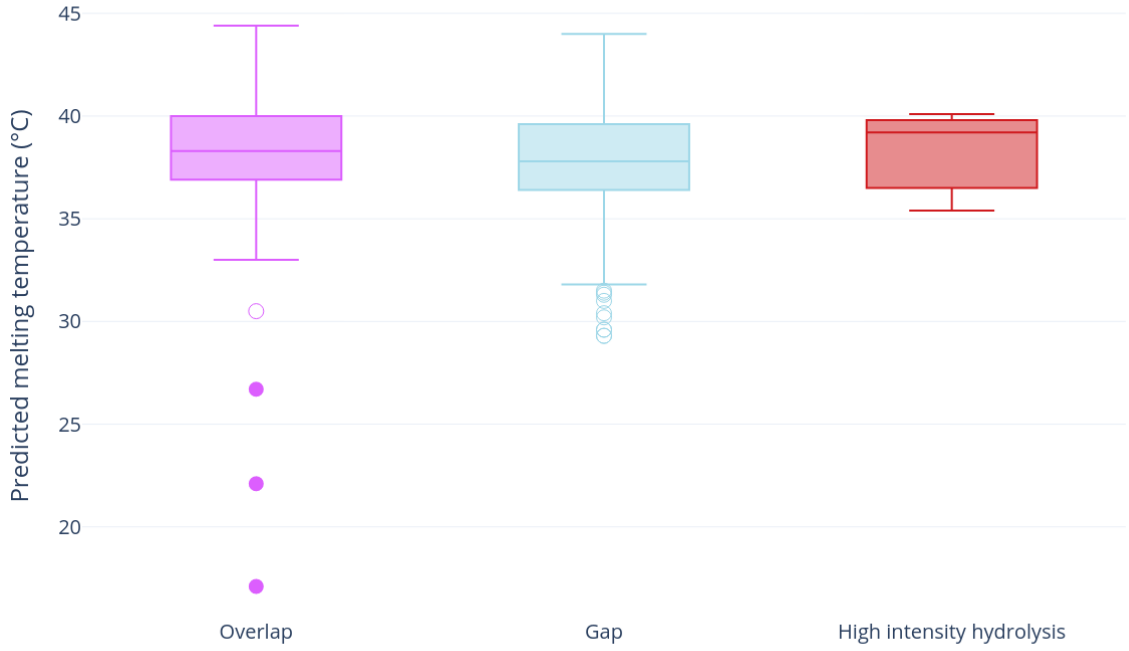


Figure 5.24: *Box plot of predicted melting temperatures of the gap, overlap, and high intensity hydrolysis (1069 - 1093) in COL1A1. Outliers are shown as filled points, while suspected outliers are represented by hollow points.*

5.3.6 Degree of conservation

In order to analyse the variability of COL1A1 and COL1A2 we compared the *Ovis aries* collagen sequences to that of 118 other mammals and quantified the number of observed variations at each position. There are positions in COL1A1 that have up to seven different variants (though only six found hydrolysed in this data) and this is up to eight in COL1A2 (though only up to seven

found hydrolysed). It could be expected that the most variable - therefore the least conserved - positions would be most susceptible to hydrolysis. As collagens' extraordinary strength is conferred by its amino acid composition (as discussed above), any variation from which might weaken the secondary structure.

If the number of different amino acids at each position had no effect on the hydrolysis, we would expect that the intensity of hydrolysis at each number of variations (also known as single amino-acid polymorphisms, or SAPs) increases fairly evenly with liming. This is largely the case for COL1A1 in the non-outlier skins (Skins 6, 8, 12, and 15 - shown in Figure 5.25.A). Sometimes, this is not the case and hydrolysis is higher before liming than after five days. This is the most extreme at hydrolysis between two and five SAPs, where the intensity before liming is very similar to the intensity after 15 days. There are only five hydrolytic events in COL1A1 between sites with 2 and 5 SAPs, and four of them are not assigned intensities. The fifth is responsible for this pattern, and it is the cleavage between 1090 A (5 SAPs) and 1091 R (2 SAPs), which is towards the end of the region of intense hydrolysis. As shown in Figure 5.4, this bond hydrolyses early but does not grow in intensity.

Figure 5.26 shows the hydrolytic events between residues with different numbers of SAPs at each time point. It is clear that most of the hydrolysis is between the conserved sites (one variant - one SAP), but by 15 days of liming hydrolysis is spread more evenly between positions with between one and four variants.

COL1A2, shown in Figure 5.25.B exhibits generally the same pattern of hydrolysis growing with liming duration. However, there is never hydrolysis before liming at cleavages between two positions which both have three or more variants. There are 22 such positions, spanning the whole COL1A2 chain including the gap and overlap regions, and there does not seem to be any structural reason for this.

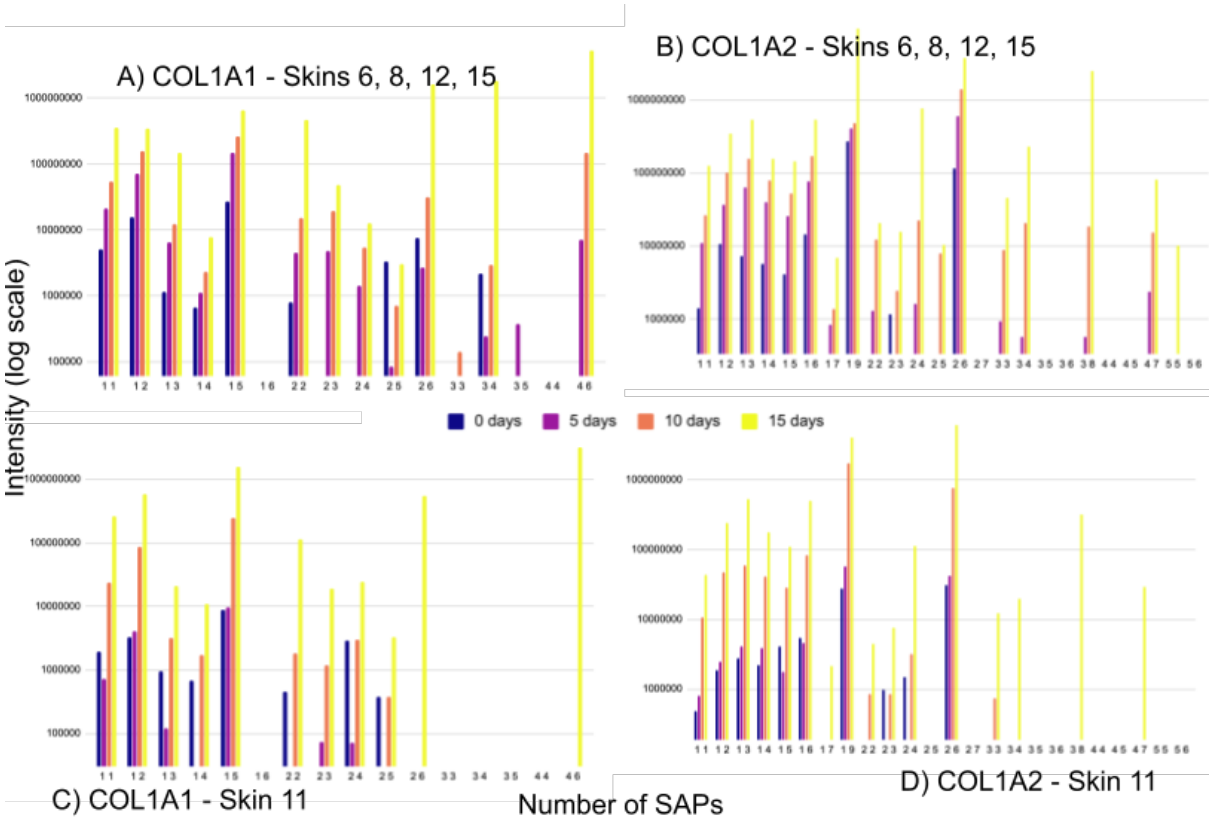


Figure 5.25: *Average hydrolysis in COL1A1 (A and C) and COL1A2 (B and D) in Skin 11 (C and D) and the average of all other skins (A and B) by how many single amino acid polymorphisms exist at each site. Numbers on the x-axis represent the number of SAPs at the sites directly before and after the hydrolysis.*

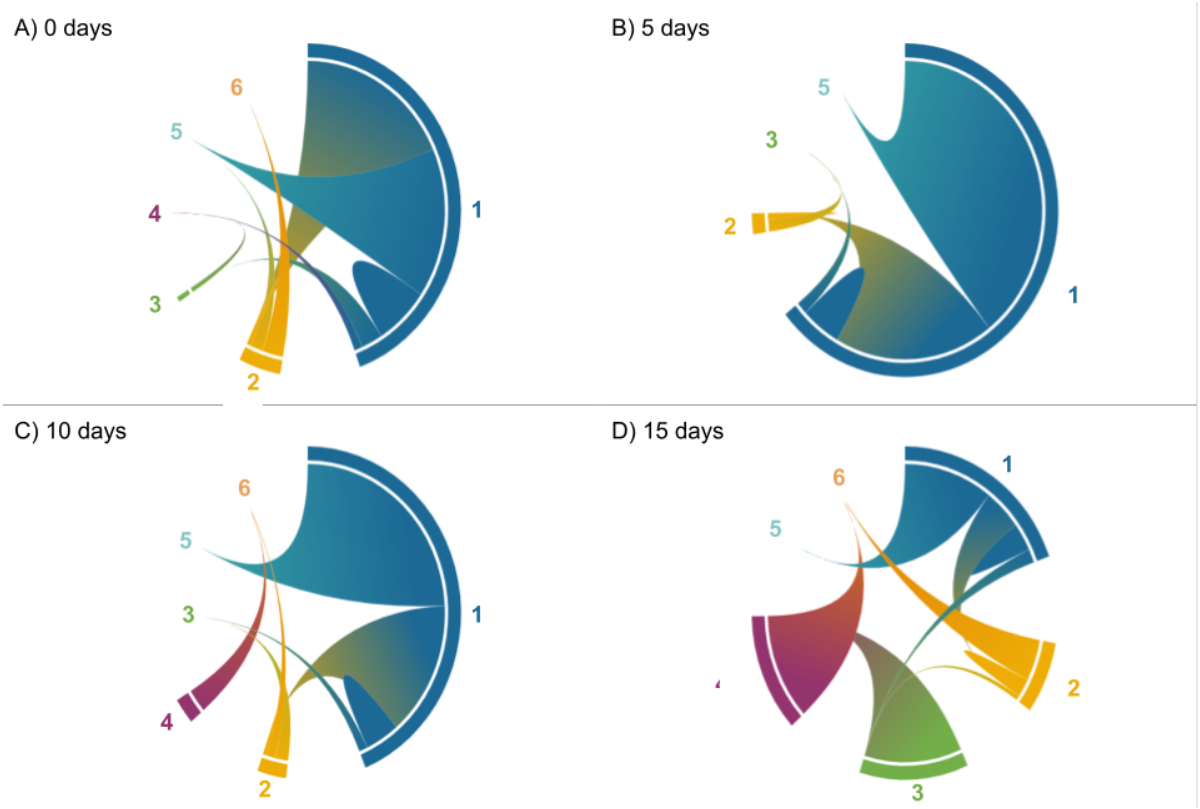


Figure 5.26: *Chord diagram of hydrolysis between sites by number of SAPs in all non-outlier skins (excluding Skin 11) in COL1A1 at all liming durations.*

As expected, Skin 11 portrays slightly different patterns. In both COL1A2 (Figure 5.25.C) and COL1A2 (Figure 5.25.D) there are combinations with no intensity after five days of liming, which fits with Figures 5.13 and 5.14, which show that after five days of liming Skin 11 contains fewer peptides than before liming. Additionally, in both COL1A1 and COL1A2 Skin 11 only exhibits hydrolysis after 15 days of liming between residues which both have high variance. For example, events between two and six variants undergo a high amount of hydrolysis in the other skins (Figure 5.25.A), but in Skin 11 this is only detected after 15 days of liming. The positions most responsible for this pattern are 179 S (2 SAPs) and 180 T (6 SAPs), which are in the propeptide region, and hydrolyse at every time point in the other skins.

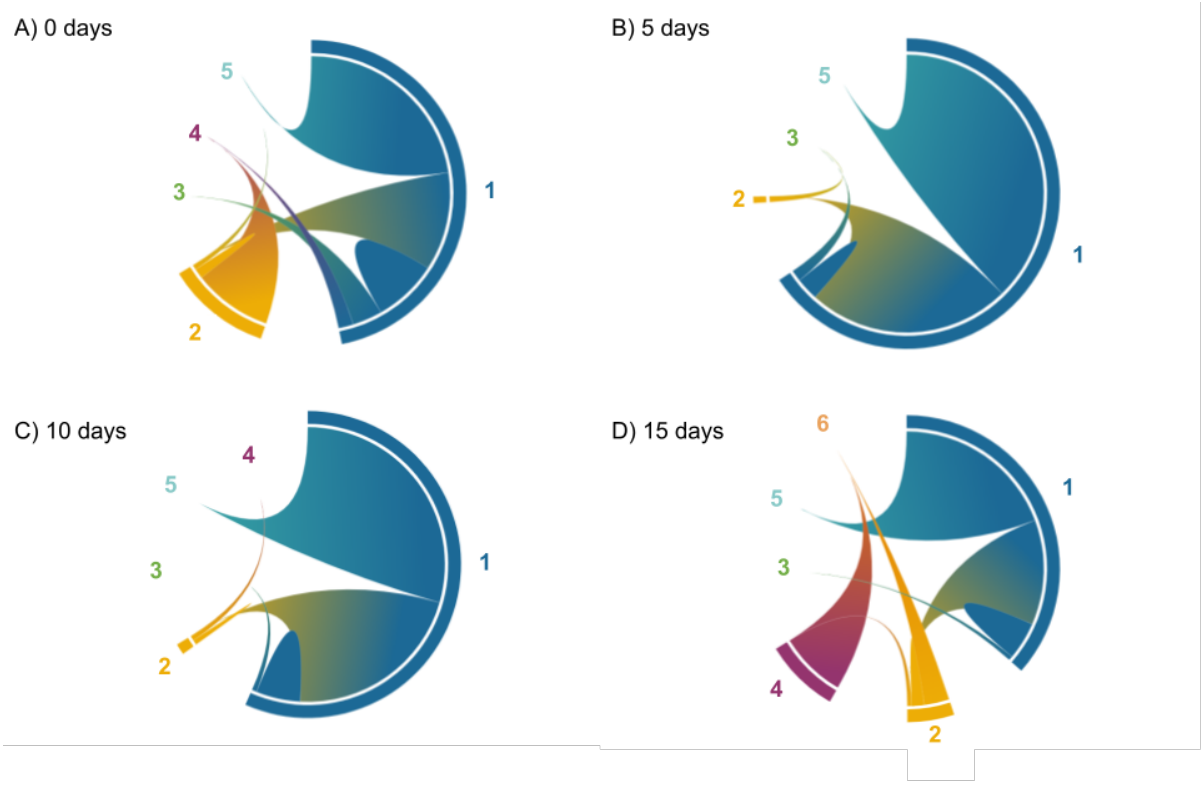


Figure 5.27: *Chord diagram of hydrolysis between sites by number of SAPs in Skin 11 in COL1A1 at all liming durations.*

As shown in Figure 5.27, Skin 11 also experiences most hydrolysis between sites with just one variant, which then decreases to a more even spread by 15 days of liming. Unlike the other skins (Figure 5.26), positions with three variants are very rarely seen in Skin 11, even after 15 days of liming (Figure 5.27.D)

5.3.6.1 Regional conservedness

In order to investigate whether the gap, overlap, or intense hydrolysis region are more variable, the number of SAPs in each region was calculated, as shown in Figure 5.28. The overlap and gap region are very similar, while the region of high intensity hydrolysis has markedly more variability, which could be the reason behind its high fragility.

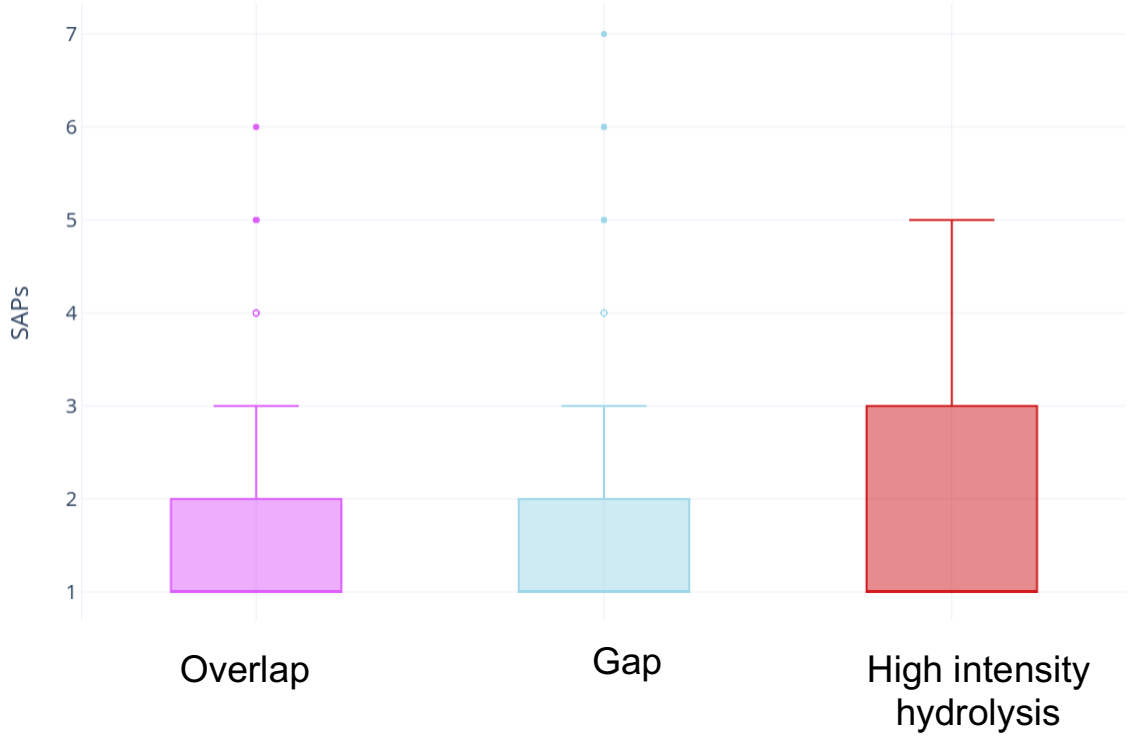


Figure 5.28: *Conservedness of overlap, gap, and high intensity hydrolysis region (1069 - 1093) in COL1A1. Outliers are shown as filled points, while suspected outliers are represented by hollow points.*

5.3.7 Conservedness vs melting temperature

As shown in Figure 5.29, there is no clear relationship between the number of variants at a given position and the predicted melting point at the position.

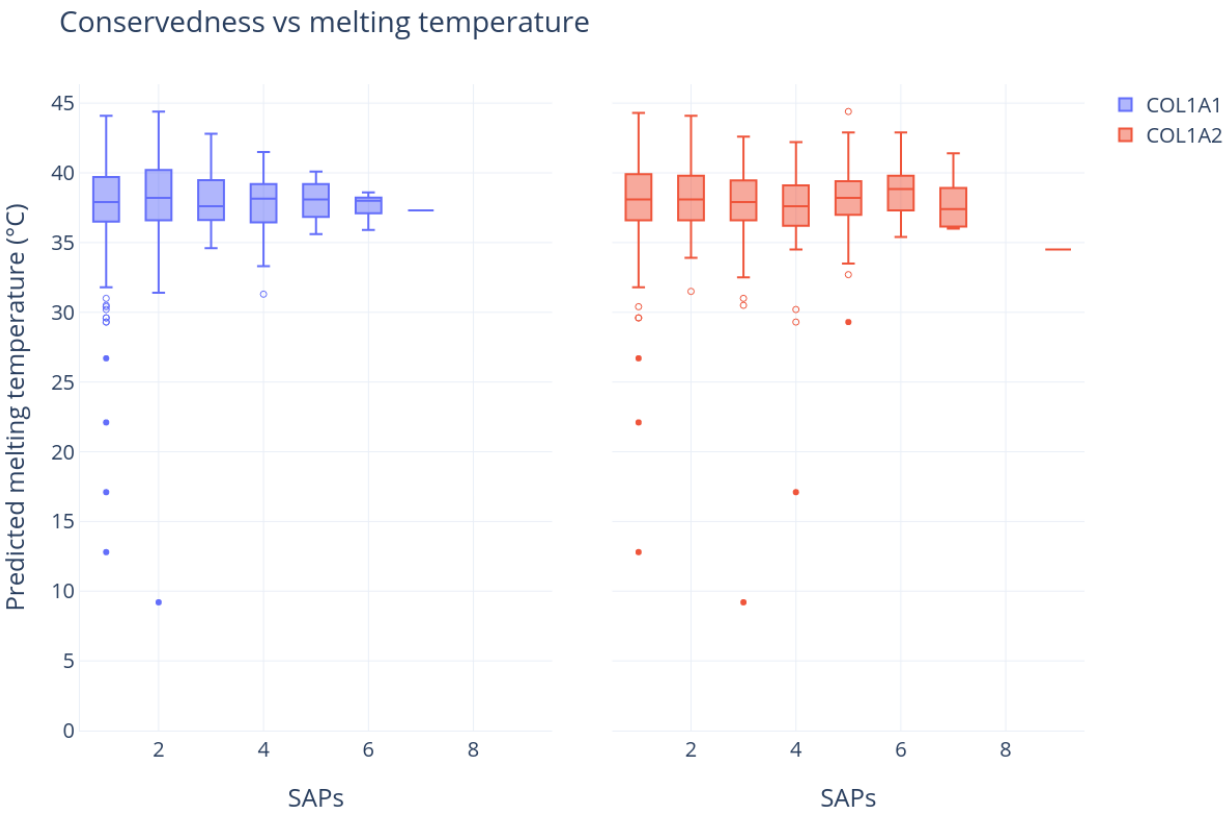


Figure 5.29: *Relationship between number of variants at a given position and the predicted melting temperature. Outliers are shown as filled points, while suspected outliers are represented by hollow points.*

5.3.8 Hydrophobicity

In order to assess whether there was a clear bias between hydrophobic and hydrophilic amino acids, the hydrolysis versus the hydropathicity was plotted in Figure 5.30. It would be expected that there is more (and earlier) hydrolysis between residues that are both hydrophilic (therefore have a low number on the x-axis). This seems to generally be the case, as, especially before liming, very hydrophilic positions seem to exhibit a generally higher cleavage intensity. However, this pattern is very weak.

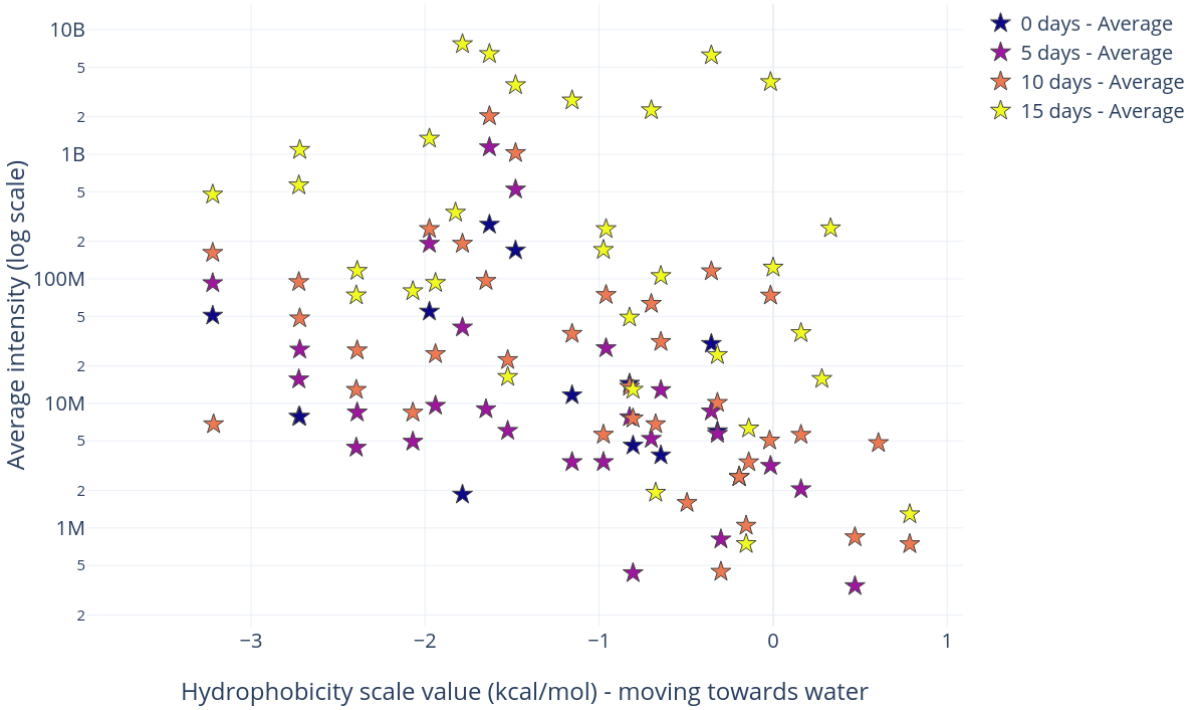


Figure 5.30: *Average hydrophobicity (Inverse Octanol scale - low value numbers are hydrophilic) for each amino acid pair and the intensity of the cleavage between them.*

5.3.9 Deamidation

Deamidation was calculated for COL1A1 and COL1A2 using deamiDATE (Ramsøe *et al.*, 2020). Figure 5.31 shows the bulk deamidation of both asparagine and glutamine in all the non-outlier skins (i.e. excluding Skin 11) at each time point. For glutamine (Figure 5.31.A and B), there is very little deamidation before liming in both COL1A1 (A) and COL1A2 (B), and deamidation gradually increases with liming duration. In asparagine (Figure 5.31.C and D) there is more deamidation before liming in both proteins (particularly apparent in COL1A1), and it too increases, until at 15 days of liming most of the asparagine residues in both proteins are deamidated.

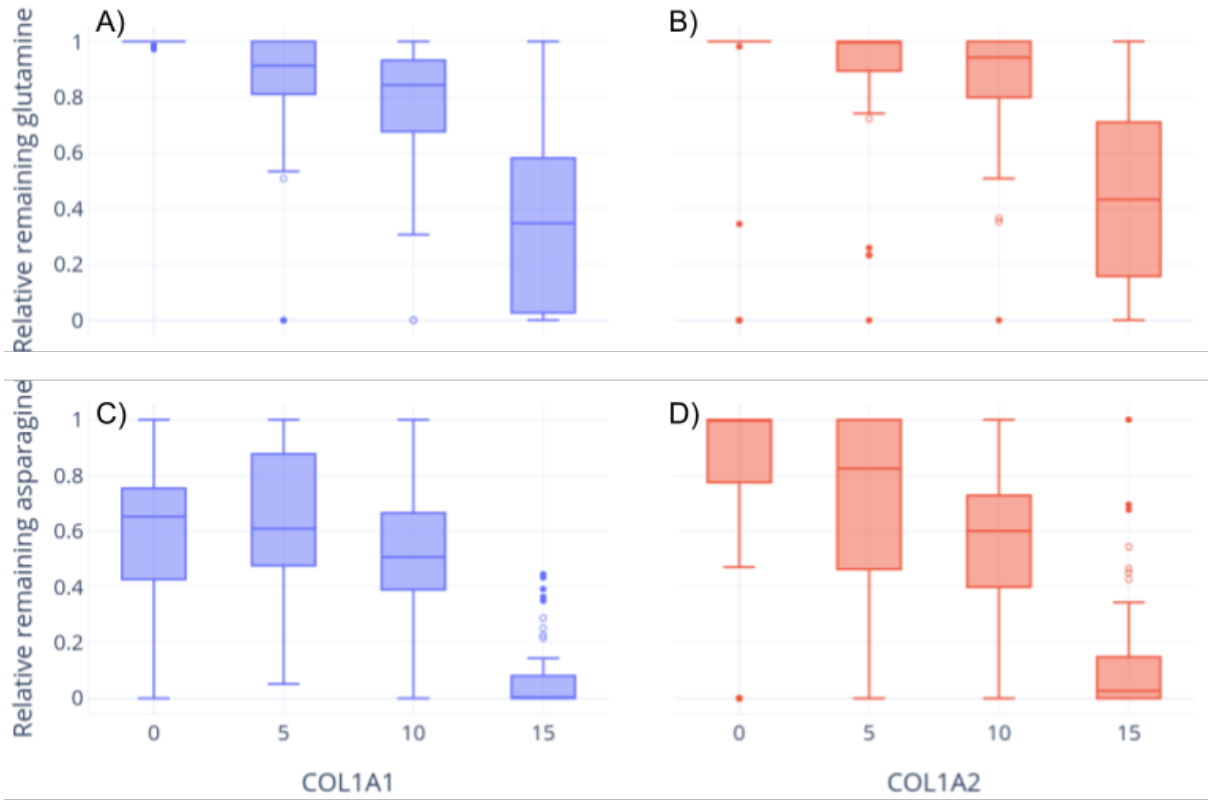


Figure 5.31: *Bulk deamidation of all non-outlier skins (Skins 6, 8, 12, and 15) for glutamine and asparagine in both COL1A1 (blue) and COL1A2 (red). The y-axis represents the relative remaining deaminating amino acid - that is, at 1 there is no deamidation, and 0 represents total deamidation. Filled points represent outliers, while hollow points show suspected outliers. A and B show glutamine deamidation in COL1A1 and COL1A2 respectively, while C and D show asparagine deamidation.*

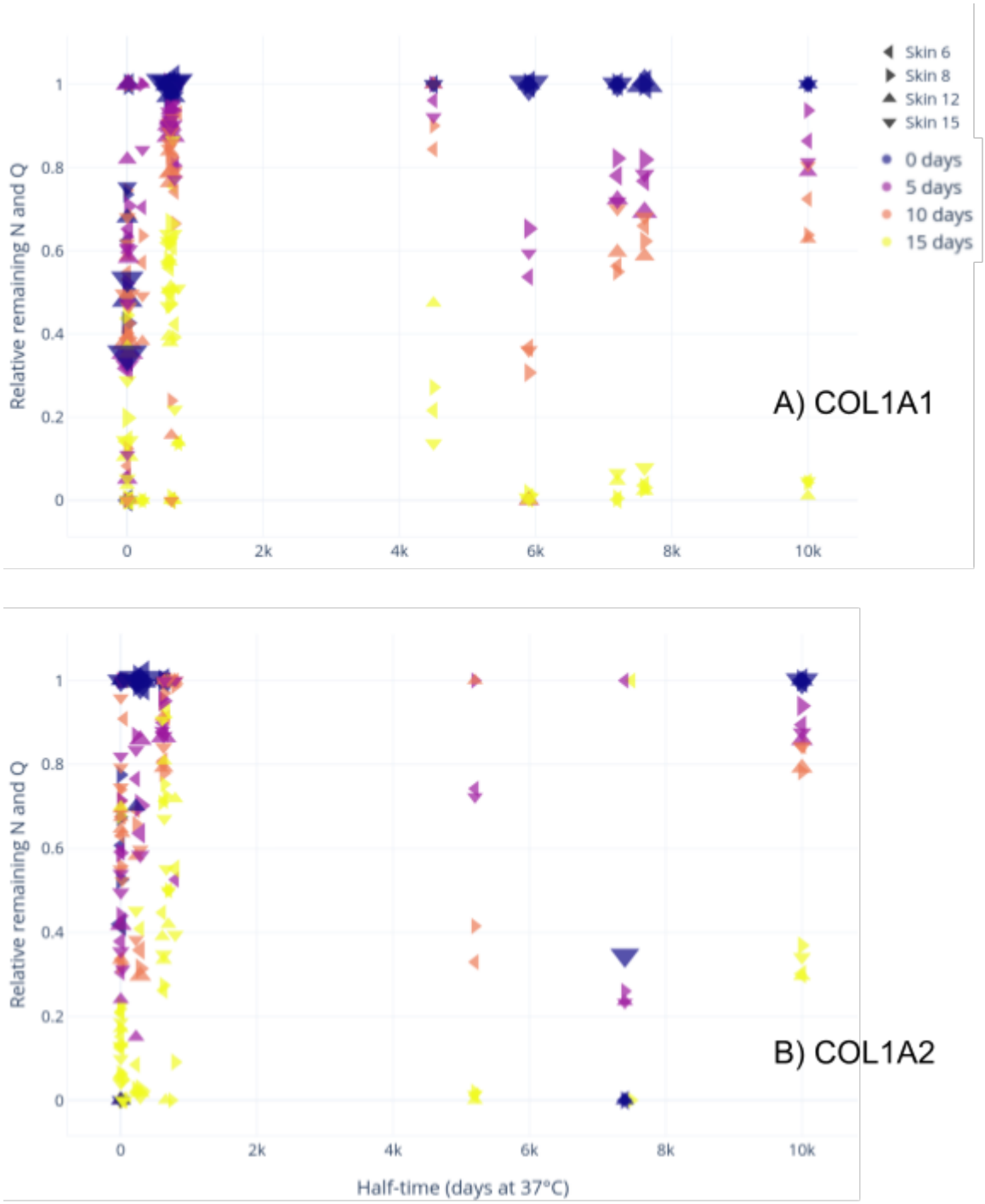


Figure 5.32: *Site-specific deamidation in COL1A1 (A) and COL1A2 (B) in Skins 6, 8, 12, and 15 at each liming time point. The x-axis represents the time it takes for a certain position to deamidate, while the y-axis shows how deamidated it is (0 remaining N and Q - total deamidation). Size of the points is scaled by their intensity.*

Figure 5.32 shows the site-specific deamidation in Skins 6, 8, 12, and 15. In COL1A1 (A), it is shown that at high half-times (>4,000 days), there is no deamidation before liming, and that deamidation at these half-times increases gradually with liming duration. By 15 days of liming, all skins show a high level of deamidation in these high half-time positions. COL1A2 (B) is a little less clear cut - although the same general trend is visible around 5,000 days and 10,000 days half-time, at 7,400 days. This half-time only occurs at position 26 of the COL1A2 chain and is not identified at the 10 or 15 day time point. As this is well before the helical region of COL1A2 (which starts at position 80), it would thus follow that this position is extremely structurally susceptible to deamidation.

When analysing only sites with 10,000 days half-time, as shown in Figure 5.33, there seems to be an exponential relationship between this deamidation and days of liming. This correlation is most accurate in COL1A2, where the non-outlier skins also portray very similar levels of deamidation. No line of best fit is shown for Skin 11, as in both proteins it is completely deamidated before liming. However, as also seen in Figure 5.34, after liming it displays similar levels of deamidation to the other skins.

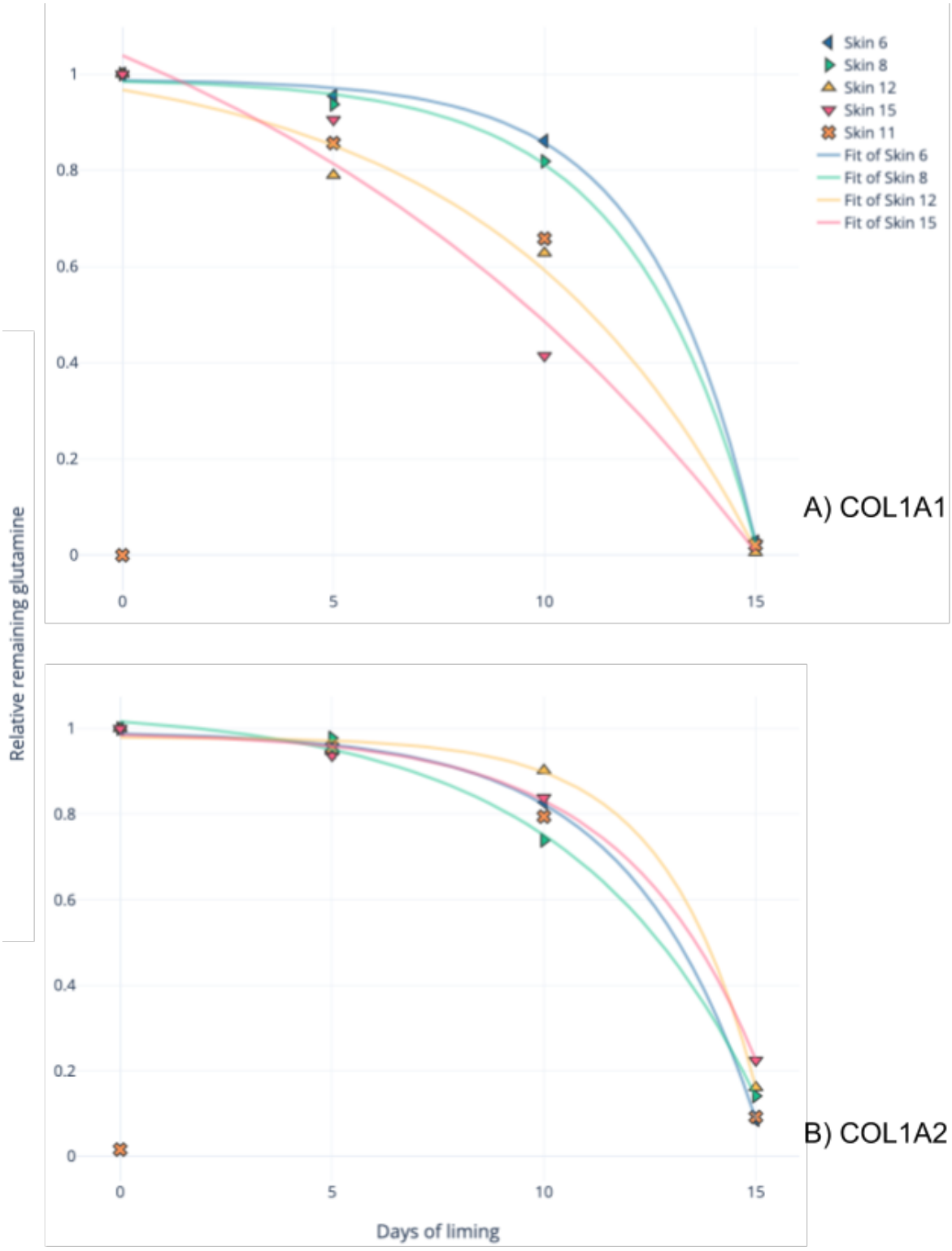


Figure 5.33: *Deamidation (relative remaining glutamine) at only sites with a half-time of 10,000 days in COL1A1 (A) and COL1A2 (B).*

5.3.9.1 Skin 11 deamidation

Bulk deamidation of just Skin 11 is shown in Figure 4.34, with the median values of the other skins represented by stars. In both COL1A1 (A) and COL1A2 (B), glutamine is almost totally deamidated before liming, but after liming displays similar values to the other skins. For asparagine, this pattern is not apparent - the deamidation before liming is not as extreme, and especially in the case of COL1A1, the distribution follows the median of the other skins very closely.

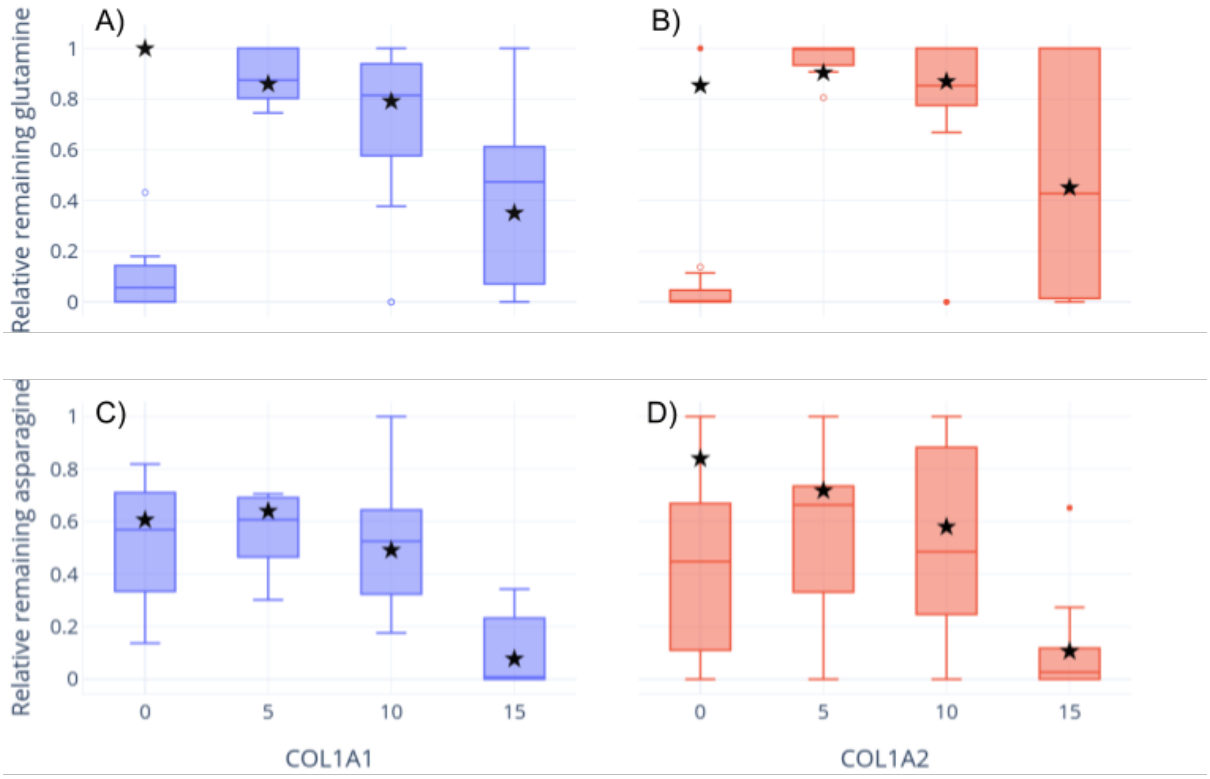


Figure 5.34: *Bulk deamidation of Skin 11 for glutamine and asparagine in both COL1A1 and COL1A2. The y-axis represents the relative remaining deaminating amino acid - that is, at 1 there is no deamidation, and 0 represents total deamidation. Filled points represent outliers, while hollow points show suspected outliers. The median value at each point of the non-outlier skins is represented by a star.*

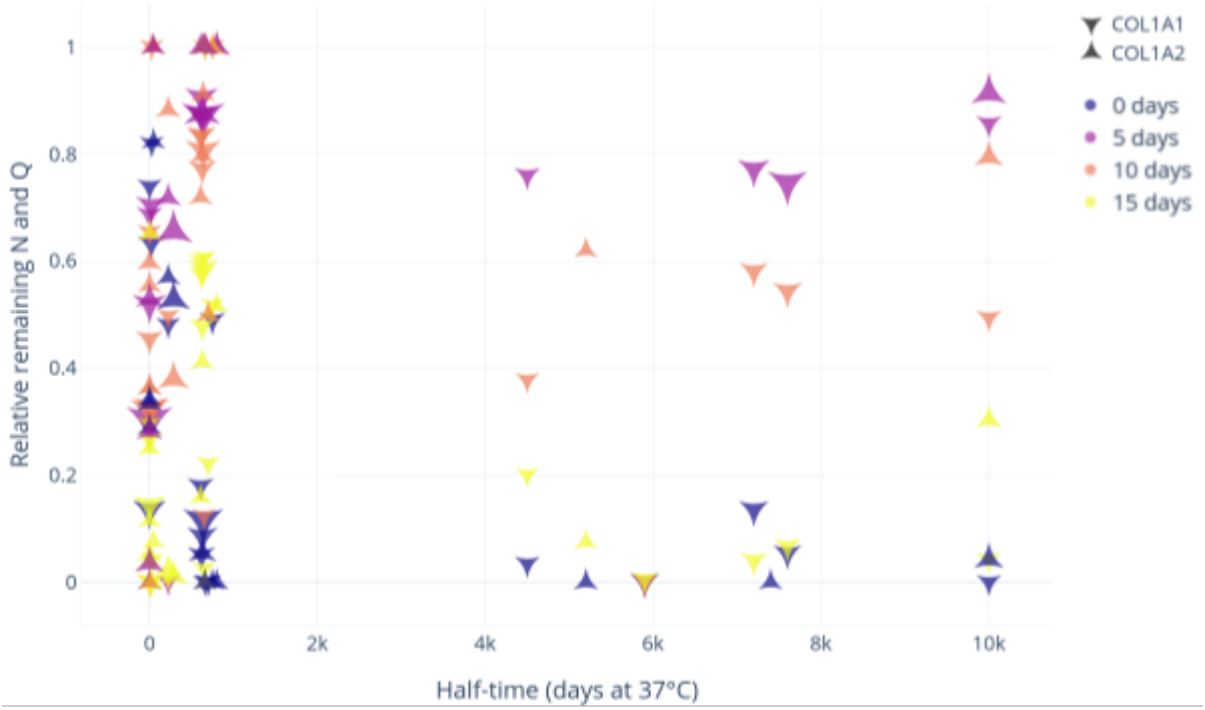


Figure 5.35: *Site-specific deamidation of COL1A1 and COL1A2 in Skin 11.*

As in Figure 5.35, there are many sites with over 4,000 day half-time that are almost totally deamidated in Skin 11 before liming. Converse to expectations, the positions which show the least deamidation before liming are asparagine positions with less than 100 days half-time, which, in theory, should be the most susceptible to random and laboratory-induced deamidation.

5.3.9.2 Deamidation in the gap and overlap regions

Figure 5.36 shows the site-specific deamidation of residues in the gap and overlap regions in the non-outlier skins at each time point. Although both proteins have similar numbers of glutamine residues in the gap and overlap zones (COL1A1: 15 gap, 12 overlap; COL1A2: 11 gap, 12 overlap), there seems to be an overrepresentation of gap deamidation after five days of liming (Figure 5.36.B). By 15 days of liming, there is not a clear difference between the gap and the overlap region, although there is a cluster of overlap residues with 10,000 days half-time that are not quite as deamidated as the gap, these all originate from one glutamine (detected in multiple skins).

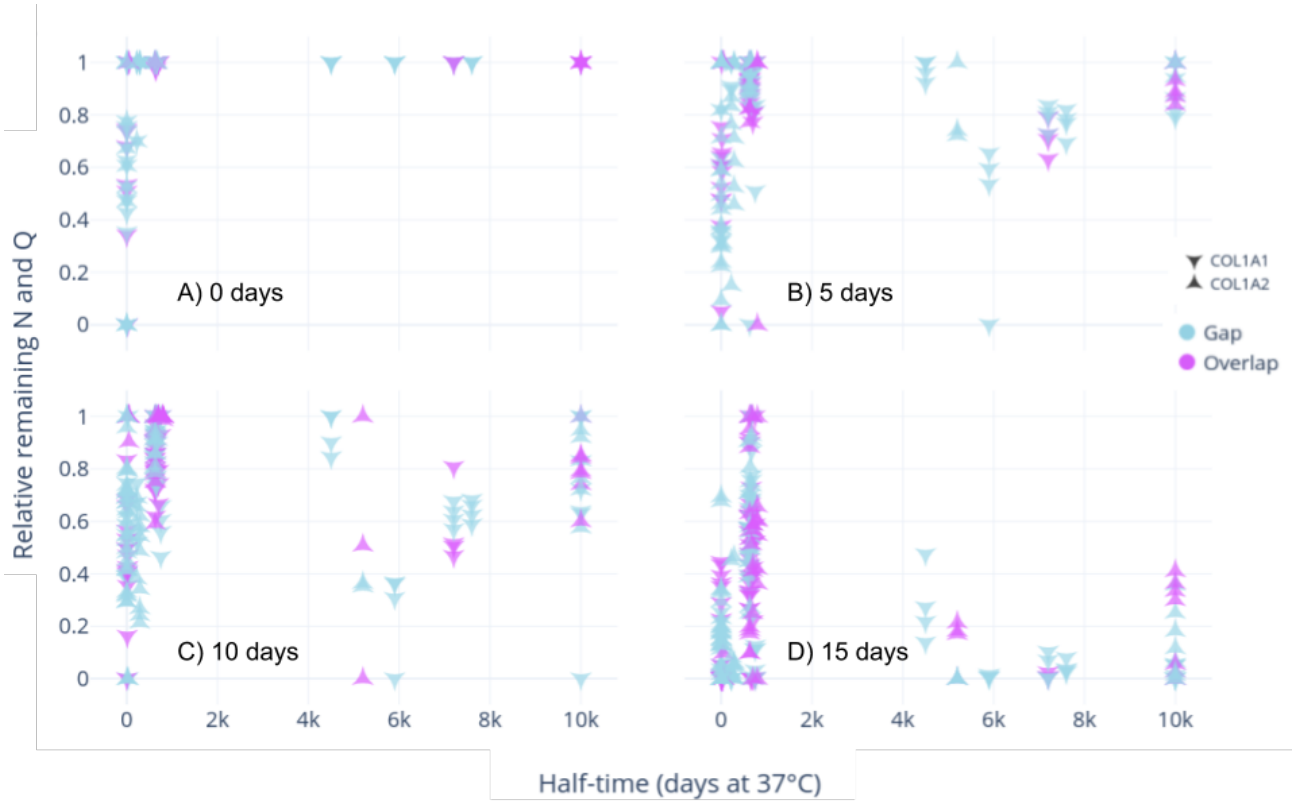


Figure 5.36: *Site-specific deamidation in all non-outlier skins (excluding Skin 11) per time point, coloured by the region of the deamidating residue.*

5.3.9.3 Deamidation near high intensity hydrolysis region

There are no deamidating (asparagine or glutamine) residues in the high intensity hydrolysis region of COL1A1 (1069 - 1093). However, proximity to this region could cause differences in the rates of deamidation. We tested this by plotting the deamidation of the four closest deamidating residues (four glutamines) - two before and two after, as shown in Table 5.7.

Position	**Half-time**	**Description**
973	7200	Next closest (<1069)
992	650	Closest (<1069)
1097	630	Closest (>1093)

1109	750	Next closest (>1093)

Table 5.7: *The four glutamine residues up- and down-chain from the region of high intensity hydrolysis*

Figure 5.37 shows these four residues compared to other similar glutamines (within 10% of the same half-time, and also in the gap region). The two residues before the region of high intensity hydrolysis deamidate very similarly to other similar glutamines (Figure 5.37.A), as does the closest residue after the region (position 1097). The next closest glutamine at position 1109, however, seems to deamidate slower than other similar glutamines - with all skins displaying no deamidation at this position until 15 days of liming.

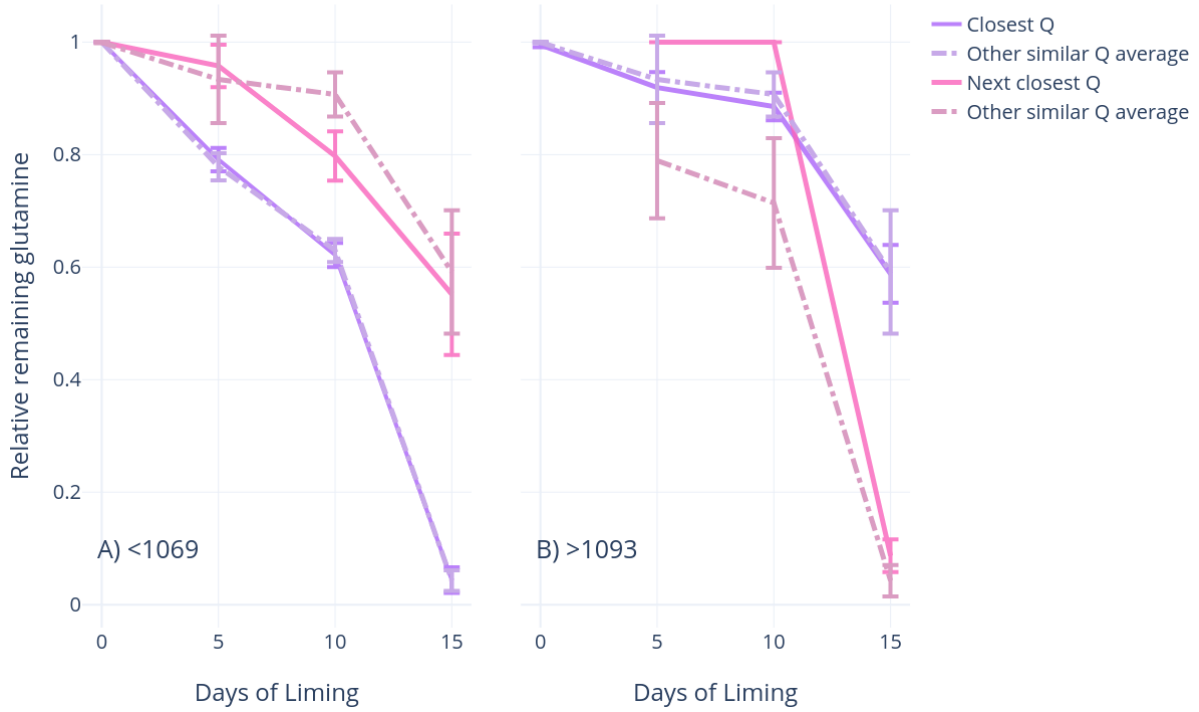


Figure 5.37: *Deamidation of the four glutamines directly before and after the region of intense hydrolysis (COL1A1: 1069 - 1093), compared to the average of similar glutamines (within 10% half-time, and also in the gap region). Error bars represent standard deviation.*

5.4 Discussion

This study sought to characterise the degradation of collagen in skins subjected to the liming process through the analysis of five experimentally limed sheep and lamb skins, sampled and characterised at four points in the liming process. We have shown that, for the sheep skins (i.e. excluding the lamb skin, Skin 11), unique peptide counts, depth, and deamidation increase with liming duration, and that whether these peptides fall in the gap or overlap region does not affect this.

Furthermore, through the analysis of hydrolytic sites in the quarter stagger structure (Figure 5.3), the breakdown of COL1A1 and COL1A2 has been investigated. As expected, the sites that are cut by trypsin (arginine and lysine) hydrolyse before liming, as trypsin was used in the sample preparation phase. After five days of liming, residues near these sites of early hydrolysis also experience cleavage (Figure 5.3.B) and show hydrolysis, which implies that this trypsin hydrolysis prompts collagen to unwind. In terms of collagen's quaternary structure, we see that the sites in the gap regions are more likely to experience hydrolysis before (and after a short duration of) liming (Figure 5.3 and Figure 5.9), but by day 15 of liming, the overlap regions are more likely to exhibit the highest levels of hydrolysis (Figure 5.10).

5.4.1 Hot spots of hydrolysis

There are two regions that are particular hot spots of hydrolysis in COL1A1: the start of the helical region at 186, and the region identified as the high intensity hydrolysis region at 1069 - 1093, as shown in Figure 5.13. The former is arguably not remarkable, as the propeptide region is cleaved before collagen fibrils are formed (Miyahara *et al.*, 1984; Kofford *et al.*, 1997). However, the high intensity hydrolysis region is a novel finding, that to our knowledge has not been reported in literature. It is an obvious hot spot in all skins at all time points in the peptigram (Figure 5.13), and is the region with the most hydrolysis before liming (Figure 5.3 and Figure 5.4). This is partly due to the short region containing three tryptic residues, but even before liming there is "random" (i.e. non-tryptic) hydrolysis in all skins, and the intensity of these cleavages, and residues directly up- and down-stream increases with liming duration. There are four pairs of amino acids in this

region that we do not detect hydrolysis between (Table 5.3 and Figure 4.6), which is unusual, as we would expect that as the fibrils become more frayed, hydrolysis becomes more likely. However, it is possible that these positions would hydrolyse giving a longer duration in lime, or that we simply do not detect the resulting peptides, due to limitations of the mass spectrometer. While, as discussed above, unique peptide count increases linearly with duration of liming, in this region it increases exponentially (Figure 5.17). Lastly, proximity to this region does not seem to affect the rate of deamidation in the four glutamines immediately up- and down-stream to this region (Figure 5.37).

The reason why this particular region experiences such intense hydrolysis is not clear. It is not a function of the region's coverage (Figure 5.19), nor its melting temperature (Figure 5.23 and Figure 5.24). Figure 5.28 shows that this region is a little more variable than the rest of the COL1A1 chain, which could indeed confer instability. The region is fairly close to a crosslink site at 1108 (Dixit *et al.*, 1978), which should increase structural strength (Oxlund *et al.*, 1995; Al-Ammar, Drummond and Bedran-Russo, 2009). It could perhaps be the case that this is the reason why we detect these peptides, as even though they are very deteriorated and easily degradable, they are at a uniquely strong point in the collagen chain, and therefore do not deteriorate beyond the limits of detection. It is also noteworthy that the collagenase cleavage site is approximately 100 residues downstream (between 960 and 961 (Chung *et al.*, 2004)), and this should have the opposite effect - conferring instability to the region. Lastly, in human COL1A1 this region is directly followed by an RGD motif starting at position 1093, which is an integrin binding site (which facilitates cell adhesion) (Taubenberger *et al.*, 2010), however, this is not the case in the *Ovis aries* sequence.

Lastly, this pattern is not unique to skins, nor alkaline treatments. As a precursor to this study, 93 bovid samples from a variety of substrates (bone, skin, and glue), locations, and ages were analysed. This pattern was found in nearly all samples analysed (apart from very modern bones, or very old samples where only a few peptides recovered). An example of this is shown in Figure 5.38 - a sheep bone from Medieval Portugal.,

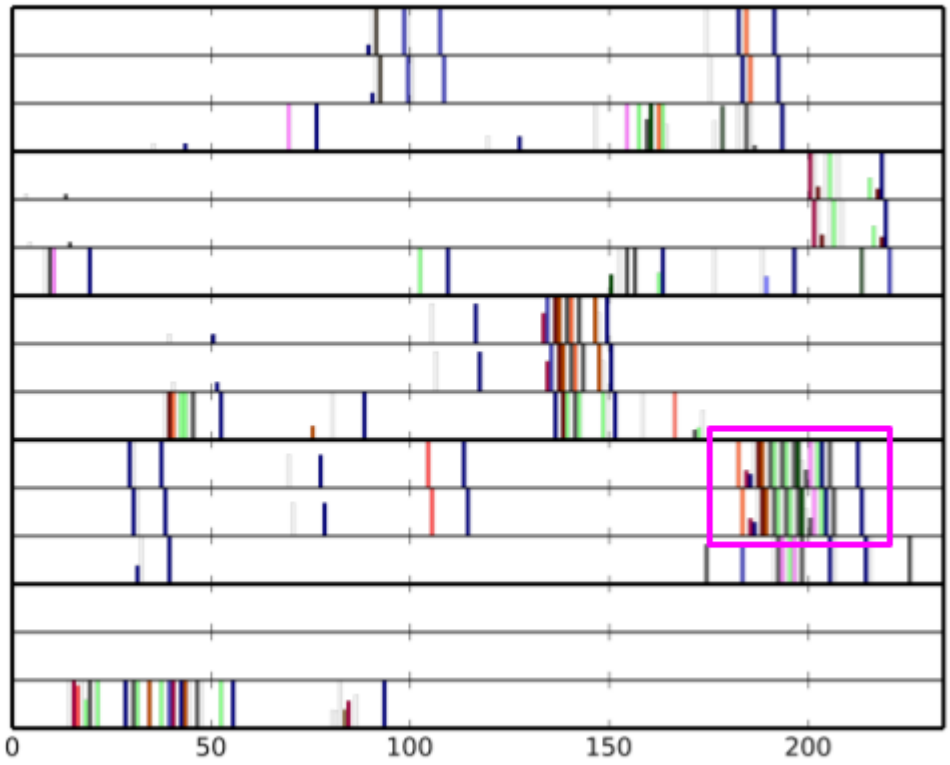


Figure 5.38: *Hydrolysis in quarter stagger structure of a Medieval sheep bone from Portugal, with the high intensity hydrolysis region shown in the pink box.*

5.4.2 Skin 11

The region of intense hydrolysis, along with many other characteristics (e.g. breadth of coverage (Figure 5.13 and Figure 5.14), unique peptide count (Figure 5.15), depth of coverage (Figure 5.21), and deamidation (Figure 5.31 and Figure 5.32)) increases evenly in all skins, apart from Skin 11, which belongs to a stillborn lamb (Table 5.1).

Before liming, Skin 11 has a much higher intensity and breadth of coverage than after five days of liming, and than other skins before liming (Figure 5.13 and Figure 5.14). We see that the depth of coverage before liming is also higher (Figure 5.21), and there are substantially more unique peptides (Figure 5.15). All of these measurements fall dramatically after five days of liming. Before liming, Skin 11 shows high hydrolysis in the high intensity hydrolysis region, but after five days only the tryptic cuts show meaningful hydrolysis (Figure 5.6). Perplexingly, before liming Skin 11 exhibits almost total deamidation in all high half-time sites (mostly glutamine residues), however the low half-time sites (mostly asparagine) only display slightly higher deamidation levels than the other skins (Figure 5.33 and Figure 5.34). This is the opposite of what we would

expect, as high half-time deamidation has been proposed as a marker of antiquity (Ramsøe *et al.*, 2020), and has never been reported in high levels in combination with low levels of short half-time deamidation.

This perplexing pattern of deamidation could be suggested to be an anomaly. However, as the average coverage of COL1A1 in Skin 11 before liming is 9.7x (Figure 5.20), and there are 62 uniquely detected deamidating residues, there is a wealth of evidence supporting this pattern. Furthermore, this would not explain the relatively normal (within error bounds) asparagine deamidation, y*et al*most complete glutamine deamidation.

A further MaxQuant search against the sheep proteome revealed that this pattern is not just in collagen, but exists in all the proteins identified in Skin 11. There was no significant differential expression of non-collagenous proteins (NCPs) detected in Skin 11. It does, however, as shown in Figure 5.39, contain a higher proportion of collagen before liming than the other skins, and conversely, retains a higher proportion of keratins and NCPs after liming than the other skins.

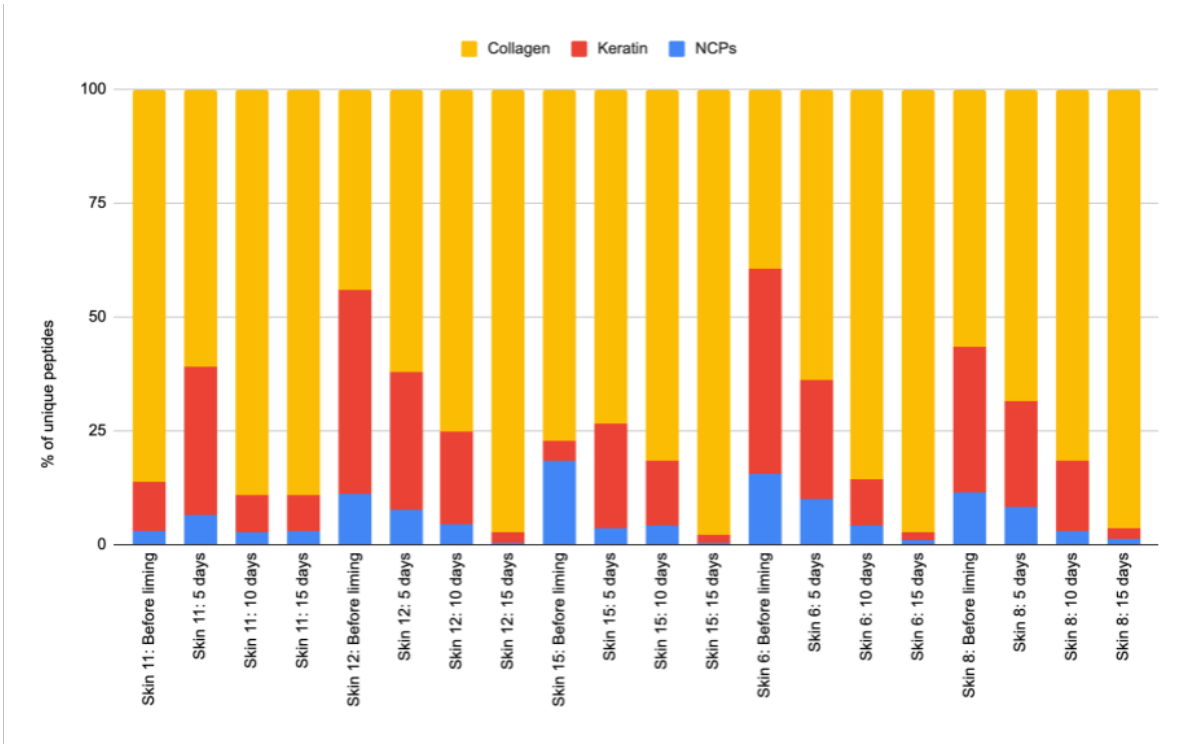


Figure 5.39: *Percentage of collagen, keratin, and other proteins in each skin at each time point.*

Skins of stillborn animals have historically been used to make parchment (in particular, uterine vellum) (Lee, 1992; Fiddyment *et al.*, 2015), so in all likelihood skins with similar patterns to Skin

11 would exist in the historical record. The fact that by 15 days of liming, all skins exhibit remarkably similar properties only reinforces this - although it is likely that Skin 11 would be thinner, it would be an equally acceptable writing surface.

5.5 Conclusion

Understanding how collagen degrades is essential for informing the conservation of parchment in historical manuscripts. This study shows that many characteristics of degradation, including peptide count, coverage, and deamidation, proceed at a predictable rate during the liming process, and that the gap region is the most susceptible to hydrolysis. We also identify a region of high intensity hydrolysis, and explore the mechanisms that may cause this pattern. Lastly, the untreated skin of stillborn animals display dramatically different patterns to that of older animals; but that after liming all skins are molecularly very similar. This supports the idea that the skin of stillborn animals was used to make vellum, as molecular similarity likely confers to structural similarity.

6. Discussion and Summary

This chapter will briefly summarise the results presented in this thesis (6.1), and then outline how the original aims of the thesis were met (6.2). It will go on to discuss the challenges faced during this research (6.3), and the implications of this work to the field as a whole (6.4). The future work will then be outlined (6.5), and the thesis will be concluded (6.6).

6.1 Summary

This thesis analysed the patterns of degradation in ancient proteomes, and examined how these characteristics can be used to develop methods to authenticate results. This section will briefly summarise each of the three main chapters.

6.1.1 Chapter 3: DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes

Palaeoproteomics suffers from a lack of widely agreed upon criteria for the authentication of results (Hendy *et al.,* 2018c). Interpreting modern contaminant proteomes as originating from ancient artefacts can obviously lead researchers to draw the wrong conclusions about the past. Conversely, discarding truly ancient proteins by labelling them as contaminants can cause researchers to miss important discoveries. It is imperative that this problem is solved.

Chapter 3 presented the development of a novel software tool, deamiDATE 1.0, that aims to use the molecular characteristics of the peptides recovered in order to differentiate between modern and ancient proteins, and thereby authenticate proteomes. It does this using two methods: bulk and site-specific deamidation. Bulk deamidation calculates the proportion of deamidated and non-deamidated glutamine and asparagine in all peptides in a given protein, and normalises this proportion to the intensity of the peptides. That is, for each protein, it outputs two values - the percentage of remaining glutamine, and the percentage of remaining asparagine. Site-specific deamidation differs as it takes into account the published rate of deamidation (which depends on amino acids flanking the deamidating residue) in order to provide more nuanced results. Rather

than the percentage of glutamine and asparagine, it produces a graph comparing the expected time of deamidation (derived largely from rates observed in pentapeptides) versus the extent of deamidation. The rule of thumb is that the slowest sites (i.e. those that take the longest to deamidate) will only be heavily altered in very old samples; whereas if a protein only exhibits deamidation at very fast sites, it may not be genuinely ancient.

Chapter 3 demonstrated the utility of this approach through the analysis of two datasets: a model dataset and a dental calculus sample. The model dataset (all from (Welker *et al.*, 2015)) comprised five modern samples and five ancient (two of which were dated to ca. 12,000 years old, while the remaining three were assigned to the Pleistocene (ca. 2,580,000 to 11,700 YBP) or Lujanian (ca. 800,000-11,000 YBP)) of South American ungulates. It was shown that the modern samples display significantly less asparagine and glutamine bulk deamidation (Figure 3.3), although this is very variable. Furthermore, only the ancient samples portrayed high levels of deamidation at high half-time sites (Figure 3.4), and there was a statistically significant difference between the extent of deamidation from ancient and modern samples at the same half-time. However, the extent of deamidation both within and between samples was extremely variable - although rough trends are visible, these are only possible as all samples analysed gave rise to a substantial amount of proteomic information.

The second dataset, the dental calculus, was extracted from a skeleton excavated from a Neolithic site near Brighton, UK. Alongside the many endogenous proteins detected, seven different proteins relating to chicken egg were identified (Table 3.2). This finding would mark the earliest ever evidence of chickens in the UK. As the chicken egg proteins were not found in the extraction blank, it was unlikely that the proteins were laboratory contaminants. Moreover, many of the peptides were specific to chicken, so the egg most likely did not originate from another bird. Using deamiDATE 1.0, the bulk deamidation of the endogenous human proteins and the chicken proteins was calculated. The deamidation of the human proteins was significantly higher than that of the chicken (Figure 3.5). Furthermore, the site-specific deamidation analysis showed that, while the endogenous human proteins showed some deamidation in high half-time sites, the vast majority of the chicken proteins showed no slow deamidation (Figure 3.6). From these analyses, it was concluded that the chicken proteins were most probably not embedded in the dental calculus during the life of the skeleton. As the remains were excavated in the 1930s, it is possible they originate from egg-based glue used by conservators at the time.

To summarise, chapter 3 presented deamiDATE 1.0, and showed that it can differentiate between proteins of various ages, and thereby authenticate results and provide a more nuanced

understanding of an artefact's history. It cautioned, however, that deamidation levels are not only influenced by time, but also by protein structure, environmental factors, and pre-depositional human activities. Furthermore, the high variation in the extent of deamidation within samples severely limits the differential power of deamiDATE. It is only with an abundance of data that these rough, yet varied patterns (such as in Figure 3.4) emerge. This precludes the effective use of this software for low-abundance data.

6.1.2 Chapter 4: Milking it? Assessing the degradation of ancient milk proteomes

In many places around the world dairy products represent a significant part of our diet, but that has not always been the case (Tishkoff *et al.,* 2007). The origins and spread of dairying has long captivated archaeologists (Craig *et al.,* 2000; Dunne *et al.,* 2012). Now, with biomolecular analysis, it is possible to detect milk in archaeological artefacts and remains, and therefore piece back together the complex history of the adoption of animal milk into the human diet (Bleasdale *et al.,* 2021). However, milk proteins are intrinsically hard to authenticate - they are typically present in very small quantities in archaeological remains, and are also commonly used laboratory reagents (Hendy *et al.,* 2018c).

Chapter 4 aimed to assess the deamidation patterns in milk proteomes, in order to ascertain whether deamidation alone can reliably differentiate between modern and contaminant milk. Four datasets were analysed in order to accomplish this - two of which were experimental datasets using modern milk.

Firstly, we quantified the deamidation of skimmed milk powder, which is a common laboratory reagent. It was shown that, despite being subject to extreme heat in the powdering process, both BLG and caseins in milk powder did not show advanced levels of deamidation (Figure 4.1), and did not have meaningful levels of deamidation at high half-time sites (a characteristic of truly ancient proteins proposed in Chapter 3) as shown in Figure 4.2. In the second experiment, five archaeological artefacts were contaminated, as the samples were prepared in a non-specialist (i.e., modern proteomic) facility. No BLG was confidently detected in any of the artefacts (using the two-peptide rule), but all samples, and the extraction blank, contained caseins (Figure 4.3).

All but one sample displayed no glutamine deamidation (and therefore no high half-time deamidation), as shown in Figures 4.4 and 4.5.

These two experimental datasets show that modern milk, even when powdered or associated with archaeological artefacts, displays very limited deamidation, with no substantial deamidation at high half-time sites. This is promising evidence that deamidation can be used to differentiate modern and ancient milk proteomes.

The last dataset consisted of 274 samples across six published articles that reported the presence of ancient milk proteins. The samples came from either pottery or dental calculus, ranged from the Neolithic to the present day, and originated from many different sites across the world. Many of the samples were radiocarbon dated, and for the rest the date was inferred from the reported archaeological period. The deamidation of both BLG and caseins in these samples was calculated, and it was shown that there is a general trend in BLG that deamidation increases over time (Figures 4.7 and 4.8). However, in these samples deamidation was extremely variable.

Though this large-scale meta-analysis can elucidate a moderate overall trend in deamidation over time, this may not translate well to individual studies, with a smaller sample size and temporal range. Even when adhering to a 'two-peptide rule' to identify ancient milk, two peptides is likely not enough to reliably authenticate the protein using deamidation. Not all peptides will even contain a deamidating residue, and at the smallest scale, deamidation is either there, or it is not. More nuanced information can only be obtained with the identification of more peptides.

Potential methods of strengthening the differentiation of ancient milk of different ages were then discussed. Firstly, a rough measure of authenticity can be made by assessing the portion of the proteome identified. For example in milk, if both beta-lactoglobulin and multiple types of casein are identified, this makes it less likely that there are spurious matches to trace amounts of milk. Secondly, it was advised that BLASTp searches are carried out to ensure that peptides that are critical to protein or species identification are unique. Lastly, it was proposed that researchers compared the deamidation of their protein of interest to the deamidation proteins that are highly likely to be endogenous to the sample, as significant similarities would imply a shared diagenetic history.

6.1.3 Chapter 5: Collagen type I degradation in limed skins

Historical manuscripts represent an important source of information about past societies and cultures at the foundation of medieval scholarship (Ryder, 1964). The biomolecular study of the parchments themselves has been recently found to provide even more insight into parchment production, dietary and husbandry practices, and previous handling and conservation practices (Fiddyment *et al.,* 2019). However, the degradation of the collagen in parchment is not well understood, and could be used to better inform conservation efforts.

Chapter 5 attempted to shed more light on how collagen in parchment breaks down, through the exhaustive data analysis of five sheep skins at four different time intervals in the liming process. It was found that four of the skins exhibited very similar characteristics in all analyses, but there was one outlier - the skin of a stillborn lamb. The work has broader implications as in almost all cases in which skin collagen is isolated (e.g. leather tanning, gelatin production), one of the first steps is dehairing in alkali.

An examination of the patterns of hydrolysis caused by liming revealed that all skins had a region of high intensity hydrolysis at COL1A1 position 1069 - 1093 in the gap zone (Figure 5.3). This damage was evident even before liming. This region, characterised by three tryptic residues, was found to always show non-tryptic cleavage, which spread to the N- and C-terminal sides during liming (Figure 5.4). Moreover, it was found that hydrolysis seems to affect some amino acids more than others, and that this differs between the two chains of collagen type I (Figure 5.7).

The number of collagen type I peptides identified in the fresh flayed skins was found to increase linearly with the duration of liming (Figure 5.15), as did the depth of amino acid coverage (Figure 5.18). This shows liming leads to the degradation of collagen, and, we speculate, the subsequent unwinding of its quaternary structure. In the high intensity hydrolysis region, the peptide count increased exponentially with liming (Figure 5.17).

It would be expected that regions of the collagen chain with the lowest melting temperature would be the first to hydrolyse, yet curiously, it was found that the melting temperature has no effect on the hydrolysis of that position (Figure 5.23). Similarly, we expected that positions with the highest number of polymorphisms (i.e. those regions most able to accept substitutions) would be the most labile and thus be more prone to hydrolysis, because collagen's structure, and therefore its strength, is conferred by its unique sequence. However, it was actually found that

the most conserved positions were more likely to be cleaved (Figure 5.26). Moreover, we found that the region of high intensity hydrolysis has a similar melting point to the rest of the collagen chain, and that it is less conserved (Figure 5.24). Lastly, there was no correlation between the melting point and the extent of conservation of residues (Figure 5.29).

The deamidation of the skins was calculated using deamiDATE 1.0. In the non-outlier skins, deamidation of both glutamine and asparagine increases during liming, and there was a clear trend of increasing deamidation at higher half-times during liming. Additionally, the deamidation of sites with 10,000 days half-time was found to have a very strong exponential correlation with time. However, as in the previous two chapters, the extent of deamidation still exhibited high variation.

The stillborn skin was an outlier as it had large amounts of non-tryptic cleavage (Figure 5.6), an increased peptide count (Figure 5.15), and an increased depth before liming (Figure 5.21), yet after the first liming time point it displayed similar characteristics to the other skins. As there was only one stillborn skin, these results must be verified with subsequent replications. This indicates that the skins of stillborn animals are biologically very different, possibly due to the breakdown of collagen (and other proteins) during autolysis - the breakdown of cells after death. We hypothesised that these damaged proteins are washed away in the liming process, which is why the skin is similar to the other skins after liming. Additionally, this is consistent with the hypothesis that stillborn skins were used to make uterine vellum (Fiddyment *et al.*, 2015), as after the liming process the skin is molecularly similar to that of the older sheep. Lastly, before liming the stillborn skin displayed almost total glutamine deamidation (including high half-time sites), but very limited asparagine deamidation. This is curious, as (chemical) asparagine deamidation is a much faster reaction than that of glutamine. We attributed this instead to biological deamidation, driven by a deamidase enzyme, though no proteomic evidence of this was found. These deamidated peptides are then thought to be removed in the subsequent liming process, which explains why they are not detected at other time points.

6.2 Achievement of aims

This thesis had two main goals:

1) Understand the diagenetic forces acting on proteins through time; and
2) Use this knowledge to develop tools and methodologies to authenticate ancient proteomes

This thesis has lead to the **increased understanding of the diagenetic forces acting on proteins through time (Aim 1)**, by characterising the extent of site-specific deamidation in samples of various ages, and thereby demonstrating its potential to differentiate between modern and ancient proteins (Chapter 3), by examining milk protein deamidation in a large dataset of experimental and archeological samples, (Chapter 4), and by characterising the patterns of hydrolysis, deamidation, and changes in proteome composition in animal skins during liming (Chapter 5). Furthermore, it has also proposed a novel **tool that authenticates ancient proteins (Aim 2)** (Chapter 3), and suggested guidelines for the reliable authentication of milk proteins (Chapter 4).

6.3 Challenges

6.3.1 Open access data

As curators are becoming increasingly sceptical of requests for destructive sampling, it is increasingly important to maximise the information gained from destructive analyses (discussed in 2.5.2 and 2.5.3). One way of achieving this is by depositing published data in online repositories, to encourage data re-use.

This thesis relied solely on the analysis of existing (both published and previously unpublished) data. Although, in Chapter 4, large comparisons were possible (with 274 milk samples from six articles), at the inception of this paper, a total of 15 articles were identified as reporting LC-MS/MS results from ancient milk proteins. However, of these only six (the six included papers) had openly shared their raw files, and the remaining nine refused a request to share their data.

Open access sharing of data allows the independent validation of results, which is commonly agreed to be a "cornerstone of science" (Moonesinghe, Khoury and Janssens, 2007). As mentioned throughout this thesis, palaeoproteomics is plagued by seemingly contradictory remarkable findings - if it was commonplace to publish open data, and therefore to verify other groups' results, perhaps this would not be the case.

It is important to note that many studies do publish some supporting data. For example, out of the eight published articles referred to above, five of them did include supplementary information with peptide identification information. This does allow some rudimentary verification of results, in that researchers can see how many peptides were found, if post-translational modifications were reported, and if they are unique to the organism of interest. However, in order to perform robust analyses that are comparable across datasets, the raw data files are required.

However, the onus to share open access data is not solely on the researchers themselves. The process of data deposition is often complex and unwieldy, and this undoubtedly does not help to encourage authors to upload their data. If this workflow was more streamlined by the data repository, in tandem with journal submittance systems, we would likely see an uptick in open access sharing. Finally, there would without a doubt be an increased pressure on the researchers to publish their data if it was a criteria for submission in all journals.

6.3.2 Bioinformatics pipelines

Another challenge of palaeoproteomic data analysis is the choice of bioinformatics pipelines. As previously mentioned, there exists a wealth of data analysis softwares, each with their strengths and weaknesses. However, no robust comparative analysis of these has been carried out for ancient protein data - and it is common that researchers will analyse the same data with different software in order to maximise results (Schroeter *et al.*, 2017; Mays *et al.*, 2018; Cappellini *et al.*, 2019).

Moreover, once the software (or software combination) is chosen, there are a multitude of different parameters that can be used to maximise peptide identification. For example, the variable and fixed modifications, the digestion mode (specific, semi-specific, or unspecific), and

the application of score filters and/or false-discovery rate cutoffs can undoubtedly affect peptide identification. Of particular interest is the digestion mode, where the user specifies the "type" of cleavage expected, which does not only rely on the protease. Specific digestion limits the search software to only seek out peptides that have been cleaved by the selected protease at both the N- and C- terminus. However, ancient proteins are expected to be fragmented prior to any protease involvement, which means that they often do not exhibit specific cleavage patterns. For this reason, semi-specific digestion is often chosen, which includes peptides that have at least been cleaved by the protease at the C-terminal side. Lastly, unspecific mode means that no protease is used in the search (regardless of whether or not it was used in the sample preparation).

That being said, the driving force behind whether or not a protein is detected is the database selection. Choosing too restricted a database will effectively force the software to identify proteins incorrectly, as the peptides better matching the spectra were not included. In these cases, the peptide score will be negatively affected due to the non-perfect match. However, as there is no commonly agreed upon score cut-off for palaeoproteomics (Hendy *et al.,* 2018c) as very few spectra-peptide matches are as conclusive as those in modern proteomics. Thus, small differences in amino acid composition are unlikely to be detected without careful manual inspection of the spectra.

Conversely, too large a database is also far from ideal. The largest databases will have many thousands of proteins that are irrelevant to the sample of interest, and as every spectrum needs to be compared to a theoretical spectrum generated from each possible peptide in the database, this increases the search space (the time and memory it takes to perform the search) exponentially. This is especially a problem while using a semi- or unspecific digestion mode, as with specific digestion assumptions can be made about the N and C termini amino acids. Additionally, peptides identified this way will be subject to an increased rate of false-positives, as this is partially calculated by matches to a decoy database.

Furthermore, search algorithms are often not optimised for palaeoproteomic data, which is often damaged, modified, and exists in low abundance. Focusing instead on modern data, all search tools will have a bias towards the characteristics of most proteomics data. There is a wealth of untapped information in every run of the mass spectrometer, as in palaeoproteomic studies, it is common that just a miniscule fraction of the spectra are matched to a theoretical mass. However, this is an ongoing area of research, with MaxQuant undergoing optimisation for ancient proteins, and researchers looking into optimising search algorithms in order to maximise spectral matches.

In this early stage of the field there can be no concrete guidelines on the "best practice" parameters, as every sample and protein mixture is different, so the result is often that multiple data analysis steps required in a "trial and error" approach (e.g. (Schroeter *et al.*, 2017; Mackie *et al.*, 2018)), where an initial search with a large database, but strict score cutoffs and a specific digestion mode is carried out in order to inform the general content of a sample. Subsequent searches with a more restrictive database, yet more inclusive match criteria are then performed in order to maximise the amount of valid information retrieved. (Froment *et al.*, 2020). Because of this, it is of the utmost importance that previously published raw data is publicly available, as it allows the reanalysis of irreplaceable data with the latest bioinformatics tools as they are released.

6.3.3 Authentication of low abundance data

Cooper and Poinar (2000) and Gilbert *et al.,* (2005) both mention 'appropriate molecular behaviour' as one of their criteria for authentication of ancient DNA sequences. That is, that truely ancient sequences should bear characteristic patterns of damage, the presence of which can be used to differentiate ancient and modern genomes (Krause *et al.*, 2010; Skoglund *et al.*, 2014). The most common form of damage used to authenticate DNA is cytosine to thymine (C-to-T) misincorporations at 5' ends (Hofreiter *et al.*, 2001; Briggs *et al.*, 2007; Key *et al.*, 2017), which is caused by the deamination of cytosines to uracils, that are then misread by DNA polymerases as thymines (Peyrégne and Prüfer, 2020). Measures of cytosine deamination at 5' overhangs are commonly used to authenticate genomic sequences (Helgason *et al.*, 2007; Skoglund *et al.*, 2014), and quantify contamination (Renaud *et al.*, 2015; Meyer *et al.*, 2016). Although peptide sequences are much more complex than that of DNA (as proteins have four times as many "bases", and the added problem of a non-uniform structure), we argue that the authentication of proteins could work in a similar manner - the more damaged a sequence, the less likely it is to be a modern contaminant.

Even with the development of software tools, such as deamiDATE 1.0, that aim to authenticate palaeoproteomics data, not all data can be reliably authenticated. The strength of authentication results from the size of the data set, if it is insufficient, it is inadvisable to draw conclusions as to the authenticity of the results. For example, taking an analogy from the field of ancient DNA, Smith and colleagues (2015) reported detecting the earliest ever evidence of wheat in the UK. However,

the coverage across the genome was not sufficient for the authentication tool mapDamage (Ginolhac *et al.*, 2011; Jónsson *et al.*, 2013) to provide an estimate of the authenticity of the sequences.

As these low abundance biomolecules are often of considerable archaeological interest, it is imperative that they are subject to some criteria of authentication. Even if a protein identification can be made using two peptides (Hendy *et al.*, 2018c), it is unlikely that said protein can be authenticated using deamidation alone. The peptides may not include a deamidating residue, and even if they both do, two data points is not sufficient. Researchers may be lucky, and have one of the proteins contain a high half-time residue that is deamidated. This would be the greatest evidence possible using deamidation alone. However, it is much more likely that they are unlucky, and there is either no deamidation, or the residue has a low half-time.

In order to combat this, multiple lines of evidence must be taken into account when authenticating this type of data. For example, whether the deamidation of proteins likely to be endogenous correlates with the deamidation of the protein of interest, whether the protein of interest is detected in the laboratory blanks, and whether the presence of the protein of interest would be easily explained by current archaeological knowledge. Using a combination of these questions, it should be possible to reliably authenticate low abundance proteins.

6.3.4 Age versus environment

Lastly, one of the biggest challenges this research faces is how to differentiate between genuinely ancient proteins and exogenously damaged proteins. Schroeter and Cleland (2016) argued that glutamine deamidation is not an indicator of the age of a sample, but of the sample's preservational quality. It is impossible to deny that there are indeed factors other than time that influence the rate of deamidation - for example environmental factors such as pH and temperature, discussed in 2.2.1.2. This logic applies to all of the methods of degradation discussed in 2.2.1.

For this reason, it is difficult to compare the extent of diagenesis in samples with very different depositional histories. This can be somewhat counteracted through the use of tools such as thermal age (Smith *et al.*, 2003), which normalise samples' age to account for the temperature of a site. However, such tools are not available for every rate influencing factor, and the calculation of the effect of a multitude of factors would likely be intractable.

Nevertheless, comparisons between different samples at the same or similar sites are less affected by these factors, as they share the same environment - therefore less of the variation is due to external factors. Lastly, perhaps the most robust comparison is between different proteins in the same artefact (such as egg and human proteins in Chapter 3). Protein groups that come from the same sample should have an intrinsically shared history. For example, it is often useful to compare the extent of deamidation in a protein that is expected to be endogenous to a sample (e.g. collagen in bone), with the deamidation in the protein of interest (e.g. dietary proteins) - if these are within the same range of deamidation values, they likely originate from the same time. However, this ignores the fact that different proteins will deamidate at different rates. As such, more analysis into the kinetics of protein degradation will shed much needed light on this problem.

6.4 Implications

One of the major products of this thesis is the development of deamiDATE 1.0 - a software tool for the authentication of proteomics results. The existence of such a tool ought to put more pressure on researchers to authenticate their data. Moreover, if raw open access data is available, the existence of deamiDATE means that researchers are able to investigate the authenticity of previously published datasets. This will lead to more robust interpretation of palaeoproteomics results, and add another layer of evidence to remarkable findings. Indeed, deamiDATE has already been used in four studies (Mackie *et al.*, 2017; Mays *et al.*, 2018; Charlton *et al.*, 2019; Wilkin *et al.*, 2020), and is used by researchers on an ad-hoc basis to verify their results. Additionally, this thesis has also explored cases where deamiDATE alone cannot provide robust authentication, and suggested guidelines for these cases. Although deamiDATE can generally discriminate between modern and contaminant proteins of high abundance, it is less suitable for use on a few, isolated peptides, such as studies aiming to identify dietary or disease-causing proteins in complex mixtures - for example in the microbiome. As discussed in 6.3.3, these cases should require multiple lines of evidence to reliably authenticate. As such, deamiDATE is best applied in the authentication of proteins of natural high abundance (e.g. collagen in bone), or proteins identified confidently, with many proteins originating from the same species or proteome (e.g. the milk proteins identified in ceramic vessels by Hendy *et al.,* (2018b)).

Lastly, this thesis has elucidated some of the patterns of damage in truly ancient proteins, and discussed some of the mechanisms behind them. This increased knowledge about how proteins fall apart can a) inform the development of other methods of authentication and b) inform conservation efforts of museum materials.

6.5 Future work

This thesis represents a step forward in the development of analytical software tools for palaeoproteomics, an area which has been hitherto somewhat neglected. As discussed in 2.4.3, there is a considerable amount of work going into optimising wet lab protocols for ancient samples (Jiang *et al.*, 2007; Wadsworth and Buckley, 2014; Schroeter *et al.*, 2016; Le Meillour *et al.*, 2018; Lanigan *et al.*, 2020); but the dry lab lags behind. Advancement in this area, which can only be made possible by large-scale analyses of the patterns of degradation, is key for the progress of the field.

This applies to all stages of proteomic data analysis - from the search algorithm to data exploration and authentication. For most ancient protein projects (certainly all I have been a part of), most of the computational time is used by the search software, rather than the downstream analysis. This is the inverse to the field of palaeogenomics, where mapping reads to the genome of interest (roughly equivalent to the database search step in proteomics) is extremely quick compared to the analytical bioinformatic pipelines. This implies that there is room for considerable growth in the downstream analyses of proteomics data. For example, as mentioned in 6.3.2, most current experiments see only a few percent of spectra successfully matched to a database. Improvements in search algorithms could increase this number, leading to more successful and reliable identifications, and therefore more archaeological insight.

Furthermore, experiments using artificial aging and closed systems could provide more insight into protein breakdown. By minimising the number of variables involved, it may be more simple to fully explore the patterns of diagenesis induced and the mechanisms behind them. Moreover, as proteins are structurally complex, different proteins will degrade in different ways and at different rates. For this reason, analysis should focus on proteins of archaeological interest. The more these strategies improve, the more essential it will be to revisit old datasets.

Lastly, the name deamiDATE 1.0 strongly implies that it is simply the first step. One of the most important updates will be to include more post-translational modifications, as they act at different rates and under different conditions, and are well characterised and detected by their unique mass-shift. Nevertheless, including more PTMs will increase the search space, and therefore peptide identifications will become less reliable as the possibility for false positives increases (Froment *et al.*, 2020). Therefore, it is imperative that deamiDATE 2.0 only quantifies PTMs that would be included in the search step for biological reasons (e.g. oxidation), rather than just as another line of evidence. One of the characteristics of authenticity used by palaeo geneticists is sequence length (Wall and Kim, 2007), as DNA sequences fragment due to hydrolysis over time. However, this is not directly applicable to palaeoproteomics, as sample preparation commonly (though not always, see (Demarchi *et al.*, 2016; Welker *et al.*, 2019)) includes the use of a protease that cleaves sequences to a shorter length (Lanigan *et al.*, 2020). Nevertheless, measures of non-tryptic cleavage or the abundance of "pac-man peptides" (those that exist in increments of one amino acid), could also be added to deamiDATE 2.0. Together, these changes will provide more nuanced information to the differentiation of ancient and modern proteomes. However, more modifications will also increase the complexity of the results. In order to combat this, deamiDATE 2.0 will compare the characteristics of a given protein with other known-age proteins in a database using machine learning methods, and be able to produce a similarity score. Thus, proteins are not classified as either modern or ancient, but can be narrowed down to a more precise date range.

6.6 Conclusion

This thesis aimed to (a) investigate the diagenetic forces acting on ancient proteins by quantifying the patterns of degradation in truly ancient proteins, and (b) use this knowledge in order to develop tools to authenticate results. In the development of deamiDATE 1.0, this study has introduced a novel tool which can aid the differentiation of ancient and modern proteomes using deamidation. Furthermore, a large-scale meta-analysis showed that the deamidation of milk proteins and collagen in the archaeological record is highly variable. Therefore, although it is a useful tool, deamiDATE alone cannot reliably authenticate such proteomes, and further guidelines for the authentication of low abundance and common contaminant proteins are suggested. This study's last contribution is an exhaustive data analysis of the degradation of collagen in limed skins, which discovered a hotspot of hydrolysis in the collagen chain. These results represent a step forward in understanding the diagenetic forces acting on truly ancient proteins, and thus enable the development of more robust methods of authentication. Although this research makes a substantial contribution, in future work it is of the utmost importance that more attention should be given to the development of software specialised to palaeoproteomic datasets. This can only be driven by the comprehensive analyses of large-scale datasets, which, in turn, are only possible if researchers deposit their published data in open-access repositories.

Appendix A, Table 1

The LC-MS/MS raw data produced in this study and Mascot search results are available in the public database on the MassIVE repository (ID: MSV000084284; http://massive.ucsd.edu). Supplementary Table 3.1 presents the run order and file names for each sample.

Extraction Blank	eBK	QE781_MSQ1032_20170512_KristaMcGrath_eBK1_calculus_5.raw	QE781_MSQ1032_20170512_KristaMcGrath_eBK1_calculus_5.mgf	QE781_MSQ1032_20170512_KristaMcGrath_eBK1_calculus_5_F149524.mzid
Injection Blank	Preblank_VVH01	QE781_Preblank_vvh01_5.raw	QE781_Preblank_vvh01_5.mgf	QE781_Preblank_vvh01_5_F149645.mzid
Whitehawk SK1	VVH01	QE781_MSQ1032_20170512_KristaMcGrath_VVH01_calculus_5.raw	QE781_MSQ1032_20170512_KristaMcGrath_VVH01_calculus_5.mgf	QE781_MSQ1032_20170512_KristaMcGrath_VVH01_calculus_5_F149525.mzid

Appendix A, Table 1: Tandem mass-spectrometry run order and information on proteomic files uploaded to the MassIVE repository

Appendix A, Table 2

Sample	Protein	AA	RelNonDeam	Protein	Name
Default	sp|C0HJN3|CO1	N	1	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.5996408806	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.8781611535	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.8860766561	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.8861693678	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.7788232723	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.8407217441	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.9107342657	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.7079992001	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.8847352688	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.7508019697	COL1A1 N	Aardvark N
Default	sp|C0HJN3|CO1	N	0.3840984356	COL1A1 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.8927951497	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.8601954151	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.602493378	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.5624433829	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.8968080748	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.4858555837	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.5830133574	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.9845976742	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.0688257655	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.77664985	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	1	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.5498419574	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	1	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	1	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.7037545819	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.6322693381	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.9798433751	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	1	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.6220346919	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.9027375964	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.830767095	COL1A2 N	Aardvark N
Default	sp|C0HJN4|CO1	N	0.9210607433	COL1A2 N	Aardvark N
Default	sp|C0HJN3|CO1	Q	1	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8382683642	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.855341226	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8395089545	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7290597091	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.6294262789	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.5248780701	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	1	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	1	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7648695709	COL1A1 Q	Aardvark Q

Default	sp|C0HJN3|CO1	Q	1	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.609252128	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7985589563	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8758019984	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.909527464	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7762897279	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7361422535	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.7573264327	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.5733095469	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	1	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8332878	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8525670791	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8819234035	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.6156022302	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.9517496105	COL1A1 Q	Aardvark Q
Default	sp|C0HJN3|CO1	Q	0.8604209263	COL1A1 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	1	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.9901317545	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.794259878	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	1	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.7459862909	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.8560210742	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	1	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.864336977	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.7861847511	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.8019477177	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.6801056538	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.9070000777	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.5000329861	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.9097258973	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	1	COL1A2 Q	Aardvark Q
Default	sp|C0HJN4|CO1	Q	0.6847153147	COL1A2 Q	Aardvark Q
1	sp|C0HJP1|CO1	N	0.918586105	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.5701218837	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.0428369285	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.151712551	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.1825344114	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.53840222	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.8664272944	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.9183261559	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	1	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	N	0.0672006294	COL1A1 N	Anteater N
1	sp|C0HJP1|CO1	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1	Q	1	COL1A1 Q	Anteater Q

1	sp|C0HJP1|CO1/	Q	0.9034075832	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.985863056	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.9712050676	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.8544104326	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.9220096093	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.7102151931	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.9411994692	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.6423473351	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.9162210331	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.7733026093	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.1802747688	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.9372535765	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.8041095271	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	0.4069653622	COL1A1 Q	Anteater Q
1	sp|C0HJP1|CO1/	Q	1	COL1A1 Q	Anteater Q
1	sp|C0HJP2|CO1/	N	0.505462476	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.9602784415	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.9397871817	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.0996001358	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.0503074545	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.0718363293	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.9959022415	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.6896981732	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.774379198	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.7358436735	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.1349100325	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.3655060263	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	1	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0.436177433	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	1	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	N	0	COL1A2 N	Anteater N
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	0.9154891168	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	0.3705928113	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	0.6731963333	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	0.8231244654	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1/	Q	1	COL1A2 Q	Anteater Q

1	sp|C0HJP2|CO1A	Q	0.9749673924	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1A	Q	0.9508954462	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1A	Q	0.7754761101	COL1A2 Q	Anteater Q
1	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP1|CO1A	N	0.8244071332	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.5984130522	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.1606622205	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.1666519686	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.9927769083	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.6326243377	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.7696883956	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	1	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	N	0.0910064784	COL1A1 N	Anteater N
2	sp|C0HJP1|CO1A	Q	0.9538633628	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9143064423	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9737268754	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9691084139	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.8469359568	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9286643821	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.6818124072	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.7691605948	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9100940633	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.8478432048	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.7636501463	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.1344400666	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9240048704	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.8100868298	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.9152186775	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	0.1461311809	COL1A1 Q	Anteater Q
2	sp|C0HJP1|CO1A	Q	1	COL1A1 Q	Anteater Q
2	sp|C0HJP2|CO1A	N	0.8612848383	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.9625455613	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.8741570355	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.0505591435	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.1841822391	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	1	COL1A2 N	Anteater N

2	sp|C0HJP2|CO1A	N	0.6736125116	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.6183325964	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.7623338917	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.1380784971	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.2227079295	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	1	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	1	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0.4463766398	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	1	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	N	0	COL1A2 N	Anteater N
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.0571489648	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.8921328403	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.8425089426	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.6710426089	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.8378986318	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	1	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.8139969906	COL1A2 Q	Anteater Q
2	sp|C0HJP2|CO1A	Q	0.9132635275	COL1A2 Q	Anteater Q
Default	sp|C0HJN5|CO1A	N	0.9461322155	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	0.9559284119	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	0	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	1	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	0.062225688	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	0.013971448	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	0.1300248429	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	N	1	COL1A1 N	Hippo N
Default	sp|C0HJN5|CO1A	Q	0.959530746	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	0	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	0.9857202608	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	0.9602018304	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	0.993978685	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	0.9887502807	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1A	Q	1	COL1A1 Q	Hippo Q

Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	0.9853772572	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	0.9919567196	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	0.997183081	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	0.8099756931	COL1A1 Q	Hippo Q
Default	sp|C0HJN5|CO1	Q	1	COL1A1 Q	Hippo Q
Default	sp|C0HJN6|CO1	N	0	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	1	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.9753481387	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	1	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	1	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	1	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.0873592633	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.0353545119	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.0884828744	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.9691681725	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	1	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	N	0.7946855468	COL1A2 N	Hippo N
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	0.9759531439	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	0.980794374	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	0.9407558338	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	0.9930762175	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo Q
Default	sp|C0HJN6|CO1	Q	1	COL1A2 Q	Hippo QSample
4	sp|C0HJP5|CO1	N	0	COL1A1 N	Macrauchenia N
4	sp|C0HJP5|CO1	N	0	COL1A1 N	Macrauchenia N
4	sp|C0HJP5|CO1	N	0	COL1A1 N	Macrauchenia N
4	sp|C0HJP5|CO1	N	0	COL1A1 N	Macrauchenia N
4	sp|C0HJP5|CO1	N	0	COL1A1 N	Macrauchenia N
4	sp|C0HJP5|CO1	Q	0.1873204731	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1	Q	0	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1	Q	0.3261758419	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1	Q	0.1067316032	COL1A1 Q	Macrauchenia Q

4	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0.4374772147	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0.083465347	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0.1934250036	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0.2494266914	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP5|CO1A	Q	0.1434902546	COL1A1 Q	Macrauchenia Q
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0.0381581054	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0.0563082577	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	N	0	COL1A2 N	Macrauchenia N
4	sp|C0HJP6|CO1A	Q	1	COL1A2 Q	Macrauchenia Q
4	sp|C0HJP6|CO1A	Q	0.1816493897	COL1A2 Q	Macrauchenia Q
4	sp|C0HJP6|CO1A	Q	0	COL1A2 Q	Macrauchenia Q
4	sp|C0HJP6|CO1A	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	N	0	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0.0249709324	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0.0813626804	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0.0341288402	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	N	0.135405089	COL1A1 N	Macrauchenia N
5	sp|C0HJP5|CO1A	Q	1	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.1517874902	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0423571488	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.2845902583	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.2030729488	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0294329638	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0137319246	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0635278053	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0774272414	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.0081884953	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.088305944	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	1	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.022323848	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1A	Q	0.1070889984	COL1A1 Q	Macrauchenia Q

5	sp|C0HJP5|CO1/	Q	0	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP5|CO1/	Q	0	COL1A1 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.031760173	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.0243522411	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.0141036994	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.1114056314	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.0937935007	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.2290198894	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	N	0.0487222034	COL1A2 N	Macrauchenia N
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.0502057171	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.2142053745	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.218220795	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.0661347497	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.0493230554	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0	COL1A2 Q	Macrauchenia Q
5	sp|C0HJP6|CO1/	Q	0.879896046	COL1A2 Q	Macrauchenia Q
Default	sp|C0HJP3|CO1/	N	0.4827820661	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0.7182704969	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0.6063379967	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0.9251224303	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0.1600653132	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	N	0.403036309	COL1A1 N	Mylodon N
Default	sp|C0HJP3|CO1/	Q	0.8592118254	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.9941414075	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.8144088317	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.7754939578	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.2716014898	COL1A1 Q	Mylodon Q

Default	sp|C0HJP3|CO1/	Q	0.7356403187	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.6455778358	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.8035037109	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.920993828	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.6839682804	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	1	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.1056827115	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.2465158128	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.0447350301	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.9897257855	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.8385773105	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.9251224303	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	1	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.6668745182	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.5973192187	COL1A1 Q	Mylodon Q
Default	sp|C0HJP3|CO1/	Q	0.5572307507	COL1A1 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	N	0.9724700715	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.1177637126	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.635577042	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.6898266624	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.2941211993	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.0308738854	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.9714338144	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.3536250957	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.9401321321	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0.2302714378	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	0	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	N	1	COL1A2 N	Mylodon N
Default	sp|C0HJP4|CO1/	Q	1	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.8650782149	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.6707812979	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.8603702279	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.4946306178	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.9719455534	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.9343696364	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.9604073487	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.9491216187	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	1	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.8995008203	COL1A2 Q	Mylodon Q
Default	sp|C0HJP4|CO1/	Q	0.7196925058	COL1A2 Q	Mylodon Q
Default	sp|C0HJN7|CO1/	N	0.8106129799	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1/	N	0	COL1A1 N	Tapir N

Default	sp|C0HJN7|CO1.	N	1	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	0.0860349817	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	1	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	0.0467395531	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	0.8873125752	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	1	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	0.9426879553	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	N	0	COL1A1 N	Tapir N
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9954983403	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.8947301183	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9854252495	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9482220718	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9842346752	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9668266916	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	0.9934651609	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN7|CO1.	Q	1	COL1A1 Q	Tapir Q
Default	sp|C0HJN8|CO1.	N	0	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.9412625174	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.0971086033	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.7986699348	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	1	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.9458608983	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.5693277856	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	1	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	N	0.0543576985	COL1A2 N	Tapir N
Default	sp|C0HJN8|CO1.	Q	0.9978808859	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1.	Q	1	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1.	Q	1	COL1A2 Q	Tapir Q

Default	sp|C0HJN8|CO1	Q	0.9735778049	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	0.9520639413	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	0.9850377126	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	0.9563306139	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	1	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	1	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	1	COL1A2 Q	Tapir Q
Default	sp|C0HJN8|CO1	Q	0.890380588	COL1A2 Q	Tapir Q
1	sp|C0HJP7|CO1	N	0.1599889589	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0.1637030861	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0.1572944643	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0.4501448332	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0.2978587226	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	N	0.1573326118	COL1A1 N	Toxodon N
1	sp|C0HJP7|CO1	Q	0.4662005488	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.5440624259	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.7357540608	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.2867856943	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.6417927266	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.366401212	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.5654945055	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.4521661919	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.8404591186	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.1719752057	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.0378151105	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.7148586855	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.3945399918	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.3728620786	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.5611628139	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.2680958346	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.4028134651	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0.132234972	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	1	COL1A1 Q	Toxodon Q
1	sp|C0HJP7|CO1	Q	0	COL1A1 Q	Toxodon Q
1	sp|C0HJP8|CO1	N	0.4005679797	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.2155836012	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.3011074672	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.2724814074	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.3196884834	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.0792547113	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1	N	0.0941951067	COL1A2 N	Toxodon N

1	sp|C0HJP8|CO1A	N	0.4079820867	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	N	0.1560753682	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	N	0.1678956744	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	N	0.1569116642	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
1	sp|C0HJP8|CO1A	Q	0.4037217706	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	0.3703217403	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	0.3630671169	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	1	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	0.4950917461	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	0.5091440205	COL1A2 Q	Toxodon Q
1	sp|C0HJP8|CO1A	Q	0.3059134933	COL1A2 Q	Toxodon Q
2	sp|C0HJP7|CO1A	N	0.2594013374	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0.0617951692	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0.1082746562	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0.6163851579	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0.155644337	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	N	0.5407783049	COL1A1 N	Toxodon N
2	sp|C0HJP7|CO1A	Q	0.6040474172	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.7087828438	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.4715772427	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.2372480177	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.5959656141	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.8059515009	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.4453621673	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.5638814898	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.9115287385	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.2676419529	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.0413811085	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.7290276745	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	1	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.4797117062	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.6608727738	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.4632114704	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	1	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	1	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.5488257346	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.2849810423	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	1	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0.5482697649	COL1A1 Q	Toxodon Q
2	sp|C0HJP7|CO1A	Q	0	COL1A1 Q	Toxodon Q
2	sp|C0HJP8|CO1A	N	0.3982210512	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.3500673397	COL1A2 N	Toxodon N

2	sp|C0HJP8|CO1A	N	0.1092162257	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.2977952879	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.1138361182	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.2011808303	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.5323249795	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.0706278488	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.1138060685	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.5957603832	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.1170751816	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.5366017714	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0.5526433842	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	N	0	COL1A2 N	Toxodon N
2	sp|C0HJP8|CO1A	Q	1	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0.4983758098	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0.5009029241	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0.560487029	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	1	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0.5918655389	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	1	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0.5518323687	COL1A2 Q	Toxodon Q
2	sp|C0HJP8|CO1A	Q	0	COL1A2 Q	Toxodon Q

Appendix A, Table 2: Bulk deamidation in model dataset, showing the protein, amino acid, relative remaining (non-deamidated proportion) and the sample.

Appendix A, Table 3

Half-time	RelNonDeam	Size	Protein	Sample	Name
650	0.0483131941	102.4594466	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
620	0.7206461859	115.8345627	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1	0.8855415691	142.5722257	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
630	0.7860995585	130.9638186	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1.18	0.6636907715	110.7015873	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
-1	0.9845976742	105.6409328	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
0.96	0	114.3368229	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1.26	0.7083374178	132.601785	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
750	0.8724233437	101.6524923	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1.14	0.7508019697	116.1302049	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
700	0.909527464	125.4228802	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
7200	0.8457247341	139.6985671	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
10000	0.6383104842	103.5389452	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
640	0.6156022302	149.576382	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
4500	0.7648695709	114.4929303	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1.45	0	105.7372737	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
-1	0.8977246013	107.4103602	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
5900	0.7284071857	125.1698262	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
660	1	100.5191168	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
-1	0.5498419574	105.6409328	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
7600	0.573028807	133.5174682	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
224	0.8750477185	101.7951287	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
11.8	0.885770462	112.4801453	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
48.2	0.8861693678	109.7896074	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
28	0.9107342657	114.3368229	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
21.1	0.8404947727	134.4878471	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
620	1	102.4493994	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
630	0.870524835	107.1035131	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
1.18	0.6418289753	122.2709412	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
610	0.9070000777	104.4010835	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
800	0.7012863131	105.3151461	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
1.02	0.5856852474	107.188194	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
1.14	0.6644718844	104.143363	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
5700	0.7459862909	155.4871622	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
1.08	0.5810316008	138.1007382	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
10000	0.5577319901	106.2537383	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
5900	0.8326053999	100.6163165	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
5900	0.6847153147	110.4739046	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
660	1	100.7594847	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
39.8	0.8968080748	101.9740338	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
224	0.0688257655	107.8958826	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
57.8	0.6200382377	117.7349196	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
-1	1	100.5569695	sp|C0HJN3|CO1A1_ORYAF	Default	Aardvark
1	1	100.3481457	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
287	0.9207787029	127.5260549	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
48.2	0.9830256132	103.5624367	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
21.1	1	100.4791456	sp|C0HJN4|CO1A2_ORYAF	Default	Aardvark
650	0.6675307861	100.7366393	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater

620	0.9480870885	113.8897787	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
630	0.7797272034	119.1799616	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
0.96	0	107.9599962	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
1.26	0.151712551	130.0794436	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
610	0.9373458794	109.9372583	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
750	0.9076420545	104.9369479	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
1.14	0.1825344114	119.6302641	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
7200	0.9081930213	152.9631369	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
10000	0.64936286	106.8052191	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
640	0.3710449109	104.7058354	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
4500	0.9708450504	117.3772166	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
1.45	0	109.2184714	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
5900	0.9134287093	118.9924413	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
660	1	101.398756	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
7600	1	133.0387844	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
-1	1	100.5268034	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
11.8	0.53840222	113.5828982	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
287	0.9136277371	105.5754457	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
48.2	0.9183261559	113.5828982	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
28	0.5660207595	109.8752711	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
21.1	0.8664272944	117.1812369	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.1	Anteater
620	1	101.1495363	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
630	0.8538199146	103.4352165	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
1.18	0.0607744072	114.682174	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
610	1	102.187537	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
750	0.9508954462	109.3616443	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
800	0.85579696	105.9165296	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
-1	0.774379198	102.187537	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
1.02	0.0996001358	101.6933161	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
1.14	0	100.8067932	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
1.08	0.1349100325	115.2160444	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
10000	0.7754761101	111.5502189	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
640	0.7067164602	103.4093606	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
5200	0.6731963333	104.8992124	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
5900	1	101.5290523	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
660	1	106.4809304	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
39.8	0.8696739742	108.7778753	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
-1	0.8778821637	109.3266049	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
57.8	0.4509092329	105.0733886	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
-1	0.3655060263	102.187537	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
287	0.9959022415	210.4013995	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
48.2	0.7358436735	105.4457306	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
28	0.9394059162	105.4351977	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.1	Anteater
620	0.9371750393	114.9591378	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
10000	0.9116183149	110.0314876	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
640	1	101.8644379	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
7200	0.9177291681	151.3729755	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
660	1	101.5116179	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
1.45	0	107.9657557	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater

-1	1	100.6190507	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
5900	0.8427340328	119.9866182	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
7600	1	139.8771137	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
11.8	0.9927769083	118.7153605	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
630	0.7271614346	118.4192479	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
287	0.8128354266	106.0440473	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
4500	0.9691084139	117.0887116	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
48.2	0.7696883956	118.7153605	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
28	0.5984130522	112.620902	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
1.14	0.1666519686	119.0187489	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
610	0.7467513967	106.4370731	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
0.96	0	112.620902	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
21.1	0.6326243377	107.2703418	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
750	0.8825750637	105.3553041	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
1.26	0.1606622205	134.9612161	sp|C0HJP1|CO1A1_CYCDI	Cyclopes.2	Anteater
620	0	100.6370238	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
630	0.7909996893	104.6806416	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
1.18	0.1122778134	115.8460266	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
610	0.9132635275	104.1911662	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
750	1	109.0348861	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
800	0.8444065582	109.4372559	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
-1	0.6183325964	104.1911662	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
1.02	0	101.6274797	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
1.14	0	101.1781911	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
1.08	0.1380784971	118.9695882	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
10000	0.8139969906	112.432984	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
640	0.9026981269	106.7646478	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
5200	0.6710426089	104.6473136	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
5900	1	100.6629749	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
660	1	106.1942345	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
39.8	0.8612848383	111.615378	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
-1	0.8728029934	108.3160042	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
57.8	0.563750051	104.8759105	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
-1	0.2227079295	104.1911662	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
287	1	190.8782259	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
48.2	0.7623338917	105.2998761	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
28	0.8741570355	104.8305971	sp|C0HJP2|CO1A2_CYCDI	Cyclopes.2	Anteater
620	1	115.4024508	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
630	0.9799159564	195.0500518	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
1.18	0.0353545119	141.5633616	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
1.08	0.0884828744	132.4395265	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
10000	1	137.2718135	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
-1	1	122.9517024	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
57.8	0.9013273528	135.355318	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
660	0.9428648557	122.1443138	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
5200	1	104.4726698	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
1.02	0.0176849041	111.4992224	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
287	0.9517785924	348.7600023	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
48.2	0.9753481387	147.8366221	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo

1.03	1	120.3909474	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
1	0	120.3767445	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
750	1	126.7415777	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
800	1	123.8490891	sp|C0HJN6|CO1A2_HIPAM	Default	Hippo
650	1	113.5237012	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
620	1	138.5024124	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
10000	0.959530746	125.7494233	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
1	0.0613637583	299.4498078	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
700	1	128.2854112	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
7200	0.9868883668	138.4602448	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
224	0.9559284119	129.6599635	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
630	0.988034031	131.699851	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
640	0.9936644339	212.9171469	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
11.8	1	100.3632142	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
1.26	0.1252989032	176.822931	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
4500	0.5162653035	105.0451031	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
1.45	0	107.4381431	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
28	1	103.9340354	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
5900	0.9833828653	148.7586219	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
610	1	133.7450433	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
21.1	0.9457136942	156.8021243	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
750	0.9976919015	132.8047061	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
7600	1	127.9389519	sp|C0HJN5|CO1A1_HIPAM	Default	Hippo
10000	0	283.069264	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
7200	0.0658276114	240.2373911	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
224	0	199.2351936	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
630	0.0849293546	216.7614423	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
640	0.2098475091	238.1637592	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
21.1	0	150.9962967	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
1.26	0	202.6080944	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
1.45	0	145.6949795	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
5900	0.1067316032	211.9939499	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
610	0.1314123074	139.9381766	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
1	0	618.0603787	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
7600	0	185.192219	sp|C0HJP5|CO1A1_MACSX	1	Macrauchenia
10000	0	294.2829431	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
1.18	0	212.6501669	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
224	0.0563082577	217.7637755	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
630	0.1585793051	277.2727828	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
-1	0	680.2178572	sp|C0HJP6|CO1A2_MACSX	4	Macrauchenia
1	0	134.8048306	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
287	0.0381581054	171.7871677	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
1.14	0	209.3715239	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
1.08	0	221.8240796	sp|C0HJP6|CO1A2_MACSX	1	Macrauchenia
650	0	106.1881466	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
10000	0	125.2199093	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
7200	0.0139527775	118.8556102	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
224	0.0341288402	133.5115708	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
630	0.0658584218	119.3437525	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia

640	0.0681462417	145.8895371	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
21.1	0	124.3798956	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
660	1	111.2832425	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
1.26	0.0820182053	133.2350368	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
4500	0.088305944	102.5089751	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
1.45	0	103.5946925	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
5900	0.1903663214	125.8654934	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
1.14	0.135405089	103.0201006	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
610	0	114.6081923	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
1	0.0252792444	323.8165142	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
750	0.0774272414	105.5583033	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
7600	0	131.8139387	sp|C0HJP5|CO1A1_MACSX	2	Macrauchenia
620	0	104.8559263	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
10000	0	121.4485859	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
1.18	0.0416370809	122.277266	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
1.08	0.0986481577	112.1890173	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
630	0.0569618979	127.4506275	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
-1	0.0240421488	142.132225	sp|C0HJP6|CO1A2_MACSX	5	Macrauchenia
57.8	0	112.3215988	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
11.8	0	105.7451768	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
224	0	135.6708158	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
5900	0	104.4925766	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
1.02	0.0421349947	106.0122119	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
287	0.2145814985	112.1037827	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
660	0.2142053745	103.3780593	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
28	0	101.5598786	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
1.14	0.0314115445	122.7699165	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
5900	0	102.650856	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
1	0	118.0599601	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
800	0.0410746435	108.3272124	sp|C0HJP6|CO1A2_MACSX	2	Macrauchenia
650	0	101.2870187	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
620	0.8832961223	128.7368515	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
630	0.5781735486	136.288288	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
700	1	128.626925	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
7200	0.6429058315	161.2051061	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
224	0.5351133472	102.8797931	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
-1	0.6898266624	130.6621005	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
10000	0.6496695442	116.9191278	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
640	0.0629314318	120.9563971	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
11.8	0	129.5438759	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
1.26	0.4074646877	151.2613044	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
4500	0.8385773105	119.5501041	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
21.1	0.7182704969	112.7342629	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
28	0	106.4285369	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
1.14	0.4827820661	123.1970579	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
610	0.8035037109	129.8566131	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
-1	0.9799806512	178.2702812	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
0.96	0	106.4285369	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
1	0.1627915271	234.5037627	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon

-1	0.0308738854	130.6621005	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
750	0.7844011499	119.9443362	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
7600	0.5973192187	182.7751371	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
620	0.9511000778	113.8700083	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
630	0.6883787625	273.8341799	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
1.18	0	165.1873114	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
610	0.9719455534	122.2352385	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
750	1	106.4983274	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
800	0.8560461919	109.5706138	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
1.02	0.1138503233	114.8173007	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
1.14	0.2941211993	153.4926563	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
-1	0.9251224303	100.7382996	sp|C0HJP3|CO1A1_MYLDA	Default	Myldodon
1.08	0.2302714378	211.1278171	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
10000	0.4946306178	155.6893399	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
640	0.9343696364	116.8603029	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
5200	0.7757353551	132.1319043	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
5900	0.8819515677	105.743872	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
660	0.9262072998	123.474223	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
224	0.3536250957	251.3901913	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
57.8	0	105.8056566	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
21.1	0	105.8056566	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
287	0.9724700715	228.7600831	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
48.2	0.9401321321	135.3036116	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
1	0	101.1653	sp|C0HJP4|CO1A2_MYLDA	Default	Myldodon
650	1	157.3842898	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
620	0.9960264477	176.219017	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
10000	0.9527504753	152.4166524	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
5900	0.9822583598	171.2289744	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
7200	0.9894951063	156.0714531	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
-1	0.9664153341	117.658747	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
224	0.9420133084	141.7727725	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
630	0.9968682904	158.5090839	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
640	1	333.5770672	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
21.1	0.8835408234	166.3540414	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
1.26	0.075042904	149.1575816	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
4500	1	109.3963647	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
1.45	0	109.6160485	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
28	0.8106129799	106.4491124	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
1.14	0	109.3963647	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
610	1	119.5148069	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
1	0.0467395531	281.0762879	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
750	0.9792276663	183.6179858	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
7600	1	127.5736038	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
620	0.9850377126	111.7895619	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
630	0.9966050646	145.9355081	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
-1	1	125.3094106	sp|C0HJN7|CO1A1_TAPTE	Default	Tapir
1.18	0	116.6860946	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
224	0.7986699348	118.8660172	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
10000	0.9722595903	112.4616704	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir

57.8	0.5693277856	107.4595093	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
5900	0.8223359982	107.2889743	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
1.02	0	109.0805577	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
287	0.9307131967	121.9485619	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
48.2	1	117.2990857	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
1.14	0.0971086033	112.3919428	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
1	0.0543576985	106.1063645	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
800	1	110.7377286	sp|C0HJN8|CO1A2_TAPTE	Default	Tapir
650	0	102.4711043	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
620	0.5163980824	104.989938	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
630	0.5843206013	129.8887497	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
700	0.6417927266	102.5236737	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
7200	0.311395869	124.0336356	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
224	0.4501448332	134.3716892	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
10000	0.1453556131	112.2828393	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
-1	0.1573326118	110.7259973	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
640	0.4293543901	138.5958445	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
21.1	0.1576520408	136.0126976	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
1.26	0.2975716707	141.0090303	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
4500	1	102.1588046	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
1.45	0	102.6703102	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
1.14	0	102.1588046	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
5900	0.3563108916	134.5376756	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
1	0.1642305945	350.5322162	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
750	0.4458995817	107.1698018	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
7600	0.0378151105	143.9806142	sp|C0HJP7|CO1A1_TOXSP	1	Toxodon
10000	0.3059134933	108.1588597	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1.18	0.1359478045	112.4447511	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1.08	0.3011074672	104.9632467	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1.45	0	103.7747589	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
-1	0.373108905	109.0499075	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
57.8	0	110.5107719	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
660	1	101.8225576	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
630	0.4980812483	158.0812943	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1.02	0	102.789801	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
287	0.1663780234	141.9974833	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
48.2	0.0842967429	104.3269258	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1.14	0	111.5653138	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
1	0.0941951067	116.7477177	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
800	0.3630671169	116.3273997	sp|C0HJP8|CO1A2_TOXSP	1	Toxodon
620	0.7577877804	108.1951155	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
630	0.581862357	122.7220577	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
700	0.8059515009	102.994022	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
7200	0.4143256064	121.9868886	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
224	0.6151614247	143.3503726	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
10000	0.2278770323	116.0657647	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
-1	0.5407783049	105.4530823	sp|C0HJP7|CO1A1_TOXSP	3	Toxodon
640	0.5680476266	146.5451953	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
21.1	0.2493232925	133.1297506	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon

660	1	114.4842348	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
1.26	0.153990491	142.6791015	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
4500	1	102.1445687	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
1.45	0	103.2301758	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
5900	0.437999251	129.5278469	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
1.14	0	102.1445687	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
610	0.4426384514	109.2289146	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
1	0.0624839252	267.6121714	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
750	0.6261008436	109.5759704	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
7600	0.0413811085	141.5967745	sp|C0HJP7|CO1A1_TOXSP	2	Toxodon
620	1	102.492717	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
630	0.580419274	139.5360091	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1.18	0.1056819095	116.078838	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1.08	0.1092162257	108.2684206	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1.45	0	104.4765922	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
-1	0.4601030314	111.1314998	sp|C0HJP8|CO1A2_TOXSP	3	Toxodon
57.8	0	128.566246	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
660	1	101.7664874	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
5900	0	102.492717	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1.02	0.1138361182	102.7691094	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
287	0.2639663997	133.3020608	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
10000	0.5518323687	113.7631722	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
48.2	0.4503622844	105.6637797	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1.14	0	115.9201188	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
1	0.1138060685	113.0865815	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon
800	0.5529101445	115.0062791	sp|C0HJP8|CO1A2_TOXSP	2	Toxodon

Appendix A, Table 3: Site-specific deamidation in model dataset, with the half-time of each detected site and the relative remaining proportion.

Appendix A, Table 4

Sample	Protein	Name	Classification	AA	RelNonDeam
sample	sp|P01009|A1AT_HUMAN	Alpha-1-antitrypsin	Human	N	0
sample	sp|P01009|A1AT_HUMAN	Alpha-1-antitrypsin	Human	Q	0
sample	sp|P18626|CASA1_CAPHI	Alpha-S1-casein	Milk	Q	0.9716493569
sample	sp|P02663|CASA2_BOVIN	Alpha-S2-casein	Milk	N	1
sample	sp|P02663|CASA2_BOVIN	Alpha-S2-casein	Milk	N	1
sample	sp|P01008|ANT3_HUMAN	Antithrombin-III	Human	N	0
sample	sp|P11839|CASB_SHEEP	Beta-casein	Milk	Q	1
sample	sp|P11839|CASB_SHEEP	Beta-casein	Milk	Q	0.9094294858
sample	sp|P67975|LACB_OVIMU	Beta-lactoglobulin	Milk	Q	0
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	N	0
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	N	0.9488883051
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	N	0.9282140981
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	Q	0.7971488839
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	Q	0
sample	sp|P08311|CATG_HUMAN	Cathepsin G	Human	Q	0.5217309946
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	N	1
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	N	0.968129779
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	N	0.7713397434
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	N	1
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	N	0.8996313136
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	Q	0.990453435
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	Q	0.3081060351
sample	sp|P01024|CO3_HUMAN	Complement C3	Human	Q	1
sample	sp|P01834|IGKC_HUMAN	Immunoglobulin kappa constant	Human	N	1
sample	sp|P01834|IGKC_HUMAN	Immunoglobulin kappa constant	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	0.4261987933
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	N	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	Q	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	Q	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	Q	1
sample	sp|P02788-2|TRFL_HUMAN	Lactotransferrin	Human	Q	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	N	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	N	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	N	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	N	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	N	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	Q	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	Q	1
sample	sp|P00698|LYSC_CHICK	Lysozyme C	Egg	Q	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	0.3760597034
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	0.5634505081
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	0.4078653635
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	0
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	N	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.9254657977

sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.7106929833
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.8534342737
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	1
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.3595079386
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.8277548451
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.7967354427
sample	sp|P05164|PERM_HUMAN	Myeloperoxidase	Human	Q	0.5435243133
sample	sp|P59665|DEF1_HUMAN	Neutrophil defensin 1	Human	Q	0.9474602552
sample	sp|P08246|ELNE_HUMAN	Neutrophil elastase	Human	N	0
sample	sp|P08246|ELNE_HUMAN	Neutrophil elastase	Human	N	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	N	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	N	0.1616466701
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	N	0.9385041731
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	N	0.9780547718
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	Q	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	Q	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	Q	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	Q	1
sample	sp|P01012|OVAL_CHICK	Ovalbumin	Egg	Q	1
sample	sp|P01014|OVALY_CHICK	Ovalbumin-related protein Y	Egg	N	1
sample	sp|P01014|OVALY_CHICK	Ovalbumin-related protein Y	Egg	N	1
sample	sp|P02789|TRFE_CHICK	Ovotransferrin	Egg	N	1
sample	sp|P02789|TRFE_CHICK	Ovotransferrin	Egg	N	1
sample	sp|P02789|TRFE_CHICK	Ovotransferrin	Egg	N	1
sample	sp|P02789|TRFE_CHICK	Ovotransferrin	Egg	Q	0
sample	sp|P05109|S10A8_HUMAN	Protein S100-A8	Human	Q	0
sample	sp|P06702|S10A9_HUMAN	Protein S100-A9	Human	N	1
sample	sp|P06702|S10A9_HUMAN	Protein S100-A9	Human	N	1
sample	sp|P06702|S10A9_HUMAN	Protein S100-A9	Human	Q	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	N	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	N	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	N	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	N	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	Q	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	Q	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	Q	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	Q	1
sample	sp|P87498|VIT1_CHICK	Vitellogenin-1	Egg	Q	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	0
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	N	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	1
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	0
sample	sp|P02845|VIT2_CHICK	Vitellogenin-2	Egg	Q	1
postblank	sp|P67975|LACB_OVIMU	Beta-lactoglobulin	Milk	Q	0

postblank	sp|P67975|LACB_OVIMU	Beta-lactoglobulin	Milk	N	0
postblank	sp|P01024|CO3_HUMAN	Complement C3	Human	Q	1
postblank	sp|P01024|CO3_HUMAN	Complement C3	Human	N	0.9024665127

Appendix A, Table 4: Bulk deamidation of protein groups in SK1.

Appendix A, Table 5

Half-time	RelNonDeam	Classification	Size	Sample	Protein
104	0	Milk	1000	postblank	sp|P67975|LACB_OVIMU
14.5	0.9024665127	Human	111.13941	postblank	sp|P01024|CO3_HUMAN
6400	1	Human	111.13941	postblank	sp|P01024|CO3_HUMAN
7700	1	Egg	103.1936076	sample	sp|P02845|VIT2_CHICK
16.4	0	Egg	103.1936076	sample	sp|P02845|VIT2_CHICK
391	1	Egg	103.3246287	sample	sp|P02845|VIT2_CHICK
1.53	1	Egg	103.0400062	sample	sp|P02845|VIT2_CHICK
83.9	1	Egg	105.8402345	sample	sp|P02845|VIT2_CHICK
-1	1	Egg	103.0408425	sample	sp|P02845|VIT2_CHICK
5500	1	Egg	103.648	sample	sp|P02845|VIT2_CHICK
-1	1	Egg	103.0408425	sample	sp|P02845|VIT2_CHICK
287	1	Egg	103.3246287	sample	sp|P02845|VIT2_CHICK
41	1	Egg	102.5337629	sample	sp|P02845|VIT2_CHICK
6200	0	Egg	101.9564336	sample	sp|P02845|VIT2_CHICK
6200	1	Egg	103.0400062	sample	sp|P02845|VIT2_CHICK
7900	1	Egg	102.5337629	sample	sp|P02845|VIT2_CHICK
18.9	1	Egg	103.5176294	sample	sp|P02789|TRFE_CHICK
5200	0	Egg	108.2078143	sample	sp|P02789|TRFE_CHICK
366	1	Egg	118.0651567	sample	sp|P02789|TRFE_CHICK
76.4	1	Egg	103.5176294	sample	sp|P02789|TRFE_CHICK
8300	1	Milk	102.6468425	sample	sp|P11839|CASB_SHEEP
6200	0.9094294858	Milk	100.9812275	sample	sp|P11839|CASB_SHEEP
22.1	1	Egg	107.9004257	sample	sp|P01014|OVALY_CHICK
14.5	1	Egg	119.5807274	sample	sp|P00698|LYSC_CHICK
4300	1	Egg	110.168578	sample	sp|P00698|LYSC_CHICK
4500	1	Egg	102.3617628	sample	sp|P00698|LYSC_CHICK
52.2	1	Egg	110.168578	sample	sp|P00698|LYSC_CHICK
6500	1	Egg	119.5807274	sample	sp|P00698|LYSC_CHICK
38.1	1	Egg	110.168578	sample	sp|P00698|LYSC_CHICK
58.8	1	Egg	110.168578	sample	sp|P00698|LYSC_CHICK
14.5	0.9385041731	Egg	131.7163694	sample	sp|P01012|OVAL_CHICK
1.04	0.1616466701	Egg	131.7163694	sample	sp|P01012|OVAL_CHICK
4300	1	Egg	103.0753169	sample	sp|P01012|OVAL_CHICK
5300	1	Egg	183.4364632	sample	sp|P01012|OVAL_CHICK
61.5	0.9780547718	Egg	183.4364632	sample	sp|P01012|OVAL_CHICK
6400	1	Egg	102.3131642	sample	sp|P01012|OVAL_CHICK
58.6	1	Egg	102.3131642	sample	sp|P01012|OVAL_CHICK
7100	1	Egg	183.4364632	sample	sp|P01012|OVAL_CHICK
5400	1	Egg	131.7163694	sample	sp|P01012|OVAL_CHICK
224	1	Egg	102.2359453	sample	sp|P87498|VIT1_CHICK
4500	1	Egg	102.2359453	sample	sp|P87498|VIT1_CHICK
83.1	1	Egg	101.2377687	sample	sp|P87498|VIT1_CHICK
7300	1	Egg	101.2377687	sample	sp|P87498|VIT1_CHICK
6200	1	Egg	102.2359453	sample	sp|P87498|VIT1_CHICK
-1	1	Egg	107.9004257	sample	sp|P01014|OVALY_CHICK
5300	1	Egg	100.1973587	sample	sp|P87498|VIT1_CHICK
4300	0.9716493569	Milk	119.0221685	sample	sp|P18626|CASA1_CAPHI
14000	1	Human	107.6775968	sample	sp|P06702|S10A9_HUMAN

73.9	1	Milk	115.6305796	sample	sp|P02663|CASA2_BOVIN
46.3	1	Human	107.6775968	sample	sp|P06702|S10A9_HUMAN
12.8	1	Human	102.1140306	sample	sp|P02788-2|TRFL_HUMAN
6200	1	Human	106.4657907	sample	sp|P02788-2|TRFL_HUMAN
2000	1	Human	106.4657907	sample	sp|P02788-2|TRFL_HUMAN
8200	1	Human	106.2803168	sample	sp|P02788-2|TRFL_HUMAN
48.7	1	Human	108.7899755	sample	sp|P02788-2|TRFL_HUMAN
48.9	1	Human	106.4657907	sample	sp|P02788-2|TRFL_HUMAN
5200	0	Milk	102.4825624	sample	sp|P67975|LACB_OVIMU
59.7	1	Human	101.0969907	sample	sp|P02788-2|TRFL_HUMAN
50.2	1	Human	101.0969907	sample	sp|P02788-2|TRFL_HUMAN
18.9	0.4261987933	Human	101.0969907	sample	sp|P02788-2|TRFL_HUMAN
-1	1	Human	108.3339105	sample	sp|P05164|PERM_HUMAN
28.4	1	Human	108.7899755	sample	sp|P02788-2|TRFL_HUMAN
6000	1	Human	101.4766732	sample	sp|P02788-2|TRFL_HUMAN
-1	1	Human	106.9659631	sample	sp|P05164|PERM_HUMAN
10000	0	Human	166.0830083	sample	sp|P05109|S10A8_HUMAN
14.5	0.968129779	Human	137.4405805	sample	sp|P01024|CO3_HUMAN
24	0.8996313136	Human	108.3400168	sample	sp|P01024|CO3_HUMAN
226	1	Human	105.5048762	sample	sp|P01024|CO3_HUMAN
7200	1	Human	102.0339312	sample	sp|P01024|CO3_HUMAN
72.4	1	Human	105.5048762	sample	sp|P01024|CO3_HUMAN
630	0.3081060351	Human	107.5812731	sample	sp|P01024|CO3_HUMAN
6400	0.990453435	Human	137.4405805	sample	sp|P01024|CO3_HUMAN
74	0.7713397434	Human	108.3400168	sample	sp|P01024|CO3_HUMAN
38.9	1	Human	104.3162078	sample	sp|P01008|ANT3_HUMAN
-1	1	Human	102.7202589	sample	sp|P05164|PERM_HUMAN
233	1	Human	109.2644393	sample	sp|P08246|ELNE_HUMAN
119	0	Human	101.0706936	sample	sp|P08246|ELNE_HUMAN
850	0.9474602552	Human	244.131355	sample	sp|P59665|DEF1_HUMAN
11000	1	Human	102.7202589	sample	sp|P05164|PERM_HUMAN
6800	0.9254657977	Human	113.7629245	sample	sp|P05164|PERM_HUMAN
247	0.5634505081	Human	121.1543632	sample	sp|P05164|PERM_HUMAN
59.7	0.4078653635	Human	106.1690511	sample	sp|P05164|PERM_HUMAN
22.1	0.3760597034	Human	121.1543632	sample	sp|P05164|PERM_HUMAN
83.1	1	Human	102.7202589	sample	sp|P05164|PERM_HUMAN
25.9	1	Human	104.2300684	sample	sp|P05164|PERM_HUMAN
6200	0.7967354427	Human	104.9268267	sample	sp|P05164|PERM_HUMAN
1.05	0	Human	112.1859318	sample	sp|P05164|PERM_HUMAN
6200	0.8277548451	Human	112.4276859	sample	sp|P05164|PERM_HUMAN
6200	0.8534342737	Human	121.2734786	sample	sp|P05164|PERM_HUMAN
59.7	1	Human	100.9288934	sample	sp|P05164|PERM_HUMAN
8900	0.5435243133	Human	112.4276859	sample	sp|P05164|PERM_HUMAN
8200	0.7106929833	Human	104.9268267	sample	sp|P05164|PERM_HUMAN
-1	0.9488883051	Human	109.2344385	sample	sp|P08311|CATG_HUMAN
6200	0	Human	100.0692554	sample	sp|P01009|A1AT_HUMAN
-1	1	Egg	101.2377687	sample	sp|P87498|VIT1_CHICK
181	0	Human	100.8818744	sample	sp|P01009|A1AT_HUMAN
5500	0.7971488839	Human	106.6294092	sample	sp|P08311|CATG_HUMAN

6100	0	Human	104.1946648	sample	sp|P08311|CATG_HUMAN
7300	0.5217309946	Human	109.0690227	sample	sp|P08311|CATG_HUMAN
-1	1	Egg	101.2377687	sample	sp|P87498|VIT1_CHICK
62.4	0	Human	102.40841	sample	sp|P08311|CATG_HUMAN
254	0.8669749923	Human	105.6782391	sample	sp|P08311|CATG_HUMAN
-1	1	Human	105.5648114	sample	sp|P01834|IGKC_HUMAN
-1	1	Human	105.5648114	sample	sp|P01834|IGKC_HUMAN

Appendix A, Table 5: Site-specific deamidation of protein groups in SK1.

Appendix A, File 1

>sp|P61626|LYSC_HUMAN Lysozyme C OS=Homo sapiens OX=9606 GN=LYZ PE=1 SV=1
MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWESGYNTRA
TNYNAGDRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVRD
PQGIRAWVAWRNRCQNRDVRQYVQGCGV
>sp|P02788|TRFL_HUMAN Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF PE=1 SV=6
MKLVFLVLLFLGALGLCLAGRRRSVQWCAVSQPEATKCFQWQRNMRKVRGPPVSCIKRDS
PIQCIQAIAENRADAVTLDGGFIYEAGLAPYKLRPVAAEVYGTERQPRTHYYAVAVVKKG
GSFQLNELQGLKSCHTGLRRTAGWNVPIGTLRPFLNWTGPPEPIEAAVARFFSASCVPGA
DKGQFPNLCRLCAGTGENKCAFSSQEPYFSYSGAFKCLRDGAGDVAFIRESTVFEDLSDE
AERDEYELLCPDNTRKPVDKFKDCHLARVPSHAVVARSVNGKEDAIWNLLRQAQEKFGKD
KSPKFQLFGSPSGQKDLLFKDSAIGFSRVPPRIDSGLYLGSGYFTAIQNLRKSEEEVAAR
RARVVWCAVGEQELRKCNQWSGLSEGSVTCSSASTTEDCIALVLKGEADAMSLDGGYVYT
AGKCGLVPVLAENYKSQQSSDPDPNCVDRPVEGYLAVAVVRRSDTSLTWNSVKGKKSCHT
AVDRTAGWNIPMGLLFNQTGSCKFDEYFSQSCAPGSDPRSNLCALCIGDEQGENKCVPNS
NERYYGYTGAFRCLAENAGDVAFVKDVTVLQNTDGNNNEAWAKDLKLADFALLCLDGKRK
PVTEARSCHLAMAPNHAVVSRMDKVERLKQVLLHQQAKFGRNGSDCPDKFCLFQSETKNL
LFNDNTECLARLHGKTTYEKYLGPQYVAGITNLKKCSTSPLLEACEFLRK
>sp|P02788-2|TRFL_HUMAN Isoform DeltaLf of Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF
MRKVRGPPVSCIKRDSPIQCIQAIAENRADAVTLDGGFIYEAGLAPYKLRPVAAEVYGTE
RQPRTHYYAVAVVKKGGSFQLNELQGLKSCHTGLRRTAGWNVPIGTLRPFLNWTGPPEPI
EAAVARFFSASCVPGADKGQFPNLCRLCAGTGENKCAFSSQEPYFSYSGAFKCLRDGAGD
VAFIRESTVFEDLSDEAERDEYELLCPDNTRKPVDKFKDCHLARVPSHAVVARSVNGKED
AIWNLLRQAQEKFGKDKSPKFQLFGSPSGQKDLLFKDSAIGFSRVPPRIDSGLYLGSGYF
TAIQNLRKSEEEVAARRARVVWCAVGEQELRKCNQWSGLSEGSVTCSSASTTEDCIALVL
KGEADAMSLDGGYVYTAGKCGLVPVLAENYKSQQSSDPDPNCVDRPVEGYLAVAVVRRSD
TSLTWNSVKGKKSCHTAVDRTAGWNIPMGLLFNQTGSCKFDEYFSQSCAPGSDPRSNLCA
LCIGDEQGENKCVPNSNERYYGYTGAFRCLAENAGDVAFVKDVTVLQNTDGNNNEAWAKD
LKLADFALLCLDGKRKPVTEARSCHLAMAPNHAVVSRMDKVERLKQVLLHQQAKFGRNGS
DCPDKFCLFQSETKNLLFNDNTECLARLHGKTTYEKYLGPQYVAGITNLKKCSTSPLLEA

CEFLRK

>sp|P01834|IGKC_HUMAN Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD
SKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>sp|P08246|ELNE_HUMAN Neutrophil elastase OS=Homo sapiens OX=9606 GN=ELANE PE=1 SV=1

MTLGRRLACLFLACVLPALLLGGTALASEIVGGRRARPHAWPFMVSLQLRGGHFCGATLI
APNFVMSAAHCVANVNVRAVRVVLGAHNLSRREPTRQVFAVQRIFENGYDPVNLLNDIVI
LQLNGSATINANVQVAQLPAQGRRLGNGVQCLAMGWGLLGRNRGIASVLQELNVTVVTSL
CRRSNVCTLVRGRQAGVCFGDSGSPLVCNGLIHGIASFVRGGCASGLYPDAFAPVAQFVN
WIDSIIQRSEDNPCPHPRDPDPASRTH

>sp|P08311|CATG_HUMAN Cathepsin G OS=Homo sapiens OX=9606 GN=CTSG PE=1 SV=2

MQPLLLLLAFLLPTGAEAGEIIGGRESRPHSRPYMAYLQIQSPAGQSRCGGFLVREDFVL
TAAHCWGSNINVTLGAHNIQRRENTQQHITARRAIRHPQYNQRTIQNDIMLLQLSRRVRR
NRNVNPVALPRAQEGLRPGTLCTVAGWGRVSMRRGTDTLREVQLRVQRDRQCLRIFGSYD
PRRQICVGDRRERKAAFKGDSGGPLLCNNVAHGIVSYGKSSGVPPEVFTRVSSFLPWIRT
TMRSFKLLDQMETPL

>sp|P01008|ANT3_HUMAN Antithrombin-III OS=Homo sapiens OX=9606 GN=SERPINC1 PE=1 SV=1

MYSNVIGTVTSGKRKVYLLSLLLIGFWDCVTCHGSPVDICTAKPRDIPMNPMCIYRSPEK
KATEDEGSEQKIPEATNRRVWELSKANSRFATTFYQHLADSKNDNDNIFLSPLSISTAFA
MTKLGACNDTLQQLMEVFKFDTISEKTSDQIHFFFAKLNCRLYRKANKSSKLVSANRLFG
DKSLTFNETYQDISELVYGAKLQPLDFKENAEQSRAAINKWVSNKTEGRITDVIPSEAIN
ELTVLVLVNTIYFKGLWKSKFSPENTRKELFYKADGESCSASMMYQEGKFRYRRVAEGTQ
VLELPFKGDDITMVLILPKPEKSLAKVEKELTPEVLQEWLDELEEMMLVVHMPRFRIEDG
FSLKEQLQDMGLVDLFSPEKSKLPGIVAEGRDDLYVSDAFHKAFLEVNEEGSEAAASTAV
VIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVK

>sp|P01009|A1AT_HUMAN Alpha-1-antitrypsin OS=Homo sapiens OX=9606 GN=SERPINA1 PE=1 SV=3

MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFS
LYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGF
QELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQ
INDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTV

KVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFL
ENEDRRSASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLSKAVHKA
VLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK
>sp|P05164|PERM_HUMAN Myeloperoxidase OS=Homo sapiens OX=9606
GN=MPO PE=1 SV=1
MGVPFFSSLRCMVDLGPCWAGGLTAEMKLLLALAGLLAILATPQPSEGAAPAVLGEVDTS
LVLSSMEEAKQLVDKAYKERRESIKQRLRSGSASPMELLSYFKQPVAATRTAVRAADYLH
VALDLLERKLRSLWRRPFNVTDVLTPAQLNVLSKSSGCAYQDVGVTCPEQDKYRTITGMC
NNRRSPTLGASNRAFVRWLPAEYEDGFSLPYGWTPGVKRNGFPVALARAVSNEIVRFPTD
QLTPDQERSLMFMQWGQLLDHDLDFTPEPAARASFVTGVNCETSCVQQPPCFPLKIPPND
PRIKNQADCIPFFRSCPACPGSNITIRNQINALTSFVDASMVYGSEEPLARNLRNMSNQL
GLLAVNQRFQDNGRALLPFDNLHDDPCLLTNRSARIPCFLAGDTRSSEMPELTSMHTLLL
REHNRLATELKSLNPRWDGERLYQEARKIVGAMVQIITYRDYLPLVLGPTAMRKYLPTYR
SYNDSVDPRIANVFTNAFRYGHTLIQPFMFRLDNRYQPMEPNPRVPLSRVFFASWRVVLE
GGIDPILRGLMATPAKLNRQNQIAVDEIRERLFEQVMRIGLDLPALNMQRSRDHGLPGYN
AWRRFCGLPQPETVGQLGTVLRNLKLARKLMEQYGTPNNIDIWMGGVSEPLKRKGRVGPL
LACIIGTQFRKLRDGDRFWWENEGVFSMQQRQALAQISLPRIICDNTGITTVSKNNIFMS
NSYPRDFVNCSTLPALNLASWREAS
>sp|P59665|DEF1_HUMAN Neutrophil defensin 1 OS=Homo sapiens
OX=9606 GN=DEFA1 PE=1 SV=1
MRTLAILAAILLVALQAQAEPLQARADEVAAAPEQIAADIPEVVVSLAWDESLAPKHPGS
RKNMACYCRIPACIAGERRYGTCIYQGRLWAFCC
>sp|P06702|S10A9_HUMAN Protein S100-A9 OS=Homo sapiens OX=9606
GN=S100A9 PE=1 SV=1
MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIE
HIMEDLDTNADKQLSFEEFIMLMARLTWASHEKMHEGDEGPGHHHKPGLGEGTP
>sp|P05109|S10A8_HUMAN Protein S100-A8 OS=Homo sapiens OX=9606
GN=S100A8 PE=1 SV=1
MLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLLETECPQYIRKKGADVWFKELDI
NTDGAVNFQEFLILVIKMGVAAHKKSHEESHKE
>sp|P01024|CO3_HUMAN Complement C3 OS=Homo sapiens OX=9606 GN=C3
PE=1 SV=2
MGPTSGPSLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVH
DFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRNKFVTVQATFGTQVVEKV
VLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPEGIPVKQDSL

SSQNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTEFEVKEYVLPSFEVIVEPTE
KFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEV
VLSRKVLLDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPIVTSPYQIHFTKT
PKYFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL
SITVRTKKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMD
RAHEAKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGA
SGQREVVADSVWVDVKDSCVGSLVVKSGQSEDRQPVPGQQMTLKIEGDHGARVVLVAVDK
GVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTAQRAEL
QCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFISLGEAC
KKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNVEDLKE
PPKNGISTKLMNIFLKDSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVV
RNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVI
VPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTVAVRTLDPERLGREGVQKE
DIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIGMTP
TVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTA
YVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKD
MALTAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMG
RLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQR
YYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLR
SEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDA
KNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSD
RNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDG
KLNKLCRDELCRCAEENCFIQKSDDKVTLEERLDKACEPGVDYVYKTRLVKVQLSNDFDE
YIMAIEQTIKSGSDEVQVGQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSY
IIGKDTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN>sp|P61626|LYSC_HUMAN
Lysozyme C OS=Homo sapiens OX=9606 GN=LYZ PE=1 SV=1
MKALIVLGLVLLSVTVQGKVFERCELARTLKRLGMDGYRGISLANWMCLAKWESGYNTRA
TNYNAGDRSTDYGIFQINSRYWCNDGKTPGAVNACHLSCSALLQDNIADAVACAKRVVRD
PQGIRAWVAWRNRCQNRDVRQYVQGCGV
>sp|P02788|TRFL_HUMAN Lactotransferrin OS=Homo sapiens OX=9606
GN=LTF PE=1 SV=6
MKLVFLVLLFLGALGLCLAGRRRSVQWCAVSQPEATKCFQWQRNMRKVRGPPVSCIKRDS
PIQCIQAIAENRADAVTLDGGFIYEAGLAPYKLRPVAAEVYGTERQPRTHYYAVAVVKKG
GSFQLNELQGLKSCHTGLRRTAGWNVPIGTLRPFLNWTGPPEPIEAAVARFFSASCVPGA
DKGQFPNLCRLCAGTGENKCAFSSQEPYFSYSGAFKCLRDGAGDVAFIRESTVFEDLSDE

AERDEYELLCPDNTRKPVDKFKDCHLARVPSHAVVARSVNGKEDAIWNLLRQAQEKFGKD
KSPKFQLFGSPSGQKDLLFKDSAIGFSRVPPRIDSGLYLGSGYFTAIQNLRKSEEEVAAR
RARVVWCAVGEQELRKCNQWSGLSEGSVTCSSASTTEDCIALVLKGEADAMSLDGGYVYT
AGKCGLVPVLAENYKSQQSSDPDPNCVDRPVEGYLAVAVVRRSDTSLTWNSVKGKKSCHT
AVDRTAGWNIPMGLLFNQTGSCKFDEYFSQSCAPGSDPRSNLCALCIGDEQGENKCVPNS
NERYYGYTGAFRCLAENAGDVAFVKDVTVLQNTDGNNNEAWAKDLKLADFALLCLDGKRK
PVTEARSCHLAMAPNHAVVSRMDKVERLKQVLLHQQAKFGRNGSDCPDKFCLFQSETKNL
LFNDNTECLARLHGKTTYEKYLGPQYVAGITNLKKCSTSPLLEACEFLRK
>sp|P02788-2|TRFL_HUMAN Isoform DeltaLf of Lactotransferrin OS=Homo sapiens OX=9606 GN=LTF
MRKVRGPPVSCIKRDSPIQCIQAIAENRADAVTLDGGFIYEAGLAPYKLRPVAAEVYGTE
RQPRTHYYAVAVVKKGGSFQLNELQGLKSCHTGLRRTAGWNVPIGTLRPFLNWTGPPEPI
EAAVARFFSASCVPGADKGQFPNLCRLCAGTGENKCAFSSQEPYFSYSGAFKCLRDGAGD
VAFIRESTVFEDLSDEAERDEYELLCPDNTRKPVDKFKDCHLARVPSHAVVARSVNGKED
AIWNLLRQAQEKFGKDKSPKFQLFGSPSGQKDLLFKDSAIGFSRVPPRIDSGLYLGSGYF
TAIQNLRKSEEEVAARRARVVWCAVGEQELRKCNQWSGLSEGSVTCSSASTTEDCIALVL
KGEADAMSLDGGYVYTAGKCGLVPVLAENYKSQQSSDPDPNCVDRPVEGYLAVAVVRRSD
TSLTWNSVKGKKSCHTAVDRTAGWNIPMGLLFNQTGSCKFDEYFSQSCAPGSDPRSNLCA
LCIGDEQGENKCVPNSNERYYGYTGAFRCLAENAGDVAFVKDVTVLQNTDGNNNEAWAKD
LKLADFALLCLDGKRKPVTEARSCHLAMAPNHAVVSRMDKVERLKQVLLHQQAKFGRNGS
DCPDKFCLFQSETKNLLFNDNTECLARLHGKTTYEKYLGPQYVAGITNLKKCSTSPLLEA
CEFLRK
>sp|P01834|IGKC_HUMAN Immunoglobulin kappa constant OS=Homo sapiens OX=9606 GN=IGKC PE=1 SV=2
RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD
SKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
>sp|P08246|ELNE_HUMAN Neutrophil elastase OS=Homo sapiens OX=9606 GN=ELANE PE=1 SV=1
MTLGRRLACLFLACVLPALLLGGTALASEIVGGRRARPHAWPFMVSLQLRGGHFCGATLI
APNFVMSAAHCVANVNVRAVRVVLGAHNLSRREPTRQVFAVQRIFENGYDPVNLLNDIVI
LQLNGSATINANVQVAQLPAQGRRLGNGVQCLAMGWGLLGRNRGIASVLQELNVTVVTSL
CRRSNVCTLVRGRQAGVCFGDSGSPLVCNGLIHGIASFVRGGCASGLYPDAFAPVAQFVN
WIDSIIQRSEDNPCPHPRDPDPASRTH
>sp|P08311|CATG_HUMAN Cathepsin G OS=Homo sapiens OX=9606 GN=CTSG PE=1 SV=2

MQPLLLLLAFLLPTGAEAGEIIGGRESRPHSRPYMAYLQIQSPAGQSRCGGFLVREDFVL
TAAHCWGSNINVTLGAHNIQRRENTQQHITARRAIRHPQYNQRTIQNDIMLLQLSRRVRR
NRNVNPVALPRAQEGLRPGTLCTVAGWGRVSMRRGTDTLREVQLRVQRDRQCLRIFGSYD
PRRQICVGDRRERKAAFKGDSGGPLLCNNVAHGIVSYGKSSGVPPEVFTRVSSFLPWIRT
TMRSFKLLDQMETPL
>sp|P01008|ANT3_HUMAN Antithrombin-III OS=Homo sapiens OX=9606
GN=SERPINC1 PE=1 SV=1
MYSNVIGTVTSGKRKVYLLSLLLIGFWDCVTCHGSPVDICTAKPRDIPMNPMCIYRSPEK
KATEDEGSEQKIPEATNRRVWELSKANSRFATTFYQHLADSKNDNDNIFLSPLSISTAFA
MTKLGACNDTLQQLMEVFKFDTISEKTSDQIHFFFAKLNCRLYRKANKSSKLVSANRLFG
DKSLTFNETYQDISELVYGAKLQPLDFKENAEQSRAAINKWVSNKTEGRITDVIPSEAIN
ELTVLVLVNTIYFKGLWKSKFSPENTRKELFYKADGESCSASMMYQEGKFRYRRVAEGTQ
VLELPFKGDDITMVLILPKPEKSLAKVEKELTPEVLQEWLDELEEMMLVVHMPRFRIEDG
FSLKEQLQDMGLVDLFSPEKSKLPGIVAEGRDDLYVSDAFHKAFLEVNEEGSEAAASTAV
VIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVK
>sp|P01009|A1AT_HUMAN Alpha-1-antitrypsin OS=Homo sapiens OX=9606
GN=SERPINA1 PE=1 SV=3
MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFS
LYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGF
QELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQ
INDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTV
KVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFL
ENEDRRSASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLSKAVHKA
VLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK
>sp|P05164|PERM_HUMAN Myeloperoxidase OS=Homo sapiens OX=9606
GN=MPO PE=1 SV=1
MGVPFFSSLRCMVDLGPCWAGGLTAEMKLLLALAGLLAILATPQPSEGAAPAVLGEVDTS
LVLSSMEEAKQLVDKAYKERRESIKQRLRSGSASPMELLSYFKQPVAATRTAVRAADYLH
VALDLLERKLRSLWRRPFNVTDVLTPAQLNVLSKSSGCAYQDVGVTCPEQDKYRTITGMC
NNRRSPTLGASNRAFVRWLPAEYEDGFSLPYGWTPGVKRNGFPVALARAVSNEIVRFPTD
QLTPDQERSLMFMQWGQLLDHDLDFTPEPAARASFVTGVNCETSCVQQPPCFPLKIPPND
PRIKNQADCIPFFRSCPACPGSNITIRNQINALTSFVDASMVYGSEEPLARNLRNMSNQL
GLLAVNQRFQDNGRALLPFDNLHDDPCLLTNRSARIPCFLAGDTRSSEMPELTSMHTLLL
REHNRLATELKSLNPRWDGERLYQEARKIVGAMVQIITYRDYLPLVLGPTAMRKYLPTYR
SYNDSVDPRIANVFTNAFRYGHTLIQPFMFRLDNRYQPMEPNPRVPLSRVFFASWRVVLE

GGIDPILRGLMATPAKLNRQNQIAVDEIRERLFEQVMRIGLDLPALNMQRSRDHGLPGYN
AWRRFCGLPQPETVGQLGTVLRNLKLARKLMEQYGTPNNIDIWMGGVSEPLKRKGRVGPL
LACIIGTQFRKLRDGDRFWWENEGVFSMQQRQALAQISLPRIICDNTGITTVSKNNIFMS
NSYPRDFVNCSTLPALNLASWREAS
>sp|P59665|DEF1_HUMAN Neutrophil defensin 1 OS=Homo sapiens
OX=9606 GN=DEFA1 PE=1 SV=1
MRTLAILAAILLVALQAQAEPLQARADEVAAAPEQIAADIPEVVVSLAWDESLAPKHPGS
RKNMACYCRIPACIAGERRYGTCIYQGRLWAFCC
>sp|P06702|S10A9_HUMAN Protein S100-A9 OS=Homo sapiens OX=9606
GN=S100A9 PE=1 SV=1
MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIE
HIMEDLDTNADKQLSFEEFIMLMARLTWASHEKMHEGDEGPGHHHKPGLGEGTP
>sp|P05109|S10A8_HUMAN Protein S100-A8 OS=Homo sapiens OX=9606
GN=S100A8 PE=1 SV=1
MLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLLETECPQYIRKKGADVWFKELDI
NTDGAVNFQEFLILVIKMGVAAHKKSHEESHKE
>sp|P01024|CO3_HUMAN Complement C3 OS=Homo sapiens OX=9606 GN=C3
PE=1 SV=2
MGPTSGPSLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVH
DFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRNKFVTVQATFGTQVVEKV
VLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPEGIPVKQDSL
SSQNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTEFEVKEYVLPSFEVIVEPTE
KFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEV
VLSRKVLLDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPIVTSPYQIHFTKT
PKYFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL
SITVRTKKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMD
RAHEAKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGA
SGQREVVADSVWVDVKDSCVGSLVVKSGQSEDRQPVPGQQMTLKIEGDHGARVVLVAVDK
GVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTAQRAEL
QCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFISLGEAC
KKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNVEDLKE
PPKNGISTKLMNIFLKDSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVV
RNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVI
VPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTVAVRTLDPERLGREGVQKE
DIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIGMTP

TVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTA
YVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKD
MALTAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMG
RLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQR
YYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLR
SEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDA
KNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSD
RNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDG
KLNKLCRDELCRCAEENCFIQKSDDKVTLEERLDKACEPGVDYVYKTRLVKVQLSNDFDE
YIMAIEQTIKSGSDEVQVGQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSY
IIGKDTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN>sp|P02754|LACB_BOVIN
Beta-lactoglobulin OS=Bos taurus OX=9913 GN=LGB PE=1 SV=3
MKCLLLALALTCGAQALIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVE
ELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLL
FCMENSAEPEQSLACQCLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEEQCHI
>sp|P67975|LACB_OVIMU Beta-lactoglobulin OS=Ovis aries musimon
OX=9938 GN=LGB PE=1 SV=1
IIVTQTMKGLDIQKVAGTWHSLAMAASDISLLDAQSAPLRVYVEELKPTPEGNLEILLQK
WENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQ
CLVRTPEVDNEALEKFDKALKALPMHIRLAFNPTQLEGQCHV
>sp|P02756|LACB_CAPHI Beta-lactoglobulin OS=Capra hircus OX=9925
GN=LGB PE=1 SV=2
MKCLLLALGLALACGIQAIIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVY
VEELKPTPEGNLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKY
LLFCMENSAEPEQSLACQCLVRTPEVDKEALEKFDKALKALPMHIRLAFNPTQLEGQCHV
>sp|P02758|LACB1_HORSE Beta-lactoglobulin-1 OS=Equus caballus
OX=9796 GN=LGB1 PE=1 SV=3
MKCLLLALGLALMCGIQATNIPQTMQDLDLQEVAGKWHSVAMAASDISLLDSESAPLRVY
IEKLRPTPEDNLEIILREGENKGCAEKKIFAEKTESPAEFKINYLDEDTVFALDTDYKNY
LFLCMKNAATPGQSLVCQYLARTQMVDEEIMEKFRRALQPLPGRVQIVPDLTRMAERCRI
>sp|P02662|CASA1_BOVIN Alpha-S1-casein OS=Bos taurus OX=9913
GN=CSN1S1 PE=1 SV=2
MKLLILTCLVAVALARPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIG
SESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYK
VPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWY

YVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW

>sp|P04653|CASA1_SHEEP Alpha-S1-casein OS=Ovis aries OX=9940 GN=CSN1S1 PE=1 SV=3

MKLLILTCLVAVALARPKHPIKHQGLSSEVLNENLLRFVVAPFPEVFRKENINELSKDIG
SESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKITMPLW

>sp|P18626|CASA1_CAPHI Alpha-S1-casein OS=Capra hircus OX=9925 GN=CSN1S1 PE=1 SV=2

MKLLILTCLVAVALARPKHPINHRGLSPEVPNENLLRFVVAPFPEVFRKENINELSKDIG
SESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKTTMPLW

>sp|P47710-4|CASA1_HUMAN Isoform 4 of Alpha-S1-casein OS=Homo sapiens OX=9606 GN=CSN1S1

MRLLILTCLVAVALARPKLPLRYPERLQNPSESSEPIPLESREEYMNGMNRRNILREKQT
DEIKDTRNESTQNCVVAEPEKMESSISSSSEEQFCRLNEYNQLQLQAAHAQEQIRRMNEN
SHVQVPFQQLNQLAAYPYAVWYYPQIMQYVPFPPFSDISNPTAHENYEKNNVMLQW

>sp|P02663|CASA2_BOVIN Alpha-S2-casein OS=Bos taurus OX=9913 GN=CSN1S2 PE=1 SV=2

MKFFIFTCLLAVALAKNTMEHVSSSEESIISQETYKQEKNMAINPSKENLCSTFCKEVVR
NANEEEYSIGSSSEESAEVATEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPIV
LNPWDQVKRNAVPITPTLNREQLSTSEENSKKTVDMESTEVFTKKTKLTEEEKNRLNFLK
KISQRYQKFALPQYLKTVYQHQKAMKPWIQPKTKVIPYVRYL

>sp|P04654|CASA2_SHEEP Alpha-S2-casein OS=Ovis aries OX=9940 GN=CSN1S2 PE=2 SV=1

MKFFIFTCLLAVALAKHKMEHVSSSEEPINISQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNADEEEYSIRSSSEESAEVAPEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL
KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL

>sp|P33049|CASA2_CAPHI Alpha-S2-casein OS=Capra hircus OX=9925 GN=CSN1S2 PE=2 SV=1

MKFFIFTCLLAVALAKHKMEHVSSSEEPINIFQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNANEEEYSIRSSSEESAEVAPEEIKITVDDKHYQKALNEINQFYQKFPQYLQYPYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL

KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL

>sp|P00711|LALBA_BOVIN Alpha-lactalbumin OS=Bos taurus OX=9913 GN=LALBA PE=1 SV=2

MMSFVSLLLVGILFHATQAEQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQAI
VQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVGI
NYWLAHKALCSEKLDQWLCEKL

>sp|P00709|LALBA_HUMAN Alpha-lactalbumin OS=Homo sapiens OX=9606 GN=LALBA PE=1 SV=1

MRFFVPLFLVGILFPAILAKQFTKCELSQLLKDIDGYGGIALPELICTMFHTSGYDTQAI
VENNESTEYGLFQISNKLWCKSSQVPQSRNICDISCDKFLDDDITDDIMCAKKILDIKGI
DYWLAHKALCTEKLEQWLCEKL

>sp|P00712|LALBA_CAPHI Alpha-lactalbumin OS=Capra hircus OX=9925 GN=LALBA PE=1 SV=1

MMSFVSLLLVGILFHATQAEQLTKCEVFQKLKDLKDYGGVSLPEWVCTAFHTSGYDTQAI
VQNNDSTEYGLFQINNKIWCKDDQNPHSRNICNISCDKFLDDDLTDDIVCAKKILDKVGI
NYWLAHKALCSEKLDQWLCEKL

>sp|P09462|LALBA_SHEEP Alpha-lactalbumin OS=Ovis aries OX=9940 GN=LALBA PE=1 SV=2

MMSFVSLLLVGILFHATQAEQLTKCEVFQELKDLKDYGGVSLPEWVCTAFHTSGYDTQAI
VQNNDSTEYGLFQINNKIWCKDDQNPHSRNICNISCDKFLDDDLTDDIMCVKKILDKVGI
NYWLAHKALCSEKLDQWLCEKL

>sp|P02666|CASB_BOVIN Beta-casein OS=Bos taurus OX=9913 GN=CSN2 PE=1 SV=2

MKVLILACLVALALARELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDEL
QDKIHPFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK
HKEMPFPKYPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQPHQPLPPTVMFPPQSVLSL
SQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFPIIV

>sp|P05814|CASB_HUMAN Beta-casein OS=Homo sapiens OX=9606 GN=CSN2 PE=1 SV=4

MKVLILACLVALALARETIESLSSSEESITEYKQKVEKVKHEDQQQGEDEHQDKIYPSFQ
PQPLIYPFVEPIPYGFLPQNILPLAQPAVVLPVPQPEIMEVPKAKDTVYTKGRVMPVLKS
PTIPFFDPQIPKLTDLENLHLPLPLLQPLMQQVPQPIPQTLALPPQPLWSVPQPKVLPIP
QQVVPYPQRAVPVQALLLNQELLLNPTHQIYPVTQPLAPVHNPISV

>sp|Q9GKK3|CASB_HORSE Beta-casein OS=Equus caballus OX=9796 GN=CSN2 PE=1 SV=3

MKILILACLVALALAREKEELNVSSETVESLSSNEPDSSSEESITHINKEKLQKFKHEGQ
QQREVERQDKISRFVQPQPVVYPYAEPVPYAVVPQSILPLAQPPILPFLQPEIMEVSQAK
ETILPKRKVMPFLKSPIVPFSERQILNPTNGENLRLPVHLIQPFMHQVPQSLLQTLMLPS
QPVLSPPQSKVAPFPQPVVPYPQRDTPVQAFLLYQDPRLGPTGELDPATQPIVAVHNPVI
V
>sp|P11839|CASB_SHEEP Beta-casein OS=Ovis aries OX=9940 GN=CSN2
PE=1 SV=3
MKVLILACLVALALAREQEELNVVGETVESLSSSEESITHINKKIEKFQSEEQQQTEDEL
QDKIHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPKVKETMVPK
HKEMPFPKYPVEPFTESQSLTLTDVEKLHLPLPLVQSWMHQPPQPLPPTVMFPPQSVLSL
SQPKVLPVPQKAVPQRDMPIQAFLLYQEPVLGPVRGPFPILV
>sp|P01012|OVAL_CHICK Ovalbumin OS=Gallus gallus OX=9031
GN=SERPINB14 PE=1 SV=2
MGSIGAASMEFCFDVFKELKVHHANENIFYCPIAIMSALAMVYLGAKDSTRTQINKVVRF
DKLPGFGDSIEAQCGTSVNVHSSLRDILNQITKPNDVYSFSLASRLYAEERYPILPEYLQ
CVKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLQPSSVDSQTAMVLVNAIV
FKGLWEKAFKDEDTQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGTM
SMLVLLPDEVSGLEQLESIINFEKLTEWTSSNVMEERKIKVYLPRMKMEEKYNLTSVLMA
MGITDVFSSSANLSGISSAESLKISQAVHAAHAEINEAGREVVGSAEAGVDAASVSEEFR
ADHPFLFCIKHIATNAVLFFGRCVSP
>sp|P01014|OVALY_CHICK Ovalbumin-related protein Y OS=Gallus
gallus OX=9031 GN=SERPINB14B PE=1 SV=1
MDSISVTNAKFCFDVFNEMKVHHVNENILYCPLSILTALAMVYLGARGNTESQMKKVLHF
DSITGAGSTTDSQCGSSEYVHNLFKELLSEITRPNATYSLEIADKLYVDKTFSVLPEYLS
CARKFYTGGVEEVNFKTAAEEARQLINSWVEKETNGQIKDLLVSSSIDFGTTMVFINTIY
FKGIWKIAFNTEDTREMPFSMTKEESKPVQMMCMNNSFNVATLPAEKMKILELPYASGDL
SMLVLLPDEVSGLERIEKTINFDKLREWTSTNAMAKKSMKVYLPRMKIEEKYNLTSILMA
LGMTDLFSRSANLTGISSVDNLMISDAVHGVFMEVNEEGTEATGSTGAIGNIKHSLELEE
FRADHPFLFFIRYNPTNAILFFGRYWSP
>sp|P01013|OVALX_CHICK Ovalbumin-related protein X (Fragment)
OS=Gallus gallus OX=9031 GN=SERPINB14C PE=3 SV=1
QIKDLLVSSSTDLDTTLVLVNAIYFKGMWKTAFNAEDTREMPFHVTKQESKPVQMMCMNN
SFNVATLPAEKMKILELPFASGDLSMLVLLPDEVSDLERIEKTINFEKLTEWTNPNTMEK
RRVKVYLPQMKIEEKYNLTSVLMALGMTDLFIPSANLTGISSAESLKISQAVHGAFMELS
EDGIEMAGSTGVIEDIKHSPESEQFRADHPFLFLIKHNPTNTIVYFGRYWSP

>sp|P02845|VIT2_CHICK Vitellogenin-2 OS=Gallus gallus OX=9031 GN=VTG2 PE=1 SV=1

MRGIILALVLTLVGSQKFDIDPGFNSRRSYLYNYEGSMLNGLQDRSLGKAGVRLSSKLEI
SGLPENAYLLKVRSPQVEEYNGVWPRDPFTRSSKITQVISSCFTRLFKFEYSSGRIGNIY
APEDCPDLCVNIVRGILNMFQMTIKKSQNVYELQEAGIGGICHARYVIQEDRKNSRIYVT
RTVDLNNCQEKVQKSIGMAYIYPCPVDVMKERLTKGTTAFSYKLKQSDSGTLITDVSSRQ
VYQISPFNEPTGVAVMEARQQLTLVEVRSERGSAPDVPMQNYGSLRYRFPAVLPQMPLQL
IKTKNPEQRIVETLQHIVLNNQQDFHDDVSYRFLEVVQLCRIANADNLESIWRQVSDKPR
YRRWLLSAVSASGTTETLKFLKNRIRNDDLNYIQTLLTVSLTLHLLQADEHTLPIAADLM
TSSRIQKNPVLQQVACLGYSSVVNRYCSQTSACPKEALQPIHDLADEAISRGREDKMKLA
LKCIGNMGEPASLKRILKFLPISSSSAADIPVHIQIDAITALKKIAWKDPKTVQGYLIQI
LADQSLPPEVRMMACAVIFETRPALALITTIANVAMKESNMQVASFVYSHMKSLSKSRLP
FMYNISSACNIALKLLSPKLDSMSYRYSKVIRADTYFDNYRVGATGEIFVVNSPRTMFPS
AIISKLMANSAGSVADLVEVGIRVEGLADVIMKRNIPFAEYPTYKQIKELGKALQGWKEL
PTETPLVSAYLKILGQEVAFININKELLQQVMKTVVEPADRNAAIKRIANQIRNSIAGQW
TQPVWMGELRYVVPSCLGLPLEYGSYTTALARAAVSVEGKMTPPLTGDFRLSQLLESTMQ
IRSDLKPSLYVHTVATMGVNTEYFQHAVEIQGEVQTRMPMKFDAKIDVKLKNLKIETNPC
REETEIVVGRHKAFAVSRNIGELGVEKRTSILPEDAPLDVTEEPFQTSERASREHFAMQG
PDSMPRKQSHSSREDLRRSTGKRAHKRDICLKMHHIGCQLCFSRRSRDASFIQNTYLHKL
IGEHEAKIVLMPVHTDADIDKIQLEIQAGSRAAARIITEVNPESEEEDESSPYEDIQAKL
KRILGIDSMFKVANKTRHPKNRPSKKGNTVLAEFGTEPDAKTSSSSSSASSTATSSSSSS
ASSPNRKKPMDEEENDQVKQARNKDASSSSRSSKSSNSSKRSSSKSSNSSKRSSSSSSSS
SSSSRSSSSSSSSSSNSKSSSSSSKSSSSSSRSRSSSKSSSSSSSSSSSSSSKSSSSRSS
SSSSKSSSHHSHSHHSGHLNGSSSSSSSSRSVSHHSHEHHSGHLEDDSSSSSSSVLSKI
WGRHEIYQYRFRSAHRQEFPKRKLPGDRATSRYSSTRSSHDTSRAASWPKFLGDIKTPVL
AAFLHGISNNKKTGGLQLVVYADTDSVRPRVQVFVTNLTDSSKWKLCADASVRNAHKAVA
YVKWGWDCRDYKVSTELVTGRFAGHPAAQVKLEWPKVPSNVRSVVEWFYEFVPGAAFMLG
FSERMDKNPSRQARMVVALTSPRTCDVVVKLPDIILYQKAVRLPLSLPVGPRIPASELQP
PIWNVFAEAPSAVLENLKARCSVSYNKIKTFNEVKFNYSMPANCYHILVQDCSSELKFLV
MMKSAGEATNLKAINIKIGSHEIDMHPVNGQVKLLVDGAESPTANISLISAGASLWIHNE
NQGFALAAPGHGIDKLYFDGKTITIQVPLWMAGKTCGICGKYDAECEQEYRMPNGYLAKN
AVSFGHSWILEEAPCRGACKLHRSFVKLEKTVQLAGVDSKCYSTEPVLRCAKGCSATKTT
PVTVGFHCLPADSANSLTDKQMKYDQKSEDMQDTVDAHTTCSCENEECST

>sp|P87498|VIT1_CHICK Vitellogenin-1 OS=Gallus gallus OX=9031 GN=VTG1 PE=1 SV=1

MRGLISALVLTLVGSQHLNYQPDFGENKVYTYNYESILFSGIPEKGLARTGIRIRSEVEI
SGIGPKLCLIRIHSIEAAEYNGIWPTSSFSRSLKLTQALTGQLSIPIKFEYSNGHVGNLM
APDSVSDDGLNIYRGILNILELSLKKMQHSYSIQEAGIGGICNTTYAIQENKKANLVDVT
KSKDLNSCEEKVQVVTGSAYTQPCQTCQQRNKNSRATATYNYKIKYTHNEAVITQAEVEE
VHQFTPFHEITGGNAIVEARQKLALIEVQKQVAEVPPKEFQKRGSLQYQFGSELLQLPVH
LFKIKDVERQIEERLQDLVETTYEQLPSDAPAKALKLMHLLRAANEENYESVWKQFSSRP
AYRRYLLDLLPAAASHRSLRFLRHKMERQELTNWEIAQTVLVALHSSSPTQEVMEEATLI
VKKHCPRSSSVLRKVCLLSYASLCHKRCSSPYSCSECLQVFHVFAGEALGKSNIEEVLLA
LKALGNVGHPASIKHIKKFLPGYAAGASELPLKVHETAVMALKSIGMRDPQMVQAITLEI
FLNHKIHPRIRMLAAVVLLETKPGLPILMILVDAVLKEPSMQVASFIYSHLRALGRSTAP
DLQMMASACRMAVRALSPKFDRSGYQFSKVFRFSMFKEFLMSGLAAKYFVLNNAGSLIPT
MAVSQLRTHFLGRVADPIEVGIAAEGLQEMFVRGYSPDKDWETNYDFREILKKLSDWKAL
PRDKPFASGYLKMFGQELLFGRLDKDTLQNVLQVWYGPDEKIPSIRRLISSLQTGIGRQW
TKALLLSEIRCIVPTCVGFPMETSFYYSSVTKVAGNVQAQITPSPRSDFRLTELLNSNVR
LRSKMSLSMAKHMTFVIGINTNMIQAGLEAHTKVNAHVPVNVVATIQMKEKSIKAEIPPC
KEETNLIIVSSKTFAVTRNIEDLAASKMTPVLLPEAVPDIMKMSFDSDSASGETDNIRDR
QSVEDVSSGNSFSFGHPSSGKEPFIQSMCSNASTFGVQVCIEKKSVHAAFIRNVPLYNAI
GEHALRMSFKPVYSDVPIEKIQVTIQAGDQAPTKMVRLVTFEDPERQESSRKEVMKRVKK
ILDDTDNQATRNSRSSSSSASSISESSESTTSTPSSSDSDNRASQGDPQINLKSRQSKAN
EKKFYPFGDSSSSGSSSSSSSSSSSSSDSSSSSRSSSSSDSSSSSSSSSSSSSSKSKSSS
RSSKSNRSSSSSNSKDSSSSSSKSNSKGSSSSSSKASGTRQKAKKQSKTTSFPHASAAEG
ERSVHEQKQETQSSSSSSSRASSNSRSTSSSTSSSSESSGVSHRQWKQDREAETKRVKSQ
FNSHSSYDIPNEWETYLPKVYRLRFRSAHTHWHSGHRTSSSSSSSSESGSSHSNSSSSD
SSSRRSHMSDSSSSSSSHRHGEKAAHSSRRSPTSRAASAHHRPGSSLTRERNFLGDVIPP
GITIVAQAVRSDNRNQGYQATAYVRSDAAKVDVQLVVVQLAETNWKACADAVILPLKAQA
RMRWGKECRDYRIAALATTGQMARKLAVQLKVQWGIIPSWIKKTSTALMRYVPGVALVLG
FSEAHQRNPSRELIVRAVATSPRSIDTVIKVPGVTLYYQGLRVPFTLALGASSSSYETRD
ITAWNFLPEIASQIAQEDQSTCEVSKGDFKTFDRMSFTCSFNKSCNVVVAQDCTEHPKFI
ITTRKVDHQSLSREVHINTSSANITICPAADSSLLVTCNKESVLSDSGVSEYEKDNIKIY
KNGKTVIVEAPIHGLKNVNFDGEILKVTVASWMRGKTCGVCGNNDREKHNELLMPNHKLA
HSCSAFVHSWVLLEETCSGGCKLQRRYVKLNRNPTIDGEESTCYSVDPVLKCMKDCTPIE
KTSVKVGFHCFPKATAVSLLEWQRSSDKKSASEDVVESVDADIDCTCTGDCS
>sp|Q91025|VIT3_CHICK Vitellogenin-3 (Fragments) OS=Gallus gallus
OX=9031 GN=VTG3 PE=2 SV=2
MRGFILALVLALVGAQKHDLEPVFSTGKTYLYDYKGLILHGLPGKGLAAAGLKLTCRLEI

SRVSRSDHLLQIENVFKVANKTRHHKKWIHSRVKAAVTDLWEEPSATPLSSSSSTDSSAE
GEEPGNKRDKDEIWQFGKKYGADSSSSSSSSSTGSGSSKTCSSSREDSSRDKHCSVDSEY
FNQQADLPIYQFWFKPADEQDPRRKVQNSSISSSSSSSSDEGISTPVSQPMFLGDSKPPV
LAAVLRAIRRNEQPTGYQLVLYTDRQASRLRVQVFVSSITESDRWKLCADASVVNSHKAS
GTLKWGKDCQDYQVATQIATGQFAAHPAIQVKLEWSEVPSSVRKTAR

>sp|P00698|LYSC_CHICK Lysozyme C OS=Gallus gallus OX=9031 GN=LYZ PE=1 SV=1

MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQA
TNRNTDGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDG
NGMNAWVAWRNRCKGTDVQAWIRGCRL

>sp|P27042|LYG_CHICK Lysozyme g OS=Gallus gallus OX=9031 PE=2 SV=1

MLGKNDPMCLVLVLLGLTALLGICQGGTGCYGSVSRIDTTGASCRTAKPEGLSYCGVRAS
RTIAERDLGSMNKYKVLIKRVGEALCIEPAVIAGIISRESHAGKILKNGWGDRGNGFGLM
QVDKRYHKIEGTWNGEAHIRQGTRILIDMVKKIQRKFPRWTRDQQLKGGISAYNAGVGNV
RSYERMDIGTLHDDYSNDVVARAQYFKQHGY

>sp|P02789|TRFE_CHICK Ovotransferrin OS=Gallus gallus OX=9031 PE=1 SV=2

MKLILCTVLSLGIAAVCFAAPPKSVIRWCTISSPEEKKCNNLRDLTQQERISLTCVQKAT
YLDCIKAIANNEADAISLDGGQAFEAGLAPYKLKPIAAEVYEHTEGSTTSYYAVAVVKKG
TEFTVNDLQGKTSCHTGLGRSAGWNIPIGTLLHRGAIEWEGIESGSVEQAVAKFFSASCV
PGATIEQKLCRQCKGDPKTKCARNAPYSGYSGAFHCLKDGKGDVAFVKHTTVNENAPDQK
DEYELLCLDGSRQPVDNYKTCNWARVAAHAVVARDDNKVEDIWSFLSKAQSDFGVDTKSD
FHLFGPPGKKDPVLKDLLFKDSAIMLKRVPSLMDSQLYLGFEYYSAIQSMRKDQLTPSPR
ENRIQWCAVGKDEKSKCDRWSVVSNGDVECTVVDETKDCIIKIMKGEADAVALDGGLVYT
AGVCGLVPVMAERYDDESQCSKTDERPASYFAVAVARKDSNVNWNNLKGKKSCHTAVGRT
AGWVIPMGLIHNRTGTCNFDEYFSEGCAPGSPPNSRLCQLCQGSGGIPPEKCVASSHEKY
FGYTGALRCLVEKGDVAFIQHSTVEENTGGKNKADWAKNLQMDDFELLCTDGRRANVMDY
RECNLAEVPTHAVVVRPEKANKIRDLLERQEKRFGVNGSEKSKFMMFESQNKDLLFKDLT
KCLFKVREGTTYKEFLGDKFYTVISSLKTCNPSDILQMCSFLEGK

>sp|P01005|IOVO_CHICK Ovomucoid OS=Gallus gallus OX=9031 PE=1 SV=1

MAMAGVFVLFSFVLCGFLPDAAFGAEVDCSRFPNATDKEGKDVLVCNKDLRPICGTDGVT
YTNDCLLCAYSIEFGTNISKEHDGECKETVPMNCSSYANTTSEDGKVMVLCNRAFNPVCG
TDGVTYDNECLLCAHKVEQGASVDKRHDGGCRKELAAVSVDCSEYPKPDCTAEDRPLCGS
DNKTYGNKCNFCNAVVESNGTLTLSHFGKC

>tr|I0J171|I0J171_CHICK OvoglobulinG2 OS=Gallus gallus OX=9031 GN=G2 PE=2 SV=1

MGALLALLDPVQPTRAPDCGGILTPLGLSYLAEVSKPHAEVVLRQDLMAQRASDLFLGSM
EPSRNRITSVKVADLWLSVIPEAGLRLGIEVELRVAPLHAVPMPVRISIRADLHVDMGPD
GNLQLLTSACRPTVQAQSTREAESKSSRSILDKVVDVDKLCLDVSKLLLFPNEQLMSLTA
LFPVTPNCQLQYLPLAAPVFSKQGIALSLQTTFQVAGAVVPVPVSPVPFSMPELASTSTS
HLILALSEHFYTSLYFTLERAGAFNMTIPSMLTTATLAQKITQVGSLYHEDLPITLSAAL
RSSPRVVLEEGRAALKLFLTVHIGAGSPDFQSFLSVSADVTAGLQLSVSDTRMMISTAVI
EDAELSLAASNVGLVRAALLEELFLAPVCQQVPAWMDDVLREGVHLPHLSHFTYTDISVV
VHKDYVLVPCKLKLRSTMA

>sp|F1NBL0|MUC6_CHICK Mucin-6 (Fragment) OS=Gallus gallus OX=9031 GN=MUC6 PE=1 SV=1

CSTWGGGHFSTFDKYQYDFTGTCNYIFATVCDESSPDFNIQFRRGLDKKIARIIIELGPS
VIIVEKDSISVRSVGVIKLPYASNGIQIAPYGRSVRLVAKLMEMELVVMWNNEDYLMVLT
EKKYMGKTCGMCGNYDGYELNDFVSEGKLLDTYKFAALQKMDDPSEICLSEEISIPAIPH
KKYAVICSQLLNLVSPTCSVPKDGFVTRCQLDMQDCSEPGQKNCTCSTLSEYSRQCAMSH
QVVFNWRTENFCSVGKCSANQIYEECGSPCIKTCSNPEYSCSSHCTYGCFCPEGTVLDDI
SKNRTCVHLEQCPCTLNGETYAPGDTMKAACRTCKCTMGQWNCKELPCPGRCSLEGGSFV
TTFDSRSYRFHGVCTYILMKSSSLPHNGTLMAIYEKSGYSHSETSLSAIIYLSTKDKIVI
SQNELLTDDDELKRLPYKSGDITIFKQSSMFIQMHTEFGLELVVQTSPVFQAYVKVSAQF
QGRTLGLCGNYNGDTTDDFMTSMDITEGTASLFVDSWRAGNCLPAMERETDPCALSQLNK
ISAETHCSILTKKGTVFETCHAVVNPTPFYKRCVYQACNYEETFPYICSALGSYARTCSS
MGLILENWRNSMDNCTITCTGNQTFSYNTQACERTCLSLSNPTLECHPTDIPIEGCNCPK
GMYLNHKNECVRKSHCPCYLEDRKYILPDQSTMTGGITCYCVNGRLSCTGKLQNPAESCK
APKKYISCSDSLENKYGATCAPTCQMLATGIECIPTKCESGCVCADGLYENLDGRCVPPE
ECPCEYGGLSYGKGEQIQTECEICTCRKGKWKCVQKSRCSSTCNLYGEGHITTFDGQRFV
FDGNCEYILAMDGCNVNRPLSSFKIVTENVICGKSGVTCSRSISIYLGNLTIILRDETYS
ISGKNLQVKYNVKKNALHLMFDIIIPGKYNMTLIWNKHMNFFIKISRETQETICGLCGNY
NGNMKDDFETRSKYVASNELEFVNSWKENPLCGDVYFVVDPCSKNPYRKAWAEKTCSIIN
SQVFSACHNKVNRMPYYEACVRDSCGCDIGGDCECMCDAIAVYAMACLDKGICIDWRTPE
FCPVYCEYYNSHRKTGSGGAYSYGSSVNCTWHYRPCNCPNQYYKYVNIEGCYNCSHDEYF
DYEKEKCMPCAMQPTSVTLPTATQPTSPSTSSASTVLTETTNPPV

>sp|Q98UI9|MUC5B_CHICK Mucin-5B OS=Gallus gallus OX=9031 GN=MUC5B PE=1 SV=1

MEIKKERSFWIFCLIWSFCKGKEPVQIVQVSTVGRSECTTWGNFHFHTFDHVKFTFPGTC

TYVFASHCNDSYQDFNIKIRRSDKNSHLIYFTVTTDGVILEVKETGITVNGNQIPLPFSL
KSILIEDTCAYFQVTSKLGLTLKWNWADTLLLDLEETYKEKICGLCGNYDGNKKNDLILD
GYKMHPRQFGNFHKVEDPSEKCPDVRPDDHTGRHPTEDDNRCSKYKKMCKKLLSRFGNCP
KVVAFDDYVATCTEDMCNCVVNSSQSDLVSSCICSTLNQYSRDCVLSKGDPGEWRTKELC
YQECPSNMEYMECGNSCADTCADPERSKICKAPCTDGCFCPPGTILDDLGGKKCVPRDSC
PCMFQGKVYSSGGTYSTPCQNCTCKGGHWSCISLPCSGSCSIDGGFHIKTFDNKKFNFHG
NCHYVLAKNTDDTFVVIGEIIQCGTSKTMTCLKNVLVTLGRTTIKICSCGSIYMNNFIVK
LPVSKDGITIFRPSTFFIKILSSAGVQIRVQMKPVMQLSITVDHSYQNRTSGLCGNFNNI
QTDDFRTATGAVEDSAAAFGNSWKTRASCFDVEDSFEDPCSNSVDKEKFAQHWCALLSNT
SSTFAACHSVVDPSVYIKRCMYDTCNAEKSEVALCSVLSTYSRDCAAAGMTLKGWRQGIC
DPSEECPETMVYNYSVKYCNQSCRSLDEPDPLCKVQIAPMEGCGCPEGTYLNDEEECVTP
DDCPCYYKGKIVQPGNSFQEDKLLCKCIQGRLDCIGETVLVKDCPAPMYYFNCSSAGPGA
IGSECQKSCKTQDMHCYVTECVSGCMCPDGLVLDGSGGCIPKDQCPCVHGGHFYKPGETI
RVDCNTCTCNKRQWNCTDNPCKGTCTVYGNGHYMSFDGEKFDFLGDCDYILAQDFCPNNM
DAGTFRIVIQNNACGKSLSICSLKITLIFESSEIRLLEGRIQEIATDPGAEKNYKVDLRG
GYIVIETTQGMSFMWDQKTTVVVHVTPSFQGKVCGLCGDFDGRSRNDFTTRGQSVEMSIQ
EFGNSWKITSTCSNINMTDLCADQPFKSALGQKHCSIIKSSVFEACHSKVNPIPYYESCV
SDFCGCDSVGDCECFCTSVAAYARSCSTAGVCINWRTPAICPVFCDYYNPPDKHEWFYKP
CGAPCLKTCRNPQGKCGNILYSLEGCYPECSPDKPYFDEERRECVSLPDCTSCNPEEKLC
TEDSKDCLCCYNGKTYPLNETIYSQTEGTKCGNAFCGPNGMIIETFIPCSTLSVPAQEQL
MQPVTSAPLLSTEATPCFCTDNGQLIQMGENVSLPMNISGHCAYSICNASCQIELIWAEC
KVVQTEALETCEPNSEACPPTAAPNATSLVPATALAPMSDCLGLIPPRKFNESWDFGNCQ
IATCLGEENNIKLSSITCPPQQLKLCVNGFPFMKHHDETGCCEVFECQCICSGWGNEHYV
TFDGTYYHFKENCTYVLVELIQPSSEKFWIHIDNYYCGAADGAICSMSLLIFHSNSLVIL
TQAKEHGKGTNLVLFNDKKVVPDISKNGIRITSSGLYIIVEIPELEVYVSYSRLAFYIKL
PFGKYYNNTMGLCGTCTNQKSDDARKRNGEVTDSFKEMALDWKAPVSTNRYCNPGISEPV
KIENYQHCEPSELCKIIWNLTECHRVVPPQPYYEACVASRCSQQHPSTECQSMQTYAALC
GLHGICVDWRGQTNGQCEATCARDQVYKPCGEAKRNTCFSREVIVDTLLSRNNTPVFVEG
CYCPDGNILLNEHDGICVSVCGCTAQDGSVKKPREAWEHDCQYCTCDEETLNISCFPRPC
AKSPPINCTKEGFVRKIKPRLDDPCCTETVCECDIKTCIINKTACDLGFQPVVAISEDGC
CPIFSCIPKGVCVSEGVEFKPGAVVPKSSCEDCVCTDEQDAVTGTNRIQCVPVKCQTTCQ
QGFRYVEKEGQCCSQCQQVACVANFPFGSVTIEVGKSYKAPYDNCTQYTCTESGGQFSLT
STVKVCLPFEESNCVPGTVDVTSDGCCKTCIDLPHKCKRSMKEQYIVHKHCKSAAPVPVP
FCEGTCSTYSVYSFENNEMEHKCICCHEKKSHVEKVELVCSEHKTLKFSYVHVDECGCVE
TKCPMRRT

>sp|P02701|AVID_CHICK Avidin OS=Gallus gallus OX=9031 GN=AVD PE=1 SV=3

MVHATSPLLLLLLLSLALVAPGLSARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITA
VTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTM
WLLRSSVNDIGDDWKATRVGINIFTRLRTQKE

Appendix A, File 2

```
#!/usr/bin/python
from __future__ import division
import numpy as np
import pandas as pd
import scipy.stats
import random

def read_bulk_df():
    """ Reads bulk stat files and returns a DF
    """
    bulk_df = pd.DataFrame()
    for f in modern_bulk_files:
        df = pd.read_csv((data_folder + f))
        df["Group"] = "Modern"
        bulk_df = bulk_df.append(df, ignore_index = False)
    for f in ancient_bulk_files:
        df = pd.read_csv((data_folder + f))
        df["Group"] = "Ancient"
        bulk_df = bulk_df.append(df, ignore_index = False)
    return bulk_df


def read_ss_df():
    """ Reads SS stat files and returns a DF
    """
    ss_df = pd.DataFrame()
    for f in modern_SS_files:
        df = pd.read_csv((data_folder + f))
        df["Group"] = "Modern"
            ss_df = ss_df.append(df, ignore_index = False, sort =
True)
    for f in ancient_SS_files:
```

```
        df = pd.read_csv((data_folder + f))
        df["Group"] = "Ancient"
            ss_df = ss_df.append(df, ignore_index = False, sort =
True)
    # Remove -1 values
    ss_df = ss_df[ss_df["Half-time"] != -1]
    return ss_df


def read_sk1_bulk():
    """ Reads SK1 bulk stat file and returns a DF
    """
    sk1_bulk_df = pd.read_csv((data_folder + bulk_stats_file))
     sk1_bulk_df = sk1_bulk_df[(sk1_bulk_df["Sample"] == "sample")]
# No blank
    return sk1_bulk_df


def read_sk1_ss():
    """ Reads SK1 SS stat file and returns a DF
    """
    sk1_ss_df = pd.read_csv((data_folder + ss_stats_file))
    sk1_ss_df = sk1_ss_df[(sk1_ss_df["Sample"] == "sample")]
    sk1_ss_df = sk1_ss_df[sk1_ss_df["Half-time"] != -1]
    return sk1_ss_df


def mann_whitney(x, y, desc):
    """ Performs Mann Whitney U
    and prints results
    """
    stat, p = scipy.stats.mannwhitneyu(x, y)
    print desc
    print "Mann Whitney U"
    print "P: \t\t\t", p
```

```
    print "Test statistic: \t", stat
    if p > ALPHA:
     print "p > %f, fail to reject H0" % (ALPHA)
    else:
     print "p < %f, reject H0" % (ALPHA)
    print ""


def bulk(bulk_df, col_name, gr1, gr2, aa):
    """ Gets bulk deamidation significance
    """
    this_df = bulk_df[(bulk_df['AA']==aa)]
    modern_df = this_df[(this_df[col_name] == gr1)]
    ancient_df = this_df[(this_df[col_name] == gr2)]
    desc = "Bulk deamidation, %s vs %s (%s)" % (gr1, gr2, aa)
                              mann_whitney(modern_df["RelNonDeam"],
ancient_df["RelNonDeam"], desc)


def pair_ss_tolerance(ss_df, col_name, gr1, gr2, tolerance=.05):
    """ Produces lists of pairs of relative deamidation levels
    at the same half-time. If tolerance > 0, use this to compare
     deamidation levels at half-times that are within x (tolerance
= x)
    percent of each other.
    """

    gr1_df = ss_df[(ss_df[col_name] == gr1)]
    gr2_df = ss_df[(ss_df[col_name] == gr2)]

    half_times = ss_df["Half-time"].unique()

    pairs_list = []

    for h1 in half_times:
```

```
        # Add tolerance
        low = h1 - (h1*tolerance)
        high = h1 + (h1*tolerance)

        # New dfs with only relevant half-times
        gr1 = gr1_df[(gr1_df["Half-time"] >= low)]
        gr1 = gr1[(gr1["Half-time"] <= high)]
        gr2 = gr2_df[(gr2_df["Half-time"] >= low)]
        gr2 = gr2[(gr2["Half-time"] <= high)]

        # If both dfs are populated
        if len(gr1.index) > 0 and len(gr2.index) > 0:

            # Get mean for later
            gr1_avg = gr1.mean(axis = 0)["RelNonDeam"]
            gr2_avg = gr2.mean(axis = 0)["RelNonDeam"]

            while len(gr1.index) > 0 and len(gr2.index) > 0:

                    # If the dfs have more than one row, generated
random row
                # to pair, and then drop

                random_gr1_index = 0
                random_gr2_index = 0

                if len(gr1.index) > 1:
                                                random_gr1_index  =
random.randrange(len(gr1.index)-1)
                g1_row = gr1.iloc[random_gr1_index]
                gr1 = gr1.drop(gr1.index[random_gr1_index])

                if len(gr2.index) > 1:
                                                random_gr2_index  =
random.randrange(len(gr2.index)-1)
```

```
                g2_row = gr2.iloc[random_gr2_index]
                gr2 = gr2.drop(gr2.index[random_gr2_index])

                # Add to pair
                            pairs_list.append([g1_row["RelNonDeam"],
g2_row["RelNonDeam"]])

            # Use average vals if one df is bigger than the other
            if len(gr1.index) > 0:
                for i, row in gr1.iterrows():
                            pairs_list.append([row["RelNonDeam"],
gr2_avg])
            if len(gr2.index) > 0:
                for i, row in gr2.iterrows():
                                        pairs_list.append([gr1_avg,
row["RelNonDeam"]])

    return pairs_list


def wilcoxon(pairs_list, desc):
    """ Performs Wilcoxon signed rank
    and prints results
    """
    pairs_list = np.array(pairs_list)
    m = pairs_list[:,0]
    a = pairs_list[:,1]
    stat, p = scipy.stats.wilcoxon(m, a)
    print desc
    print "Wilcoxon signed rank test"
    print "N: ", len(pairs_list)
    print "P: \t\t\t", p
    print "Test statistic: \t", stat
    if p > ALPHA:
     print "p > %f, fail to reject H0" % (ALPHA)
```

```
    else:
     print "p < %f, reject H0" % (ALPHA)
    print ""



data_folder = "/home/abby/Documents/1.
DeamiDATE/REVISIONS/STATS_datasets/"

ALPHA = .05

### Model

modern_bulk_files = ["Model Dataset - Aardvark - BULK.csv",
"Model Dataset - Anteater - BULK.csv",
    "Model Dataset - Hippo - BULK.csv", "Model Dataset - Tapir -
BULK.csv"]
modern_SS_files = ["Model Dataset - Aardvark - SS.csv", "Model
Dataset - Anteater - SS.csv",
    "Model Dataset - Hippo - SS.csv", "Model Dataset - Tapir -
SS.csv"]
ancient_bulk_files = ["Model Dataset - Macrauchenia - BULK.csv",
"Model Dataset - Mylodon - BULK.csv",
    "Model Dataset - Toxodon - BULK.csv"]
ancient_SS_files = ["Model Dataset - Macrauchenia - SS.csv",
"Model Dataset - Mylodon - SS.csv",
    "Model Dataset - Toxodon - SS.csv"]



#Bulk
bulk_df = read_bulk_df()
bulk(bulk_df, "Group", "Modern", "Ancient", "N")
bulk(bulk_df, "Group", "Modern", "Ancient", "Q")
```

```
# SS
ss_df = read_ss_df()
pairs_list = pair_ss_tolerance(ss_df, "Group", "Modern",
"Ancient", tolerance=0)
wilcoxon(pairs_list, "Site specific deamidation, modern vs
ancient")


### SK 1

bulk_stats_file = "SK1 - Bulk.csv"
ss_stats_file = "SK1 - SS.csv"


#Bulk
sk1_bulk_df = read_sk1_bulk()
bulk(sk1_bulk_df, "Classification", "Human", "Egg", "N")
bulk(sk1_bulk_df, "Classification", "Human", "Egg", "Q")

#SS
sk1_ss_df = read_sk1_ss()
pairs_list = pair_ss_tolerance(sk1_ss_df, "Classification",
"Human", "Egg", tolerance=.05)
wilcoxon(pairs_list, "Site specific deamidation, human vs egg")
```

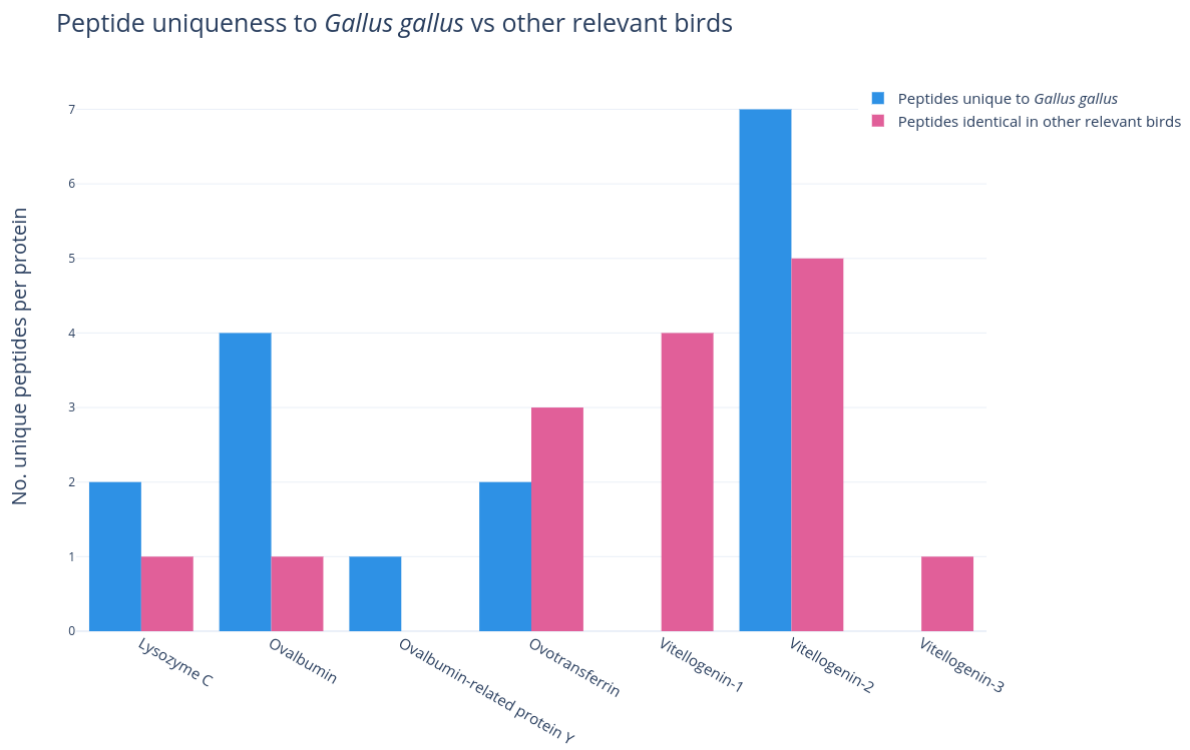
Appendix A, Table 6

In order to ascertain whether or not the chicken peptides were conserved across other bird species, and could thus originate from wild bird species available in Neolithic Britain, each peptide was taxonomically identified through NCBI BLASTp (Altschul *et al.*, 1990). In total, 16 peptides identified as being unique to chicken, while 15 matched identically to other relevant birds too - most commonly *Anas platyrhynchos*.

		Gallus gallus			
P00698	Lysozyme C	2	166.56	1	214.00
P01012	Ovalbumin	4	136.76	1	91.47
P01014	Ovalbumin-related protein Y	1	105.52	0	N/A
P02789	Ovotransferrin	2	127.55	3	122.11
P87498	Vitellogenin-1	0	N/A	4	89.32
P02845	Vitellogenin-2	7	91.50	5	103.29
Q91025	Vitellogenin-3	0	N/A	1	61.997

Appendix A, Table 6 Peptide count in proteins unique to Gallus gallus versus those identical in other relevant birds and their scores.

Appendix A, Figure 1

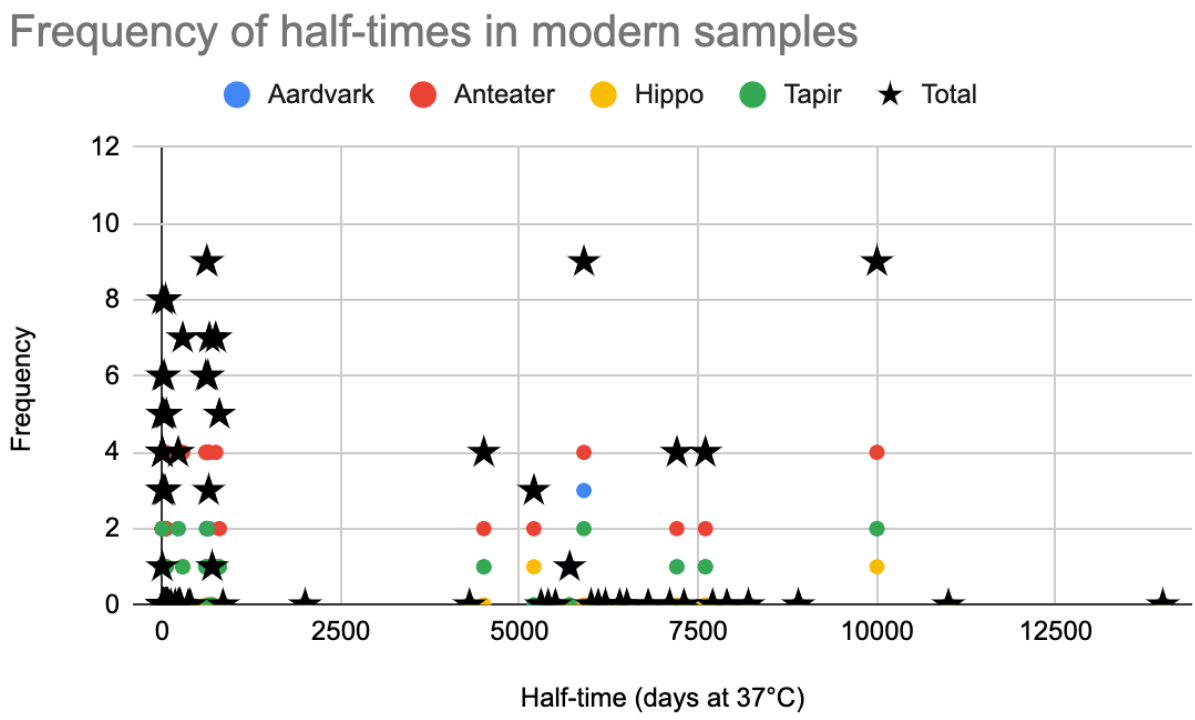


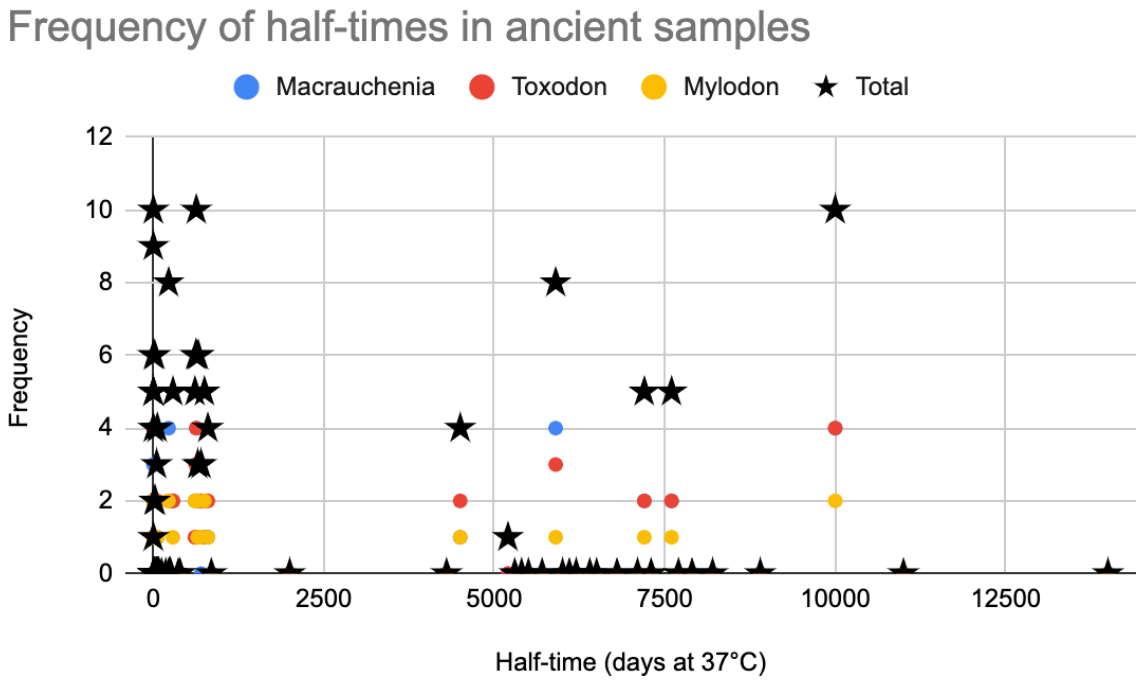
Appendix A, Figure 1: Peptide count in egg endogenous proteins. Peptide count unique to chicken is shown in blue, while peptides that are identical in other relevant birds are shown in blue.

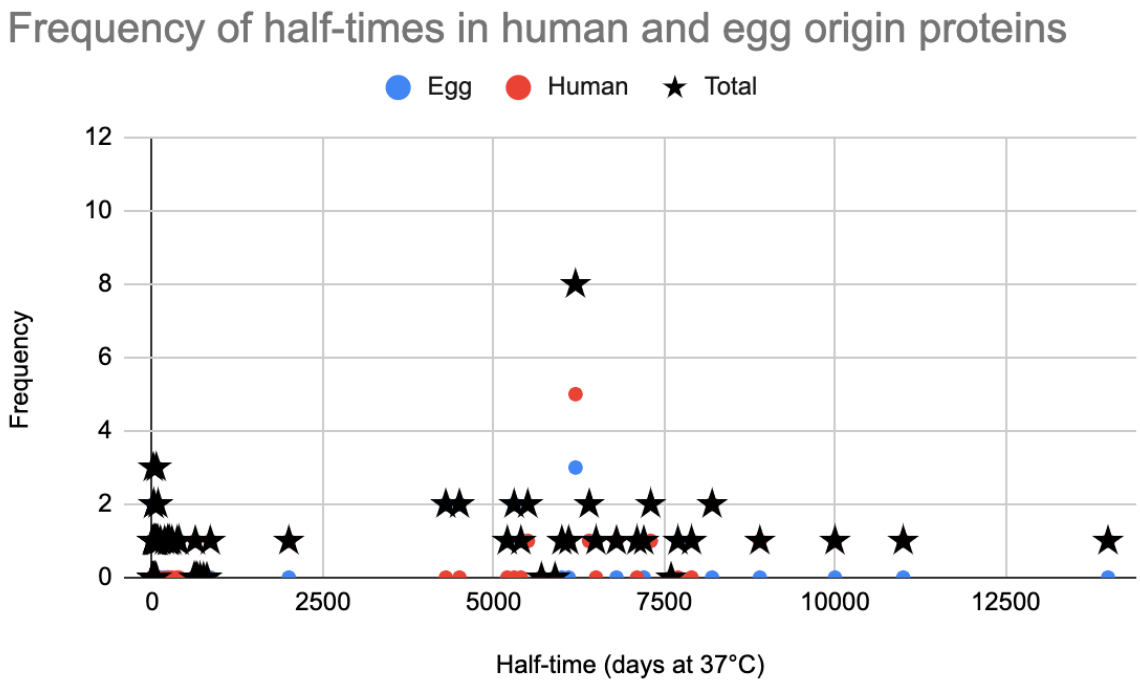
Appendix A, Figure 2

The frequency of different half-times in the datasets was analysed to investigate whether the deamidation, or lack thereof, at different half-times is only dependent on the original distribution of half-times. The distribution of half-times in the model dataset groups is very similar - this is most likely due to the fact that they both represent only collagen type 1 peptides.

The egg and human origin proteins from SK1 show a wider variation in reported half-times, with the human endogenous proteins especially being represented at almost every half-time







Appendix A, Figure 2: The frequency of the Robinson and Robinson (2004) half-times found in the SK1 protein groups and the modern and ancient model dataset.

Appendix B, Table 1

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
3DT21 (CS6)	Hendy et al. 2018a	UK - Driffield Terrace, York, York	Roman	44-410 CE	1723	No
3DT26 (CS5)	Hendy et al. 2018a	UK - Driffield Terrace, York, York	Roman	44-410 CE	1723	No
6DT21 (CS9)	Hendy et al. 2018a	UK - Driffield Terrace, York, York	Roman	44-410 CE	1723	No
6DT3 (CS7)	Hendy et al. 2018a	UK - Driffield Terrace, York, York	Roman	44-410 CE	1723	No
6DT7 (CS8)	Hendy et al. 2018a	UK - Driffield Terrace, York, York	Roman	44-410 CE	1723	No
CS11	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS12	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS18	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS19	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS21	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS31	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS40	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS46	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS6	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
CS7	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1485-1872 CE	271.5	No
FAO1	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO10	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO11	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO12	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO13	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO14	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO15	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO16	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO17	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO18	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO2	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO3	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO4	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO5	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO6	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO7	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO8	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FAO9	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1485-1872 CE	271.5	No
FW192	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW217	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW268	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW303	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW331	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW435	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
FW450	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW53	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW68	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW88	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW9	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
FW98	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, UK - Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1485-1872 CE	271.5	No
ML1032	Hendy et al. 2018a	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML1489	Hendy et al. 2018a	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML1823	Hendy et al. 2018a	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML3890	Hendy et al. 2018a	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
NBS262	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NBS325	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NBS410	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM093	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM099	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM18	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
OX01	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX03	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX04	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX05	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX06	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX09	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX10	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
OX12	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	43-410 CE	1723	No
RLH 103	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 131	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 135	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 208	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 340	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 349	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 365	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 367	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 386	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 397	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 421	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
RLH 572	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1485-1872 CE	271.5	No
SP2182	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2295	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2300	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
SP2301	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2369	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2468	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2477	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2484	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2647	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2748	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
SP2860	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1485-1872 CE	271.5	No
TKAC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	410-1484 CE	1003	No
TKDC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	410-1484 CE	1003	No
TKEC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	410-1484 CE	1003	No
TKFC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	410-1484 CE	1003	No
UT103	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	2018 CE	0	No - but reliable due to modernity
UT111	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	2018 CE	0	No - but reliable due to modernity
UT116	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	2018 CE	0	No - but reliable due to modernity
UT73	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	2018 CE	0	No - but reliable due to modernity
WG1082	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1252	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1483	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1561	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1566	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1585	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
Z100	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z81	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z82	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z83	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z84	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z85	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z86	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z87	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z88	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z89	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z90	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
Z94	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	2018 CE	0	No - but reliable due to modernity
CW10 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW10 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW10 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW11 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW18 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
CW18 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW20 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW20 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW20 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW21 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW22 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW22 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW22 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW23 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW24 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW24 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW27 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
CW8 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7950-7550 BCE	7750	Yes - indirectly from context
ARS007	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3235-3076 BCE	3155.5	Yes
ARS008	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3235-3076 BCE	3155.5	Yes
ARS009	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	average of other	3176.0625	No
ARS010	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3028-2879 BCE	2953.5	Yes
ARS013	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3355-3230 BCE	3292.5	Yes
ARS014	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3335-3168 BCE	3251.5	Yes
ARS016	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3455-3350 BCE	3402.5	Yes
ARS017	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3343-3183 BCE	3263	Yes
ARS018	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	3002-2867 BCE	2934.5	Yes
VVH02	Mays et al. 2018	UK - West Amesbury Farm	Middle Bronze A(	3346-3431 YBP	3388.5	Yes
VVH03	Mays et al. 2018	UK - West Amesbury Farm	Middle Bronze A(	3312-3382 YBP	3347	Yes
VVH04	Mays et al. 2018	UK - West Amesbury Farm	Middle Neolithic	4869-4955 YBP	4912	Yes
3DT21 (CS6)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	44-410 CE	1723	No
3DT26 (CS5)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	44-410 CE	1723	No
6DT21 (CS9)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	44-410 CE	1723	No
6DT3 (CS7)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	44-410 CE	1723	No
6DT7 (CS8)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	44-410 CE	1723	No
FW435	Warinner et al. 2014	UK - Fewston	Victorian	1835-1895 CE	85	No
FW53	Warinner et al. 2014	UK - Fewston	Victorian	1835-1895 CE	85	No
JH1	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH1	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH10	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH10	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH11	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH11	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH12	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
JH12	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH13	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH13	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH14	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH14	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH15	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH15	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH16	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH16	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH17	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH17	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH18	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH18	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH19	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH19	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH2	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH2	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH20	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH20	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH3	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH3	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH4	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH4	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH5	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH5	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH6	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH6	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH7	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH7	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH8	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH8	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1840-1872 CE	94	No
ML1032	Warinner et al. 2014	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML1489	Warinner et al. 2014	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML1823	Warinner et al. 2014	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
ML3890	Warinner et al. 2014	UK - Melton	Iron Age	800 BCE to 80 C	2310	No
NBS262	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
NBS325	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NBS410	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM093	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM099	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
NEM18	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	ca. 500-910 CE	1245	No
OX01	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX03	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX05	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX09	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX10	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX12	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX4	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
OX6	Warinner et al. 2014	UK - Leicester	Roman	44-410 CE	1723	No
RISE 432	Warinner et al. 2014	Denmark - Gjerrild	Neolithic	ca. 3000-1500 B	4200	No
RISE 460	Warinner et al. 2014	Denmark - Øster Harup	Neolithic	ca. 2800-2400 B	4550	No
RISE363	Warinner et al. 2014	Hungary - Szöreg-C (Sziv Utca)	Bronze Age	ca. 3000-1500 B	4200	No
RISE368	Warinner et al. 2014	Hungary - Szöreg-C (Sziv Utca)	Bronze Age	ca. 3000-1500 B	4200	No
RISE387	Warinner et al. 2014	Russia - Bulanovo	Bronze Age	ca. 3000-1500 B	4200	No
SCR5070	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR832	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
WG1082	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1252	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1483	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1561	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1566	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
WG1585	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	ca. 1000-1550 C	675	No
Z44	Warinner et al. 2014	UK - Fewston	Victorian	1835-1895 CE	85	No
Z45	Warinner et al. 2014	UK - Fewston	Victorian	1835-1895 CE	85	No
Z5	Warinner et al. 2014	Switzerland - Zurich	Modern	2011 CE	3	No - but reliable due to modernity
Z6	Warinner et al. 2014	Switzerland - Zurich	Modern	2011 CE	3	No - but reliable due to modernity
Z7	Warinner et al. 2014	Switzerland - Zurich	Modern	2011 CE	3	No - but reliable due to modernity
Z8	Warinner et al. 2014	Switzerland - Zurich	Modern	2011 CE	3	No - but reliable due to modernity
RISE307	Warinner et al. 2014	Russia - West Caucasus, Marchenkova Gora	Bronze Age	ca. 3000-2000 B	4450	No
RISE307	Warinner et al. 2014	Russia - West Caucasus, Marchenkova Gora	Bronze Age	ca. 3000-2000 B	4450	No
RISE417	Warinner et al. 2014	Armenia - Hatsarat	Bronze Age	ca. 2000 BCE	3950	No
RISE420	Warinner et al. 2014	Armenia - Noraduz	Iron Age	ca. 700 BCE	2650	No
RISE415	Warinner et al. 2014	Armenia - Nerkin Getashen	Bronze Age	ca. 1200-1300 B	3200	No
RISE414	Warinner et al. 2014	Armenia - Noraduz	Bronze Age	ca. 1200-1300 B	3200	No
RISE472	Warinner et al. 2014	Germany - Nersingen	Bronze Age	ca. 3000-1500 B	4200	No

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
RISE473	Warinner et al. 2014	Germany - Regensburg-Dechbetten	Bronze Age	ca. 3000-1500 BC	4200	No
Y47-B27	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
Y48-B40	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
Y49-B85	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
B78	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
G12	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
B17	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
B78	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
B61	Warinner et al. 2014	Germany - Dalheim	Medieval	ca. 950-1200 CE	875	No
Z35	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	ca. 1290-1430 CI	590	No
Z36	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	ca. 1290-1430 CI	590	No
Z37	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	ca. 1290-1430 CI	590	No
Z38	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	ca. 1290-1430 CI	590	No
Z39	Warinner et al. 2014	Greenland - Ø29a Brattahlið	Medieval	ca. 890-1230 CE	890	No
Z40	Warinner et al. 2014	Greenland - Ø29a Brattahlið	Medieval	ca. 890-1230 CE	890	No
RISE466	Warinner et al. 2014	Italy - Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
RISE467	Warinner et al. 2014	Italy - Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN19	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN207	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN271	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN361	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN424	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN458	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
ODN98	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	ca. 2700-1350 BC	3625	No
SCR277	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR250	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR264	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR323	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR5028	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
SCR5042	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	50-200 CE	1825	No
BL13217	Charlton et al. 2019	UK - Banbury Lane	Neolithic	3360–3090 BCE	5175	Yes - indirectly from context
BL30912	Charlton et al. 2019	UK - Banbury Lane	Neolithic	3360–3090 BCE	5175	Yes - indirectly from context
HH3188	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	3650–3370 BCE	5460	Yes - indirectly from context
HH3	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	3680–3310 BCE	5445	Yes - indirectly from context
HH610	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	3680–3310 BCE	5445	Yes - indirectly from context
HN5037-1	Charlton et al. 2019	UK - Hazleton North	Neolithic	3800–3620 BCE	5660	Yes - indirectly from context
HN7656	Charlton et al. 2019	UK - Hazleton North	Neolithic	3800–3620 BCE	5660	Yes - indirectly from context
HN11456	Charlton et al. 2019	UK - Hazleton North	Neolithic	3800–3620 BCE	5660	Yes - indirectly from context
HN4786	Charlton et al. 2019	UK - Hazleton North	Neolithic	3800–3620 BCE	5660	Yes - indirectly from context

ID	Paper	Site	Age	Date Range	Mid of range (YE	RC date?
HN7387	Charlton et al. 2019	UK - Hazleton North	Neolithic	3800–3620 BCE	5660	Yes - indirectly from context

Appendix B, Table 1: Radiocarbon and inferred dates of samples from all archaeological datasets

Appendix B, Table 2

ID	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
Milk Powder	295	72	90.83	95.56	62.11	88.27

Appendix B, Table 2: Deamidation and peptide count in milk powder

Appendix B, Table 3

ID	Half-time	Remaining AA	Size	Protein
Milk Powder	-1	0.9691052666	104.2223646	BLG
Milk Powder	1.14	0.3606003063	104.5051933	BLG
Milk Powder	10	0	100	BLG
Milk Powder	15.7	0.2844494928	109.9325373	BLG
Milk Powder	48.3	1	100	BLG
Milk Powder	50.2	1	122.7099211	BLG
Milk Powder	75.7	0.9567864389	122.7099211	BLG
Milk Powder	104	1	1000	BLG
Milk Powder	300	0	100	BLG
Milk Powder	800	1	109.020425	BLG
Milk Powder	2100	0.8108389566	104.6366516	BLG
Milk Powder	4000	1	171.1546966	BLG
Milk Powder	5200	0.9000192751	104.3315886	BLG
Milk Powder	5200	1	107.1529474	BLG
Milk Powder	5700	0.9982989777	104.5051933	BLG
Milk Powder	6100	1	1000	BLG
Milk Powder	6100	1	100.2780195	BLG
Milk Powder	6100	1	109.3702632	BLG
Milk Powder	6100	1	106.8427546	BLG
Milk Powder	-1	0	100	Casein
Milk Powder	-1	0	100	Casein
Milk Powder	-1	1	133.0717462	Casein
Milk Powder	-1	0.999394259	128.8832862	Casein
Milk Powder	-1	0.9965107014	132.1820048	Casein
Milk Powder	-1	1	128.5574808	Casein
Milk Powder	-1	1	128.5574808	Casein
Milk Powder	-1	1	127.3831087	Casein
Milk Powder	-1	0.9998621035	128.5574808	Casein
Milk Powder	-1	0.7350080699	128.8832862	Casein
Milk Powder	-1	1	129.8320684	Casein
Milk Powder	-1	0.7994730488	167.2336046	Casein
Milk Powder	-1	0.9393384008	109.6784119	Casein
Milk Powder	-1	1	167.2336046	Casein
Milk Powder	-1	1	101.1732887	Casein
Milk Powder	-1	0.983613673	110.9198439	Casein
Milk Powder	-1	1	101.1732887	Casein
Milk Powder	-1	1	100.5872975	Casein
Milk Powder	-1	0.9984034957	111.1428425	Casein
Milk Powder	-1	1	108.8576632	Casein
Milk Powder	-1	1	173.5291251	Casein
Milk Powder	-1	1	162.3269128	Casein
Milk Powder	-1	0.9963390309	167.4544558	Casein
Milk Powder	15.7	0.9427376321	103.1472924	Casein
Milk Powder	15.7	0.9439537807	100.5269487	Casein
Milk Powder	18.9	0.9278185369	106.5625141	Casein

Milk Powder	18.9	1	108.8576632	Casein
Milk Powder	22.5	0.9733653145	132.7652459	Casein
Milk Powder	36.8	1	100.936776	Casein
Milk Powder	45.7	0.9238611784	121.1270652	Casein
Milk Powder	46.3	1	100.4237236	Casein
Milk Powder	56.7	0.2345581476	100.183335	Casein
Milk Powder	59.6	1	100.139765	Casein
Milk Powder	68.3	1	100.3524708	Casein
Milk Powder	71.1	1	101.6434975	Casein
Milk Powder	73.9	0.9313473188	136.3942803	Casein
Milk Powder	75.7	0.9332352967	179.5894436	Casein
Milk Powder	75.7	0.9418921862	110.3410569	Casein
Milk Powder	80.9	0.8520160699	101.2566905	Casein
Milk Powder	92.6	0.9445353347	133.9672061	Casein
Milk Powder	116	0.9216917453	133.1388014	Casein
Milk Powder	116	0.9997060199	110.6989911	Casein
Milk Powder	116	1	100.944151	Casein
Milk Powder	226	1	129.6858185	Casein
Milk Powder	850	1	109.7809322	Casein
Milk Powder	850	0.9778508195	111.4301226	Casein
Milk Powder	4000	1	100.5073443	Casein
Milk Powder	4000	1	102.7382393	Casein
Milk Powder	4100	0.8450146926	101.9333427	Casein
Milk Powder	4200	1	121.0844249	Casein
Milk Powder	4300	1	102.9133313	Casein
Milk Powder	4300	1	123.9842788	Casein
Milk Powder	4600	1	100.1265149	Casein
Milk Powder	4600	1	102.2580639	Casein
Milk Powder	5100	1	126.1005238	Casein
Milk Powder	5500	1	107.7216157	Casein
Milk Powder	5800	0	550	Casein
Milk Powder	5800	1	106.4523601	Casein
Milk Powder	5900	1	200.5892357	Casein
Milk Powder	6000	1	105.5949543	Casein
Milk Powder	6100	1	100.1854306	Casein
Milk Powder	6200	1	121.9652449	Casein
Milk Powder	6200	1	739.1277493	Casein
Milk Powder	6200	1	172.3787111	Casein
Milk Powder	6400	1	106.8296028	Casein
Milk Powder	6600	1	109.6784119	Casein
Milk Powder	6800	0.8892301769	110.5185922	Casein
Milk Powder	7100	1	100.6810439	Casein
Milk Powder	7200	1	100.5119952	Casein
Milk Powder	7400	1	100.5073443	Casein
Milk Powder	7500	1	155.6266298	Casein
Milk Powder	7800	1	100.7057242	Casein

Milk Powder	7800	0.9976667221	157.9145895	Casein
Milk Powder	8100	1	100.936776	Casein
Milk Powder	8200	0.9931581514	120.0394928	Casein
Milk Powder	8300	1	307.6258303	Casein
Milk Powder	8300	0.9989741101	178.2993698	Casein
Milk Powder	8400	1	100.3502067	Casein
Milk Powder	9300	0.930644892	101.5849107	Casein
Milk Powder	10000	1	128.5574808	Casein
Milk Powder	10000	0.9899527175	140.2524925	Casein
Milk Powder	10000	1	112.6709579	Casein
Milk Powder	10000	1	106.8385762	Casein
Milk Powder	11000	1	103.9083362	Casein
Milk Powder	12000	0.6949068523	144.7247746	Casein
Milk Powder	12000	1	110.266809	Casein
Milk Powder	16000	0.9978810503	119.9772883	Casein

Appendix B, Table 3: Site-specific deamidation of caseins and BLG in milk powder.

Appendix B, Table 4

ID	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
Blank	5	1	No N in Caseins	100.00	Not enough unique peptides	Not enough unique peptides
Dental Calculus A (8786)	2	0	20.00	57.14	No BLG	No BLG
Dental Calculus B (8836)	2	1	33.33	100.00	Not enough unique peptides	Not enough unique peptides
Dental Calculus C (8756)	6	0	15.55	100.00	No BLG	No BLG
Egg Shell (ES001)	2	0	No N in Caseins	100.00	No BLG	No BLG
Limescale (11287)	8	1	100.00	100.00	Not enough unique peptides	Not enough unique peptides

Appendix B, Table 4: Deamidation and peptide count in contamination experiment

Appendix B, Table 5

ID	Half-time	Remaining AA	Size	Protein
Blank	7400.00	0.00	179.12	BLG
Blank	75.70	1.00	100.00	Casein
Blank	116.00	1.00	100.00	Casein
Blank	850.00	1.00	100.00	Casein
Blank	4300.00	1.00	177.24	Casein
Blank	6200.00	1.00	507.74	Casein
Blank	6800.00	1.00	100.00	Casein
Blank	8300.00	1.00	507.74	Casein
Blank	10000.00	1.00	1000.00	Casein
Blank	12000.00	1.00	155.84	Casein
Dental Calculus A (8786)	4300.00	1.00	100.00	Casein
Dental Calculus A (8786)	6400.00	1.00	1000.00	Casein
Dental Calculus B (8836)	10000.00	0.00	1000.00	BLG
Dental Calculus B (8836)	75.70	0.16	198.46	Casein
Dental Calculus B (8836)	116.00	0.16	198.46	Casein
Dental Calculus B (8836)	4300.00	1.00	208.65	Casein
Dental Calculus C (8756)	73.90	1.00	446.35	Casein
Dental Calculus C (8756)	75.70	0.00	100.00	Casein
Dental Calculus C (8756)	75.70	0.00	100.00	Casein
Dental Calculus C (8756)	116.00	0.00	100.00	Casein
Dental Calculus C (8756)	116.00	0.00	100.00	Casein
Dental Calculus C (8756)	4300.00	1.00	1000.00	Casein
Dental Calculus C (8756)	12000.00	1.00	304.20	Casein
Egg Shell (ES001)	10000.00	1.00	1000.00	Casein
Egg Shell (ES001)	12000.00	1.00	100.00	Casein
Limescale (11287)	50.20	1.00	1000.00	BLG

Limescale (11287)	75.70	0.00	1000.00	BLG
Limescale (11287)	-1.00	1.00	151.96	Casein
Limescale (11287)	75.70	0.00	100.00	Casein
Limescale (11287)	75.70	0.00	673.24	Casein
Limescale (11287)	116.00	0.00	100.00	Casein
Limescale (11287)	116.00	0.00	673.24	Casein
Limescale (11287)	850.00	1.00	151.96	Casein
Limescale (11287)	4300.00	1.00	100.00	Casein
Limescale (11287)	5400.00	0.00	208.29	Casein
Limescale (11287)	6400.00	0.00	144.51	Casein
Limescale (11287)	6600.00	0.00	151.96	Casein
Limescale (11287)	8900.00	1.00	144.51	Casein
Limescale (11287)	10000.00	1.00	388.03	Casein

Appendix B, Table 5: Site-specific deamidation in contamination experiment

Appendix B, Table 6

ID	Paper	Site	Age	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
Z40	Warinner et al. 2014	Greenland - Ø29a Brattahlið	Medieval	0	7	No Caseins	No Caseins	66.67	50.00
OX12	Warinner et al. 2014	UK - Leicester	Roman	0	7	No Caseins	No Caseins	83.88	71.15
RISE368	Warinner et al. 2014	Hungary - Szöreg-C (Sziv Utca)	Bronze Age	7	6	62.37	91.62	80.49	81.72
OX4	Warinner et al. 2014	UK - Leicester	Roman	0	6	No Caseins	No Caseins	100.00	100.00
Z39	Warinner et al. 2014	Greenland - Ø29a Brattahlið	Medieval	1	5	Not enough unique peptides	Not enough unique peptides	50.00	No Q in BLG
NEM18	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	1	4	Not enough unique peptides	Not enough unique peptides	60.52	100.00
Z5	Warinner et al. 2014	Switzerland - Zurich	Modern	3	3	0.00	100.00	100.00	100.00
ML1032	Warinner et al. 2014	UK - Melton	Iron Age	2	3	No N in Caseins	50.00	No N in BLG	No Q in BLG
RISE387	Warinner et al. 2014	Russia - Bulanovo	Bronze Age	1	3	Not enough unique peptides	Not enough unique peptides	100.00	53.07
WG1082	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	1	3	Not enough unique peptides	Not enough unique peptides	No N in BLG	No Q in BLG
ML1823	Warinner et al. 2014	UK - Melton	Iron Age	0	3	No Caseins	No Caseins	100.00	100.00
WG1561	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	3	No Caseins	No Caseins	50.00	100.00
RISE363	Warinner et al. 2014	Hungary - Szöreg-C (Sziv Utca)	Bronze Age	4	2	No N in Caseins	No Q in Caseins	100.00	100.00
B61	Warinner et al. 2014	Germany - Dalheim	Medieval	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
Z37	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
ODN207	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
ODN271	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
Z45	Warinner et al. 2014	UK - Fewston	Victorian	0	2	No Caseins	No Caseins	No N in BLG	100.00
3DT21 (CS6)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	0	2	No Caseins	No Caseins	50.00	No Q in BLG
WG1483	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	2	No Caseins	No Caseins	No N in BLG	#N/A
RISE467	Warinner et al. 2014	Italy - Olmo di Nogara, Italy	Bronze Age	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FW53	Warinner et al. 2014	UK - Fewston	Victorian	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
Y48-B40	Warinner et al. 2014	Germany - Dalheim	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
B78	Warinner et al. 2014	Germany - Dalheim	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
B17	Warinner et al. 2014	Germany - Dalheim	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
B78	Warinner et al. 2014	Germany - Dalheim	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
Z36	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
ODN424	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
JH5	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
Z44	Warinner et al. 2014	UK - Fewston	Victorian	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW435	Warinner et al. 2014	UK - Fewston	Victorian	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
NEM099	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
WG1566	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
WG1585	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
RISE466	Warinner et al. 2014	Italy - Olmo di Nogara, Italy	Bronze Age	3	0	100.00	100.00	No BLG	No BLG
Z6	Warinner et al. 2014	Switzerland - Zurich	Modern	2	0	0.00	No Q in Caseins	No BLG	No BLG
RISE473	Warinner et al. 2014	Germany - Regensburg-Dechbetten	Bronze Age	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
SCR264	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
JH19	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
JH8	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
OX10	Warinner et al. 2014	UK - Leicester	Roman	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
RISE 460	Warinner et al. 2014	Denmark - Øster Harup	Neolithic	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
Z8	Warinner et al. 2014	Switzerland - Zurich	Modern	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
WG1252	Warinner et al. 2014	UK - Wighill (Syningthwaite Priory Farm)	Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
RISE307	Warinner et al. 2014	Russia - West Caucasus, Marchenkova Gora	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE307	Warinner et al. 2014	Russia - West Caucasus, Marchenkova Gora	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE417	Warinner et al. 2014	Armenia - Hatsarat	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE420	Warinner et al. 2014	Armenia - Noraduz	Iron Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE415	Warinner et al. 2014	Armenia - Nerkin Getashen	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE414	Warinner et al. 2014	Armenia - Noraduz	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
RISE472	Warinner et al. 2014	Germany - Nersingen	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
Y47-B27	Warinner et al. 2014	Germany - Dalheim	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
Y49-B85	Warinner et al. 2014	Germany - Dalheim	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
G12	Warinner et al. 2014	Germany - Dalheim	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
Z35	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG

ID	Paper	Site	Age	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
Z38	Warinner et al. 2014	Greenland - W51 Sandnes	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
ODN19	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
ODN361	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
ODN458	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
ODN98	Warinner et al. 2014	Italy - Olmo di Nogara	Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR277	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR250	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR323	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR5028	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR5042	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR5070	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
SCR832	Warinner et al. 2014	Italy - Isola Sacra (Portus Romae)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
JH1	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH10	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH11	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH12	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH13	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH14	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH15	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH16	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH17	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH18	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH2	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH20	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH3	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH4	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH5	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH6	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH7	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH1	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH10	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH11	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH12	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH13	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH14	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH15	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH16	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH17	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH18	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH19	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH2	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH20	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH3	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH4	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH6	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH7	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH8	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
JH9	Warinner et al. 2014	St Helena - Rupert's Valley	Victorian	0	0	No Caseins	No Caseins	No BLG	No BLG
OX01	Warinner et al. 2014	UK - Leicester	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX03	Warinner et al. 2014	UK - Leicester	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX05	Warinner et al. 2014	UK - Leicester	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX6	Warinner et al. 2014	UK - Leicester	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX09	Warinner et al. 2014	UK - Leicester	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG

ID	Paper	Site	Age	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
RISE 432	Warinner et al. 2014	Denmark - Gjerrild	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
Z7	Warinner et al. 2014	Switzerland - Zurich	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
3DT26 (CS5)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT3 (CS7)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT7 (CS8)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT21 (CS9)	Warinner et al. 2014	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
ML1489	Warinner et al. 2014	UK - Melton	Iron Age	0	0	No Caseins	No Caseins	No BLG	No BLG
ML3890	Warinner et al. 2014	UK - Melton	Iron Age	0	0	No Caseins	No Caseins	No BLG	No BLG
NEM093	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS410	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS262	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS325	Warinner et al. 2014	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
VVH03	Mays et al. 2018	UK - West Amesbury Farm	Middle Bronze Age	3	4	0.00	0.00	94.64	69.97
VVH04	Mays et al. 2018	UK - West Amesbury Farm	Middle Neolithic	2	8	No N in Caseins	No Q in Caseins	67.17	46.38
VVH02	Mays et al. 2018	UK - West Amesbury Farm	Middle Bronze Age	2	3	No N in Caseins	100.00	81.87	32.93
ARS016	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	6	53	100.00	100.00	58.61	55.41
ARS018	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	4	10	No N in Caseins	No Q in Caseins	100.00	74.35
ARS008	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	2	13	No N in Caseins	No Q in Caseins	50.00	50.00
ARS017	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	1	8	Not enough unique peptides	Not enough unique peptides	33.33	50.00
ARS007	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	1	3	Not enough unique peptides	Not enough unique peptides	No N in BLG	No Q in BLG
ARS010	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	0	10	No Caseins	No Caseins	33.33	100.00
ARS014	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	0	4	No Caseins	No Caseins	No N in BLG	No Q in BLG
ARS009	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
ARS013	Jeong et al. 2018	Mongolia - Khövsgöl	Late Bronze Age	0	0	No Caseins	No Caseins	No BLG	No BLG
CW23 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	39	16	10.18	63.38	37.76	31.70
CW8 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	23	16	2.75	65.94	38.42	20.88
CW21 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	17	5	27.44	86.54	0.00	75.73
CW27 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	13	0	20.00	62.93	No BLG	No BLG
CW18 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	12	3	42.95	79.46	44.89	0.00
CW20 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	9	6	71.43	37.50	66.67	66.67
CW11 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7	4	No N in Caseins	85.07	100.00	66.67
CW10 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	7	1	0.00	33.33	Not enough unique peptides	Not enough unique peptides
CW22 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	6	89	0.00	100.00	52.36	34.36
CW24 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	4	3	0.00	100.00	40.48	33.33
CW18 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
CW22 Calcite	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	8	No Caseins	No Caseins	33.52	42.66
CW22 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	7	No Caseins	No Caseins	87.88	57.66
CW10 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
CW20 Outer cera	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
CW10 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
CW20 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
CW24 Ceramic M	Hendy et al. 2018b	Turkey - Çatalhöyük	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
SP2300	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	4	3	50.00	100.00	100.00	100.00
UT111	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	3	7	0.00	0.00	100.00	-100.00
FW98	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	3	3	No N in Caseins	100.00	100.00	66.67
FAO18	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	3	1	0.00	No Q in Caseins	Not enough unique peptides	Not enough unique peptides
CS7	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	3	1	0.00	100.00	Not enough unique peptides	Not enough unique peptides
UT73	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	3	1	0.00	100.00	Not enough unique peptides	Not enough unique peptides
ML1032	Hendy et al. 2018a	UK - Melton	Iron Age	2	3	No N in Caseins	50.00	No N in BLG	No Q in BLG
FW192	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	2	2	No N in Caseins	66.67	100.00	100.00
FAO11	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	2	0.00	100.00	100.00	66.67
FAO17	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	2	0.00	100.00	100.00	100.00
SP2468	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	2	1	0.00	0.00	Not enough unique peptides	Not enough unique peptides
FW53	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	2	1	No N in Caseins	No Q in Caseins	Not enough unique peptides	Not enough unique peptides
FAO1	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	1	0.00	100.00	Not enough unique peptides	Not enough unique peptides

ID	Paper	Site	Age	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
FAO5	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	1	0.00	No Q in Caseins	Not enough unique peptides	Not enough unique peptides
CS18	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	2	1	0.00	100.00	Not enough unique peptides	Not enough unique peptides
SP2182	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	2	0	0.00	No Q in Caseins	No BLG	No BLG
SP2301	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	2	0	0.00	100.00	No BLG	No BLG
SP2477	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	2	0	0.00	100.00	No BLG	No BLG
FAO14	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	0	0.00	88.17	No BLG	No BLG
FAO2	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	2	0	0.00	No Q in Caseins	No BLG	No BLG
CS19	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	2	0	0.00	100.00	No BLG	No BLG
FW450	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1	8	Not enough unique peptides	Not enough unique peptides	85.74	71.14
UT116	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	1	7	Not enough unique peptides	Not enough unique peptides	100.00	-100.00
FW68	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1	5	Not enough unique peptides	Not enough unique peptides	100.00	72.02
NEM18	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	1	4	Not enough unique peptides	Not enough unique peptides	60.52	100.00
FAO4	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	3	Not enough unique peptides	Not enough unique peptides	100.00	77.25
WG1082	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	1	3	Not enough unique peptides	Not enough unique peptides	No N in BLG	No Q in BLG
FW268	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1	2	Not enough unique peptides	Not enough unique peptides	100.00	100.00
FW88	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	1	2	Not enough unique peptides	Not enough unique peptides	100.00	No Q in BLG
SP2369	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO10	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO12	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO13	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO3	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO7	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
FAO9	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
Z100	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
Z94	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
CS31	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1	1	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides	Not enough unique peptides
SP2295	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
SP2484	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
SP2748	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
SP2860	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
FAO16	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
FAO8	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
CS40	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
CS12	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
RLH 131	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
RLH 397	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
WG1252	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
OX10	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	1	0	Not enough unique peptides	Not enough unique peptides	No BLG	No BLG
OX12	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	7	No Caseins	No Caseins	83.88	71.15
OX04	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	6	No Caseins	No Caseins	100.00	100.00
ML1823	Hendy et al. 2018a	UK - Melton	Iron Age	0	3	No Caseins	No Caseins	100.00	100.00
WG1561	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	3	No Caseins	No Caseins	50.00	100.00
3DT21 (CS6)	Hendy et al. 2018a	UK - Driffield Terrace, York	Roman	0	2	No Caseins	No Caseins	50.00	No Q in BLG
CS46	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	0	2	No Caseins	No Caseins	0.00	100.00
TKDC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	0	2	No Caseins	No Caseins	No N in BLG	100.00
WG1483	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	2	No Caseins	No Caseins	No N in BLG	No Q in BLG
SP2647	Hendy et al. 2018a	UK - Christ Church, Spitalfields, London	Post-Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW217	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW303	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW331	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW435	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
NEM099	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
TKEC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
TKFC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
UT103	Hendy et al. 2018a	USA - University of Tenneesee Forensic Anthropology Center	Modern	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides

ID	Paper	Site	Age	Total Caseins	Total BLG	Caseins N Remaining %	Caseins Q Remaining %	BLG N Remaining %	BLG Q Remaining %
WG1566	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
WG1585	Hendy et al. 2018a	UK - Wighill (Syningthwaite Priory Farm)	Medieval	0	1	No Caseins	No Caseins	Not enough unique peptides	Not enough unique peptides
FW9	Hendy et al. 2018a	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
3DT26 (CS5)	Hendy et al. 2018a	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT3 (CS7)	Hendy et al. 2018a	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT7 (CS8)	Hendy et al. 2018a	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
6DT21 (CS9)	Hendy et al. 2018a	UK - Driffield Terrace, York	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
FAO15	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
FAO6	Hendy et al. 2018a	UK - Lower St Bride's Churchyard, London (FAO90)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
ML1489	Hendy et al. 2018a	UK - Melton	Iron Age	0	0	No Caseins	No Caseins	No BLG	No BLG
ML3890	Hendy et al. 2018a	UK - Melton	Iron Age	0	0	No Caseins	No Caseins	No BLG	No BLG
NEM093	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS410	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS262	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
NBS325	Hendy et al. 2018a	UK - Norton-on-Tees	Anglo-Saxon	0	0	No Caseins	No Caseins	No BLG	No BLG
Z90	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z84	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z85	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z86	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z81	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z82	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z83	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z87	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z88	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
Z89	Hendy et al. 2018a	USA - Oklahoma Dental Patients	Modern	0	0	No Caseins	No Caseins	No BLG	No BLG
CS11	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
CS21	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
CS6	Hendy et al. 2018a	UK - Radcliffe Infirmary, Oxfordshire	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 103	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 135	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 208	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 340	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 349	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 365	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 367	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 386	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 421	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
RLH 572	Hendy et al. 2018a	UK - Royal London Hospital (RLP05)	Post-Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
TKAC	Hendy et al. 2018a	UK - St. Mary's Church, Tickhill, South Yorkshire (NGR SK 592931)	Medieval	0	0	No Caseins	No Caseins	No BLG	No BLG
OX01	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX03	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX05	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX06	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
OX09	Hendy et al. 2018a	UK - Oxford Street, Leicestershire (NGR SK 586038)	Roman	0	0	No Caseins	No Caseins	No BLG	No BLG
HH3	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	3	1	No N in Caseins	No Q in Caseins	Not enough unique peptides	Not enough unique peptides
HH610	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	1	17	Not enough unique peptides	Not enough unique peptides	1.26	0.00
HN7387	Charlton et al. 2019	UK - Hazleton North	Neolithic	1	12	Not enough unique peptides	Not enough unique peptides	4.64	45.44
HN4786	Charlton et al. 2019	UK - Hazleton North	Neolithic	0	5	No Caseins	No Caseins	0.00	No Q in BLG
HN11456	Charlton et al. 2019	UK - Hazleton North	Neolithic	0	4	No Caseins	No Caseins	0.00	0.00
BL30912	Charlton et al. 2019	UK - Banbury Lane	Neolithic	0	3	No Caseins	No Caseins	0.00	No Q in BLG
HH3188	Charlton et al. 2019	UK - Hambledon Hill	Neolithic	0	1	No Caseins	No Caseins	No N in BLG	Not enough unique peptides
HN7656	Charlton et al. 2019	UK - Hazleton North	Neolithic	0	1	No Caseins	No Caseins	Not enough unique peptides	No Q in BLG
BL13217	Charlton et al. 2019	UK - Banbury Lane	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG
HN5037-1	Charlton et al. 2019	UK - Hazleton North	Neolithic	0	0	No Caseins	No Caseins	No BLG	No BLG

Appendix B, Table 6: Deamidation and peptide count for BLG and caseins in ancient milk

Appendix B, Table 7

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
3DT21 (CS6)	Hendy et al. 2018a	75.7	0	550	UK - Driffield Terrace, York	BLG
3DT21 (CS6)	Hendy et al. 2018a	50.2	1	550	UK - Driffield Terrace, York	BLG
CS18	Hendy et al. 2018a	4300	1	162.7892793	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS18	Hendy et al. 2018a	116	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS18	Hendy et al. 2018a	75.7	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS19	Hendy et al. 2018a	4300	1	100	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS19	Hendy et al. 2018a	116	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS19	Hendy et al. 2018a	75.7	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS46	Hendy et al. 2018a	5700	1	100	UK - Radcliffe Infirmary, Oxfordshire	BLG
CS46	Hendy et al. 2018a	300	0	100	UK - Radcliffe Infirmary, Oxfordshire	BLG
CS7	Hendy et al. 2018a	4300	1	100	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS7	Hendy et al. 2018a	116	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
CS7	Hendy et al. 2018a	75.7	0	1000	UK - Radcliffe Infirmary, Oxfordshire	Casein
FAO1	Hendy et al. 2018a	4300	1	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO1	Hendy et al. 2018a	116	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO1	Hendy et al. 2018a	75.7	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO11	Hendy et al. 2018a	7400	0	157.5378945	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO11	Hendy et al. 2018a	5200	1	127.5242901	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO11	Hendy et al. 2018a	4300	1	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO11	Hendy et al. 2018a	2100	1	127.5242901	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO11	Hendy et al. 2018a	116	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO11	Hendy et al. 2018a	75.7	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO14	Hendy et al. 2018a	4300	0.8816960391	387.6487163	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO14	Hendy et al. 2018a	116	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO14	Hendy et al. 2018a	75.7	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO17	Hendy et al. 2018a	5200	1	198.0428999	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO17	Hendy et al. 2018a	5200	1	187.2211784	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO17	Hendy et al. 2018a	4300	1	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO17	Hendy et al. 2018a	2100	1	198.0428999	UK - Lower St Bride's Churchyard, London (FAO90)	BLG
FAO17	Hendy et al. 2018a	116	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO17	Hendy et al. 2018a	75.7	0	1000	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO18	Hendy et al. 2018a	116	0	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO18	Hendy et al. 2018a	75.7	0	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO2	Hendy et al. 2018a	116	0	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO2	Hendy et al. 2018a	75.7	0	100	UK - Lower St Bride's Churchyard, London (FAO90)	Casein
FAO4	Hendy et al. 2018a	5200	0.5746758444	357.7184494	UK - Lower St Bride's Churchyard, London (FAO90)	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
FAO4	Hendy et al. 2018a	5200	0.8824475298	552.7053534	UK - Lower St Bride’s Churchyard, London (FAO90)	BLG
FAO4	Hendy et al. 2018a	2100	0.8603853988	357.7184494	UK - Lower St Bride’s Churchyard, London (FAO90)	BLG
FAO5	Hendy et al. 2018a	116	0	100	UK - Lower St Bride’s Churchyard, London (FAO90)	Casein
FAO5	Hendy et al. 2018a	75.7	0	100	UK - Lower St Bride’s Churchyard, London (FAO90)	Casein
FW192	Hendy et al. 2018a	10000	1	1000	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Casein
FW192	Hendy et al. 2018a	5200	1	209.7864892	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW192	Hendy et al. 2018a	4300	1	100	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Casein
FW192	Hendy et al. 2018a	2100	1	209.7864892	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW268	Hendy et al. 2018a	5200	1	634.227649	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW268	Hendy et al. 2018a	2100	1	634.227649	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW268	Hendy et al. 2018a	75.7	1	1000	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW268	Hendy et al. 2018a	50.2	1	1000	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	6100	0.4862740942	266.3288304	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	5200	0.3102912462	430.0367496	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	5200	0.9028911234	576.4045097	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	2100	0.8575652049	430.0367496	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	800	1	271.4662086	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	75.7	1	638.3007715	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	50.2	1	638.3007715	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW450	Hendy et al. 2018a	15.7	0.5137259058	266.3288304	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW68	Hendy et al. 2018a	10000	0	472.0218656	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW68	Hendy et al. 2018a	5200	0.6774621682	649.1729718	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW68	Hendy et al. 2018a	5200	0.9233721704	238.5065289	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW68	Hendy et al. 2018a	2100	1	649.1729718	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW68	Hendy et al. 2018a	800	1	218.9943923	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW88	Hendy et al. 2018a	75.7	1	100	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW88	Hendy et al. 2018a	50.2	1	100	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW98	Hendy et al. 2018a	10000	0	1000	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW98	Hendy et al. 2018a	5200	1	523.2101892	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
FW98	Hendy et al. 2018a	4300	1	100	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	Casein
FW98	Hendy et al. 2018a	2100	1	523.2101892	UK - Church of St Michael and St Lawrence, Fewston, North Yorkshire (SE 1947 5411)	BLG
ML1032	Hendy et al. 2018a	5900	1	100	UK - Melton	Casein
ML1032	Hendy et al. 2018a	5200	0	1000	UK - Melton	Casein
ML1823	Hendy et al. 2018a	5200	1	100	UK - Melton	BLG
ML1823	Hendy et al. 2018a	75.7	1	1000	UK - Melton	BLG
ML1823	Hendy et al. 2018a	50.2	1	1000	UK - Melton	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
NEM18	Hendy et al. 2018a	7400	1	100	UK - Norton-on-Tees	BLG
NEM18	Hendy et al. 2018a	75.7	0.2103999977	222.2069683	UK - Norton-on-Tees	BLG
NEM18	Hendy et al. 2018a	50.2	1	222.2069683	UK - Norton-on-Tees	BLG
OX04	Hendy et al. 2018a	5200	1	1000	UK - Oxford Street, Leicestershire (NGR SK 586038)	BLG
OX04	Hendy et al. 2018a	5200	1	100	UK - Oxford Street, Leicestershire (NGR SK 586038)	BLG
OX04	Hendy et al. 2018a	2100	1	1000	UK - Oxford Street, Leicestershire (NGR SK 586038)	BLG
OX04	Hendy et al. 2018a	75.7	1	111.9812423	UK - Oxford Street, Leicestershire (NGR SK 586038)	BLG
OX04	Hendy et al. 2018a	50.2	1	111.9812423	UK - Oxford Street, Leicestershire (NGR SK 586038)	BLG
OX12	Hendy et al. 2018a	6100	1	1000	UK - Leicester	BLG
OX12	Hendy et al. 2018a	5200	0.4230362036	249.0074584	UK - Leicester	BLG
OX12	Hendy et al. 2018a	104	1	1000	UK - Leicester	BLG
OX12	Hendy et al. 2018a	75.7	0.5163849603	272.0137694	UK - Leicester	BLG
OX12	Hendy et al. 2018a	50.2	1	272.0137694	UK - Leicester	BLG
SP2182	Hendy et al. 2018a	116	0	100	UK - Christ Church, Spitalfields, London	Casein
SP2182	Hendy et al. 2018a	75.7	0	100	UK - Christ Church, Spitalfields, London	Casein
SP2300	Hendy et al. 2018a	10000	1	701.3498399	UK - Christ Church, Spitalfields, London	Casein
SP2300	Hendy et al. 2018a	5600	1	100	UK - Christ Church, Spitalfields, London	BLG
SP2300	Hendy et al. 2018a	5200	1	286.9410597	UK - Christ Church, Spitalfields, London	BLG
SP2300	Hendy et al. 2018a	2100	1	286.9410597	UK - Christ Church, Spitalfields, London	BLG
SP2300	Hendy et al. 2018a	116	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2300	Hendy et al. 2018a	75.7	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2300	Hendy et al. 2018a	58.6	1	642.9220869	UK - Christ Church, Spitalfields, London	Casein
SP2301	Hendy et al. 2018a	4300	1	100	UK - Christ Church, Spitalfields, London	Casein
SP2301	Hendy et al. 2018a	116	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2301	Hendy et al. 2018a	75.7	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2468	Hendy et al. 2018a	8100	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2468	Hendy et al. 2018a	391	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2477	Hendy et al. 2018a	10000	1	100	UK - Christ Church, Spitalfields, London	Casein
SP2477	Hendy et al. 2018a	116	0	1000	UK - Christ Church, Spitalfields, London	Casein
SP2477	Hendy et al. 2018a	75.7	0	1000	UK - Christ Church, Spitalfields, London	Casein
TKDC	Hendy et al. 2018a	7400	1	100	UK - St. Mary’s Church, Tickhill, South Yorkshire (NGR SK 592931)	BLG
UT111	Hendy et al. 2018a	5800	0	1000	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT111	Hendy et al. 2018a	116	0	100	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT111	Hendy et al. 2018a	116	0	100	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT111	Hendy et al. 2018a	75.7	1	1000	USA - University of Tenneesee Forensic Anthropology Center	BLG
UT111	Hendy et al. 2018a	75.7	0	100	USA - University of Tenneesee Forensic Anthropology Center	Casein

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
UT111	Hendy et al. 2018a	75.7	0	100	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT111	Hendy et al. 2018a	50.2	1	1000	USA - University of Tenneesee Forensic Anthropology Center	BLG
UT116	Hendy et al. 2018a	75.7	1	100	USA - University of Tenneesee Forensic Anthropology Center	BLG
UT116	Hendy et al. 2018a	50.2	1	100	USA - University of Tenneesee Forensic Anthropology Center	BLG
UT73	Hendy et al. 2018a	4300	1	100	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT73	Hendy et al. 2018a	116	0	1000	USA - University of Tenneesee Forensic Anthropology Center	Casein
UT73	Hendy et al. 2018a	75.7	0	1000	USA - University of Tenneesee Forensic Anthropology Center	Casein
WG1561	Hendy et al. 2018a	5200	1	1000	UK - Wighill (Syningthwaite Priory Farm)	BLG
WG1561	Hendy et al. 2018a	75.7	0	100	UK - Wighill (Syningthwaite Priory Farm)	BLG
WG1561	Hendy et al. 2018a	50.2	1	100	UK - Wighill (Syningthwaite Priory Farm)	BLG
CW10 Calcite	Hendy et al. 2018b	8500	1	118.2513628	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	7500	1	116.646505	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	6400	0	100	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	5500	0	100	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	4300	0	1000	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	116	0	391.9458469	Turkey - Çatalhöyük	Casein
CW10 Calcite	Hendy et al. 2018b	75.7	0	391.9458469	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	13000	1	105.9581509	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	10000	0	107.3194116	Turkey - Çatalhöyük	BLG
CW11 Calcite	Hendy et al. 2018b	9300	1	101.8413949	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	8300	1	103.4827444	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	6200	1	103.4827444	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	5600	1	105.9581509	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	5200	1	1000	Turkey - Çatalhöyük	BLG
CW11 Calcite	Hendy et al. 2018b	4300	0.8059216108	187.4547804	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	4100	1	101.8413949	Turkey - Çatalhöyük	Casein
CW11 Calcite	Hendy et al. 2018b	2100	1	1000	Turkey - Çatalhöyük	BLG
CW18 Calcite	Hendy et al. 2018b	12000	0.8296602564	462.9791611	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	8100	1	122.2830441	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	7200	1	149.2652699	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	5800	1	130.4417184	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	5200	0	103.2409099	Turkey - Çatalhöyük	BLG
CW18 Calcite	Hendy et al. 2018b	4300	0.7322268869	407.128076	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	116	0.288513747	177.6493682	Turkey - Çatalhöyük	Casein
CW18 Calcite	Hendy et al. 2018b	75.7	1	725.6506466	Turkey - Çatalhöyük	BLG
CW18 Calcite	Hendy et al. 2018b	75.7	0	177.6493682	Turkey - Çatalhöyük	Casein

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
CW18 Calcite	Hendy et al. 2018b	50.2	0	725.6506466	Turkey - Çatalhöyük	BLG
CW18 Calcite	Hendy et al. 2018b	15.7	1	149.2652699	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	10000	0	113.5943123	Turkey - Çatalhöyük	BLG
CW20 Calcite	Hendy et al. 2018b	7200	0	141.5464225	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	6400	0	590.7727707	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	5500	1	590.7727707	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	5200	1	191.8287054	Turkey - Çatalhöyük	BLG
CW20 Calcite	Hendy et al. 2018b	4600	0	100	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	4300	1	112.2300013	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	2100	1	191.8287054	Turkey - Çatalhöyük	BLG
CW20 Calcite	Hendy et al. 2018b	327	1	141.4091781	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	92.6	0	182.0030228	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	75.7	1	576.3316096	Turkey - Çatalhöyük	BLG
CW20 Calcite	Hendy et al. 2018b	58.6	1	141.4091781	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	50.2	0	576.3316096	Turkey - Çatalhöyük	BLG
CW20 Calcite	Hendy et al. 2018b	46.3	0	100	Turkey - Çatalhöyük	Casein
CW20 Calcite	Hendy et al. 2018b	15.7	1	141.5464225	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	12000	0.8517718533	226.9182038	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	8300	1	nan	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	6800	1	116.5776353	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	6600	1	154.9125023	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	6400	1	1000	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	6100	1	178.7234681	Turkey - Çatalhöyük	BLG
CW21 Calcite	Hendy et al. 2018b	5800	1	141.6073675	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	5500	1	1000	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	5200	1	100	Turkey - Çatalhöyük	BLG
CW21 Calcite	Hendy et al. 2018b	4300	0.6673062193	155.4156052	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	4300	0	131.6932332	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	850	1	154.9125023	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	800	0.7866167186	336.0311627	Turkey - Çatalhöyük	BLG
CW21 Calcite	Hendy et al. 2018b	116	0.1950283685	190.2880823	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	116	0	116.5776353	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	75.7	0	190.2880823	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	75.7	0	116.5776353	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	62.1	0	131.6932332	Turkey - Çatalhöyük	Casein
CW21 Calcite	Hendy et al. 2018b	32	1	155.5974555	Turkey - Çatalhöyük	Casein

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
CW22 Calcite	Hendy et al. 2018b	6100	0	124.0018796	Turkey - Çatalhöyük	BLG
CW22 Calcite	Hendy et al. 2018b	5200	0.7062380058	225.727923	Turkey - Çatalhöyük	BLG
CW22 Calcite	Hendy et al. 2018b	800	1	454.5894392	Turkey - Çatalhöyük	BLG
CW22 Calcite	Hendy et al. 2018b	15.7	0	175.3106617	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	12000	1	104.3245155	Turkey - Çatalhöyük	Casein
CW22 Ceramic n	Hendy et al. 2018b	6100	0.3610829329	106.7040504	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	6100	0.5978929157	110.5693159	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	5500	1	107.6187384	Turkey - Çatalhöyük	Casein
CW22 Ceramic n	Hendy et al. 2018b	5200	0.2949805956	102.8934096	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	5200	0.3633246386	150.3545318	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	5200	0	102.5024212	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	5200	0.5241748423	114.1082909	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	5200	0.2972252368	101.6448553	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	1900	0.3013526885	141.702048	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	1900	0.3075014421	110.4475356	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	800	0.5562635003	115.0582155	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	800	0.7696414234	223.5470431	Turkey - Çatalhöyük	BLG
CW22 Ceramic n	Hendy et al. 2018b	116	0	101.4367453	Turkey - Çatalhöyük	Casein
CW22 Ceramic n	Hendy et al. 2018b	75.7	0	101.4367453	Turkey - Çatalhöyük	Casein
CW22 Ceramic n	Hendy et al. 2018b	15.7	0.0195348333	108.0244599	Turkey - Çatalhöyük	BLG
CW22 Ceramic N	Hendy et al. 2018b	15.7	0	118.1964952	Turkey - Çatalhöyük	BLG
CW22 Outer cera	Hendy et al. 2018b	6100	0	154.308094	Turkey - Çatalhöyük	BLG
CW22 Outer cera	Hendy et al. 2018b	5200	0.4509566327	492.3971001	Turkey - Çatalhöyük	BLG
CW22 Outer cera	Hendy et al. 2018b	1900	0.5357657905	423.854382	Turkey - Çatalhöyük	BLG
CW22 Outer cera	Hendy et al. 2018b	800	1	154.308094	Turkey - Çatalhöyük	BLG
CW22 Outer cera	Hendy et al. 2018b	800	1	100	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	13000	0	118.7129442	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	12000	1	120.0742878	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	10000	1	182.1955934	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	8300	0.9098717455	345.6137832	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	8300	0.5265254883	107.0778966	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	7500	0	118.7129442	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	6800	1	127.8905272	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	6400	0.5755824829	182.9793525	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	6200	1	101.0005521	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	6100	0	103.3981476	Turkey - Çatalhöyük	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
CW23 Calcite	Hendy et al. 2018b	5500	1	182.9793525	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	5200	0.1376866357	116.9442492	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	5200	0	113.0516691	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	4300	0.6494239109	267.8161827	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	4300	0.2113997422	155.7375906	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	1900	0	114.5108593	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	800	0.8665752232	148.9476146	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	116	1	100	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	116	0.0494935925	150.6564997	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	116	0	103.3406096	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	116	0.1717961904	127.8905272	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	92	0	103.3406096	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	75.7	1	522.8490653	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	75.7	0	100	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	75.7	0	150.6564997	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	75.7	0	127.8905272	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	73.9	0	108.7474835	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	62.1	0	155.7375906	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	50.2	0	522.8490653	Turkey - Çatalhöyük	BLG
CW23 Calcite	Hendy et al. 2018b	22.5	0	131.299523	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	22.1	0	204.5180064	Turkey - Çatalhöyük	Casein
CW23 Calcite	Hendy et al. 2018b	15.7	0	102.1010424	Turkey - Çatalhöyük	BLG
CW24 Calcite	Hendy et al. 2018b	12000	1	540.6434668	Turkey - Çatalhöyük	Casein
CW24 Calcite	Hendy et al. 2018b	5200	0	148.6320858	Turkey - Çatalhöyük	BLG
CW24 Calcite	Hendy et al. 2018b	4300	1	113.1319764	Turkey - Çatalhöyük	Casein
CW24 Calcite	Hendy et al. 2018b	1900	0	167.7281681	Turkey - Çatalhöyük	BLG
CW24 Calcite	Hendy et al. 2018b	800	1	1000	Turkey - Çatalhöyük	BLG
CW24 Calcite	Hendy et al. 2018b	22.1	0	370.6172029	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	12000	0.9113353735	568.1534375	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	10000	1	130.0366347	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	8300	1	131.5624455	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	5800	1	701.6176354	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	4600	0	130.0366347	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	4300	0.7526190229	318.4378717	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	4300	0	128.375228	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	116	1	183.9067481	Turkey - Çatalhöyük	Casein

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
CW27 Calcite	Hendy et al. 2018b	75.7	0	183.9067481	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	62.1	0	128.375228	Turkey - Çatalhöyük	Casein
CW27 Calcite	Hendy et al. 2018b	22.1	0	320.4361272	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	12000	0.8617403222	180.4917257	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	10000	0	100.404165	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	10000	1	100.064618	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	8300	0.4412243163	101.0612042	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	6400	0	140.8208812	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	6200	1	100.73264	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	6100	0.2128565525	101.2562828	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	5800	1	107.9015403	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	5500	1	140.8208812	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	5200	0.1216919288	106.7911699	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	4300	0.7662544878	249.011089	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	4300	0	108.0522868	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	1900	0	101.8858874	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	800	0.7095475598	115.3011948	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	116	0.1652385983	111.6118393	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	75.7	1	190.2316401	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	75.7	0	111.6118393	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	62.1	0	108.0522868	Turkey - Çatalhöyük	Casein
CW8 Calcite	Hendy et al. 2018b	50.2	0	190.2316401	Turkey - Çatalhöyük	BLG
CW8 Calcite	Hendy et al. 2018b	22.1	0	102.1501474	Turkey - Çatalhöyük	Casein
ARS008	Jeong et al. 2018	6100	1	108.3791912	Mongolia - Khövsgöl	BLG
ARS008	Jeong et al. 2018	5200	0	101.4270588	Mongolia - Khövsgöl	BLG
ARS008	Jeong et al. 2018	104	1	108.3791912	Mongolia - Khövsgöl	BLG
ARS008	Jeong et al. 2018	75.7	0	101.1334418	Mongolia - Khövsgöl	BLG
ARS008	Jeong et al. 2018	50.2	0	101.1334418	Mongolia - Khövsgöl	BLG
ARS008	Jeong et al. 2018	48.3	1	124.6013189	Mongolia - Khövsgöl	BLG
ARS010	Jeong et al. 2018	6100	1	111.4332189	Mongolia - Khövsgöl	BLG
ARS010	Jeong et al. 2018	104	1	111.4332189	Mongolia - Khövsgöl	BLG
ARS010	Jeong et al. 2018	75.7	0	118.8827149	Mongolia - Khövsgöl	BLG
ARS010	Jeong et al. 2018	50.2	0	118.8827149	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	16000	1	191.2627789	Mongolia - Khövsgöl	Casein
ARS016	Jeong et al. 2018	8200	1	191.2627789	Mongolia - Khövsgöl	Casein
ARS016	Jeong et al. 2018	6100	1	102.2044955	Mongolia - Khövsgöl	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
ARS016	Jeong et al. 2018	6100	0.1831248447	151.6461751	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	6100	0.2520588783	101.8781351	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0.525314695	104.9343422	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0.7007028321	101.2080521	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0	100.8453067	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0	102.1159366	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	1	102.4640773	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0.6891159157	102.0358488	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5200	0.3638088371	127.8123389	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	5100	1	191.2627789	Mongolia - Khövsgöl	Casein
ARS016	Jeong et al. 2018	2100	0.6245682361	104.9343422	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	800	1	103.7828726	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	800	1	104.3213745	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	104	1	151.6461751	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	75.7	0.2463425337	107.4608146	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	75.7	0	101.2525783	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	50.2	0.467004838	107.4608146	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	50.2	0	101.2525783	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	48.3	1	1900-5	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	48.3	1	107.2225234	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	48.3	1	110.1773818	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	31.8	1	191.2627789	Mongolia - Khövsgöl	Casein
ARS016	Jeong et al. 2018	15.7	0	104.3033525	Mongolia - Khövsgöl	BLG
ARS016	Jeong et al. 2018	15.7	0	103.4793761	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	6100	1	116.2753796	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	5200	1	120.0349855	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	1900	0	120.0349855	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	1900	0.8665985011	127.8123389	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	104	1	116.2753796	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	75.7	0	180.6658476	Mongolia - Khövsgöl	BLG
ARS017	Jeong et al. 2018	50.2	0	180.6658476	Mongolia - Khövsgöl	BLG
VVH02	Mays et al. 2018	5200	0.4701959536	517.7518754	UK - West Amesbury Farm,	BLG
VVH02	Mays et al. 2018	5200	0.598573089	504.9782245	UK - West Amesbury Farm,	BLG
VVH02	Mays et al. 2018	5200	0	403.6515913	UK - West Amesbury Farm,	BLG
VVH02	Mays et al. 2018	4300	1	100	UK - West Amesbury Farm,	Casein
VVH02	Mays et al. 2018	1900	0.3957328274	517.7518754	UK - West Amesbury Farm,	BLG

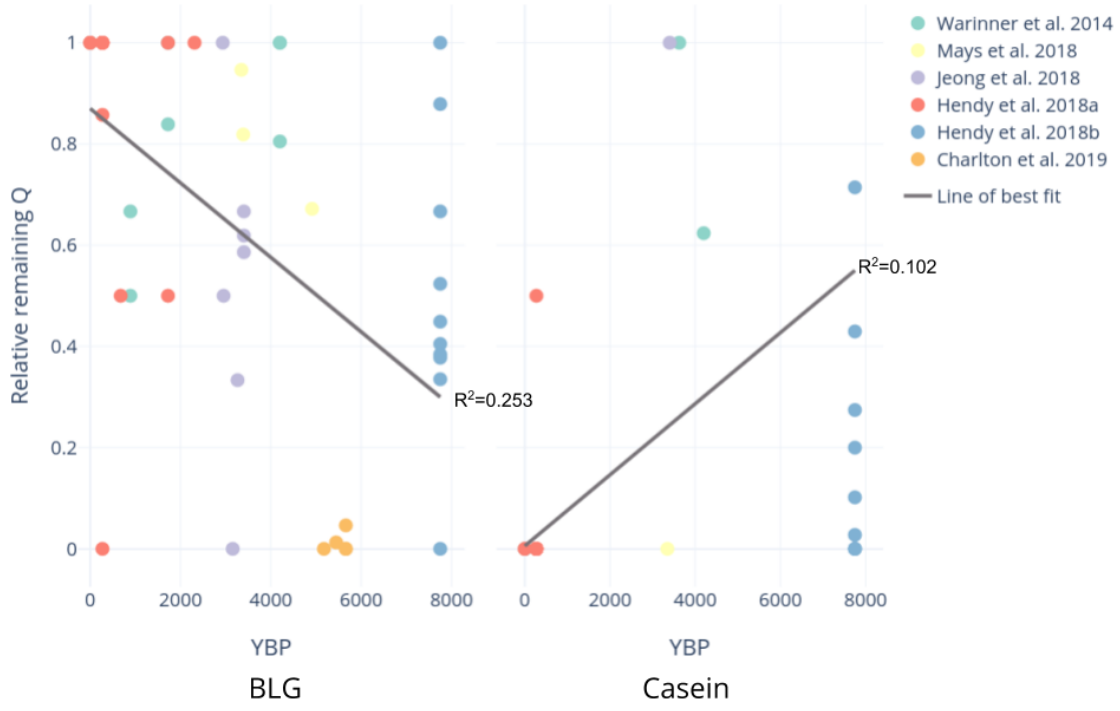
ID	Paper	Half-time	Remaining AA	Size	Site	Protein
VVH02	Mays et al. 2018	1900	0.181955882	504.9782245	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	8100	0	159.4801549	UK - West Amesbury Farm,	Casein
VVH03	Mays et al. 2018	5200	1	130.7637948	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	5200	0.3443128776	404.1862611	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	5200	1	203.7682887	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	2100	1	203.7682887	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	1900	0.5326817349	130.7637948	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	1900	0.3209630468	404.1862611	UK - West Amesbury Farm,	BLG
VVH03	Mays et al. 2018	391	0	159.4801549	UK - West Amesbury Farm,	Casein
VVH03	Mays et al. 2018	94.5	0	1000	UK - West Amesbury Farm,	Casein
VVH04	Mays et al. 2018	6100	1	100	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	5200	0.2578457053	657.6760068	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	5200	0.4139973242	264.2695926	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	5200	0	648.8161869	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	5200	0.2166928082	374.7789602	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	2100	0.4895761813	657.6760068	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	2100	0.3326082883	648.8161869	UK - West Amesbury Farm,	BLG
VVH04	Mays et al. 2018	800	1	100	UK - West Amesbury Farm,	BLG
Z45	Warinner et al. 2014	850	1	100	UK - Fewston	BLG
Z45	Warinner et al. 2014	75.7	0	100	UK - Fewston	BLG
RISE363	Warinner et al. 2015	1900	1	100	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE363	Warinner et al. 2016	5200	1	100	Hungary - Szöreg-C (Sziv Utca)	BLG
Z5	Warinner et al. 2017	2100	1	100	Switzerland - Zurich	BLG
Z5	Warinner et al. 2018	5200	1	100	Switzerland - Zurich	BLG
RISE387	Warinner et al. 2019	2100	0.563743065	558.7734333	Russia - Bulanovo	BLG
RISE387	Warinner et al. 2020	5200	0.4976185491	558.7734333	Russia - Bulanovo	BLG
Z39	Warinner et al. 2021	75.7	0	550	Greenland - Ø29a Brattahlið	BLG
Z39	Warinner et al. 2022	50.2	1	550	Greenland - Ø29a Brattahlið	BLG
RISE368	Warinner et al. 2023	2100	0.9610400379	278.2133508	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2024	5200	0.6899290596	278.2133508	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2025	1900	0.7490425916	189.4536909	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2026	5200	0.746238671	189.4536909	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2027	75.7	0	145.4133303	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2028	5200	0.8133373515	445.628656	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2029	50.2	1	145.4133303	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2030	2100	1	445.628656	Hungary - Szöreg-C (Sziv Utca)	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
RISE368	Warinner et al. 2031	1900	0.8094754387	211.3971871	Hungary - Szöreg-C (Sziv Utca)	BLG
RISE368	Warinner et al. 2032	5200	0.7685353028	211.3971871	Hungary - Szöreg-C (Sziv Utca)	BLG
OX4	Warinner et al. 2033	5200	1	1000	UK - Leicester	BLG
OX4	Warinner et al. 2034	5200	1	100	UK - Leicester	BLG
OX4	Warinner et al. 2035	75.7	1	111.9812423	UK - Leicester	BLG
OX4	Warinner et al. 2036	50.2	1	111.9812423	UK - Leicester	BLG
OX4	Warinner et al. 2037	2100	1	1000	UK - Leicester	BLG
Z40	Warinner et al. 2038	75.7	0	708.6940106	Greenland - Ø29a Brattahlið	BLG
Z40	Warinner et al. 2039	5200	0	100	Greenland - Ø29a Brattahlið	BLG
Z40	Warinner et al. 2040	50.2	1	708.6940106	Greenland - Ø29a Brattahlið	BLG
Z40	Warinner et al. 2041	2100	1	100	Greenland - Ø29a Brattahlið	BLG
OX12	Warinner et al. 2042	75.7	0.5163849603	272.0137694	UK - Leicester	BLG
OX12	Warinner et al. 2043	50.2	1	272.0137694	UK - Leicester	BLG
OX12	Warinner et al. 2044	5200	0.4230362036	249.0074584	UK - Leicester	BLG
OX12	Warinner et al. 2045	104	1	1000	UK - Leicester	BLG
OX12	Warinner et al. 2046	6100	1	1000	UK - Leicester	BLG
RISE368	Warinner et al. 2046	56.7	0.6798085418	120.8057347	Hungary - Szöreg-C (Sziv Utca)	Casein
RISE368	Warinner et al. 2047	4300	1	137.1201863	Hungary - Szöreg-C (Sziv Utca)	Casein
RISE368	Warinner et al. 2048	56.7	0.1912240844	104.4169943	Hungary - Szöreg-C (Sziv Utca)	Casein
RISE368	Warinner et al. 2049	5800	1	147.303515	Hungary - Szöreg-C (Sziv Utca)	Casein
RISE368	Warinner et al. 2050	4300	0.748643508	181.11693	Hungary - Szöreg-C (Sziv Utca)	Casein
RISE466	Warinner et al. 2051	6100	1	501.7411003	Italy - Olmo di Nogara, Italy	Casein
RISE466	Warinner et al. 2052	10000	1	1000	Italy - Olmo di Nogara, Italy	Casein
RISE466	Warinner et al. 2053	9300	1	100	Italy - Olmo di Nogara, Italy	Casein
RISE466	Warinner et al. 2054	4100	1	100	Italy - Olmo di Nogara, Italy	Casein
Z5	Warinner et al. 2055	75.7	0	443.872	Switzerland - Zurich	Casein
Z5	Warinner et al. 2056	116	0	443.872	Switzerland - Zurich	Casein
Z5	Warinner et al. 2057	4300	1	1000	Switzerland - Zurich	Casein
Z6	Warinner et al. 2058	75.7	0	100	Switzerland - Zurich	Casein
Z6	Warinner et al. 2059	116	0	100	Switzerland - Zurich	Casein
BL30912	Charlton et al. 2019	75.7	0	100	UK - Banbury Lane	BLG
BL30912	Charlton et al. 2019	50.2	0	100	UK - Banbury Lane	BLG
HH3188	Charlton et al. 2019	5200	0	100	UK - Hambledon Hill	BLG
HH610	Charlton et al. 2019	75.7	0	290.4856498	UK - Hambledon Hill	Casein
HH610	Charlton et al. 2019	116	0	290.4856498	UK - Hambledon Hill	Casein
HH610	Charlton et al. 2019	1.14	0	206.4769344	UK - Hambledon Hill	BLG

ID	Paper	Half-time	Remaining AA	Size	Site	Protein
HH610	Charlton et al. 2019	75.7	0	434.176627	UK - Hambledon Hill	BLG
HH610	Charlton et al. 2019	50.2	0.03785105596	434.176627	UK - Hambledon Hill	BLG
HH610	Charlton et al. 2019	5200	0	288.1117038	UK - Hambledon Hill	BLG
HH610	Charlton et al. 2019	5700	0	206.4769344	UK - Hambledon Hill	BLG
HN11456	Charlton et al. 2019	75.7	0	550	UK - Hazleton North	BLG
HN11456	Charlton et al. 2019	50.2	0	550	UK - Hazleton North	BLG
HN11456	Charlton et al. 2019	5200	0	578.8234769	UK - Hazleton North	BLG
HN4786	Charlton et al. 2019	75.7	0	100	UK - Hazleton North	BLG
HN4786	Charlton et al. 2019	50.2	0	100	UK - Hazleton North	BLG
HN7387	Charlton et al. 2019	75.7	0	1000	UK - Hazleton North	Casein
HN7387	Charlton et al. 2019	116	0	1000	UK - Hazleton North	Casein
HN7387	Charlton et al. 2019	75.7	0	390.824963	UK - Hazleton North	BLG
HN7387	Charlton et al. 2019	50.2	0.09271938126	390.824963	UK - Hazleton North	BLG
HN7387	Charlton et al. 2019	5200	0.4543905636	343.1789988	UK - Hazleton North	BLG
HN7656	Charlton et al. 2019	75.7	0	550	UK - Hazleton North	BLG
HN7656	Charlton et al. 2019	50.2	0.3921092285	550	UK - Hazleton North	BLG

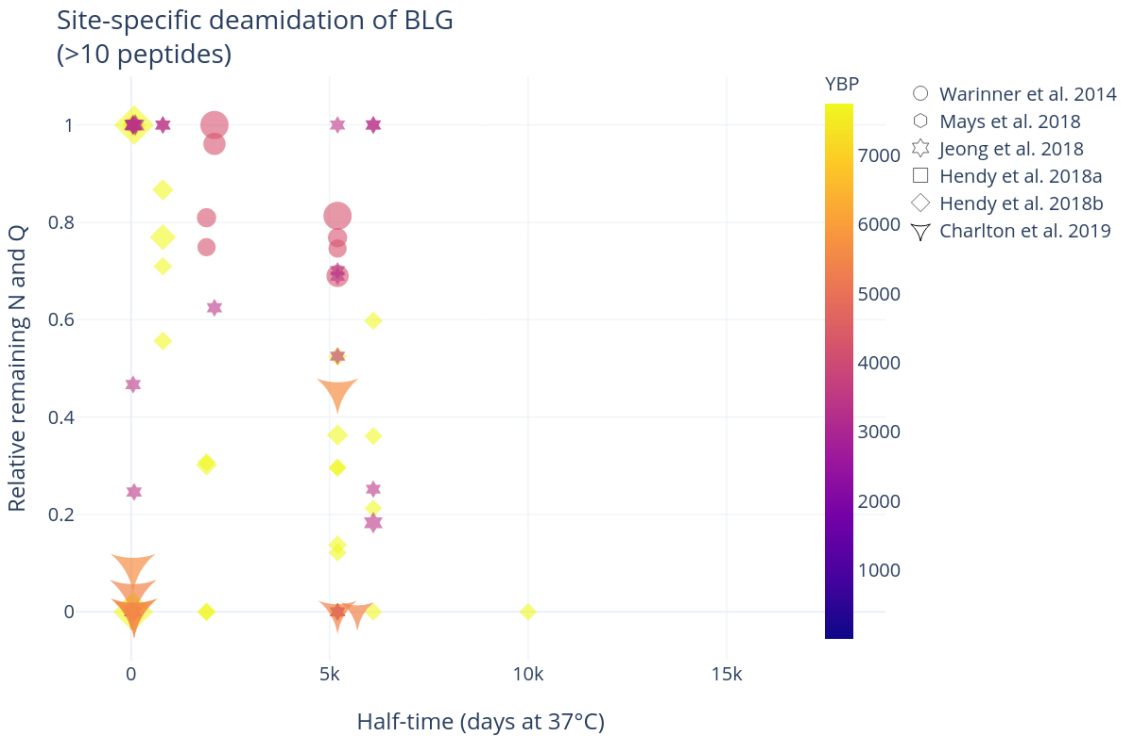
Appendix B, Table 7: Site-specific deamidation in caseins and BLG in ancient milk

Appendix B, Figure 1



Appendix B, Figure 1: Asparagine deamidation of BLG and caseins in archaeological milk

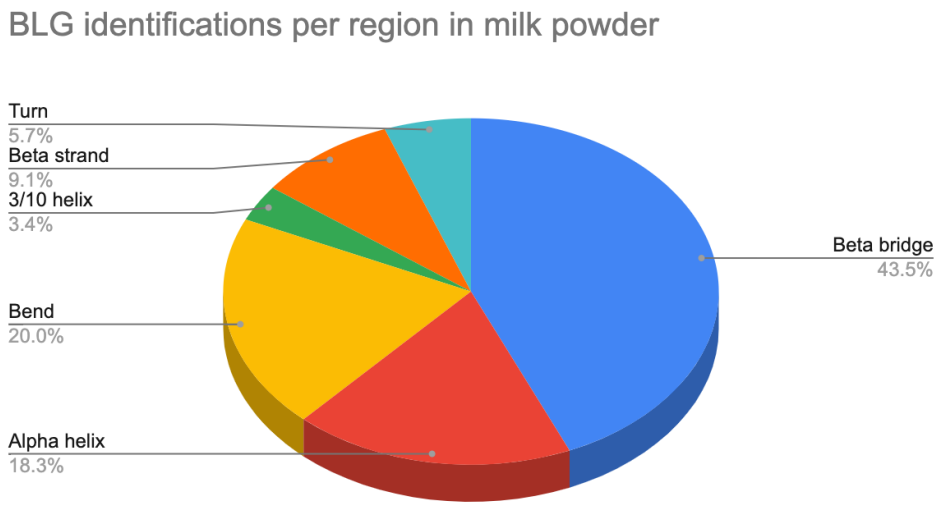
Appendix B, Figure 2



Appendix B, Figure 2: Site-specific deamidation of BLG peptides in archaeological milk, where the sample in question has more than 10 unique BLG peptides.

Appendix B, Figures 3 and 4

All BLG identifications reported in this study were further investigated in order to examine whether certain regions of BLG undergo preferential preservation. According to the Protein Data Bank (Berman, 2000), the regions of BLG are beta bridge, bend, turn, beta strand, 3/10 helix, and alpha helix. The coverage of each region in each sample was calculated and averaged in relation to the coverage across the protein as a whole.

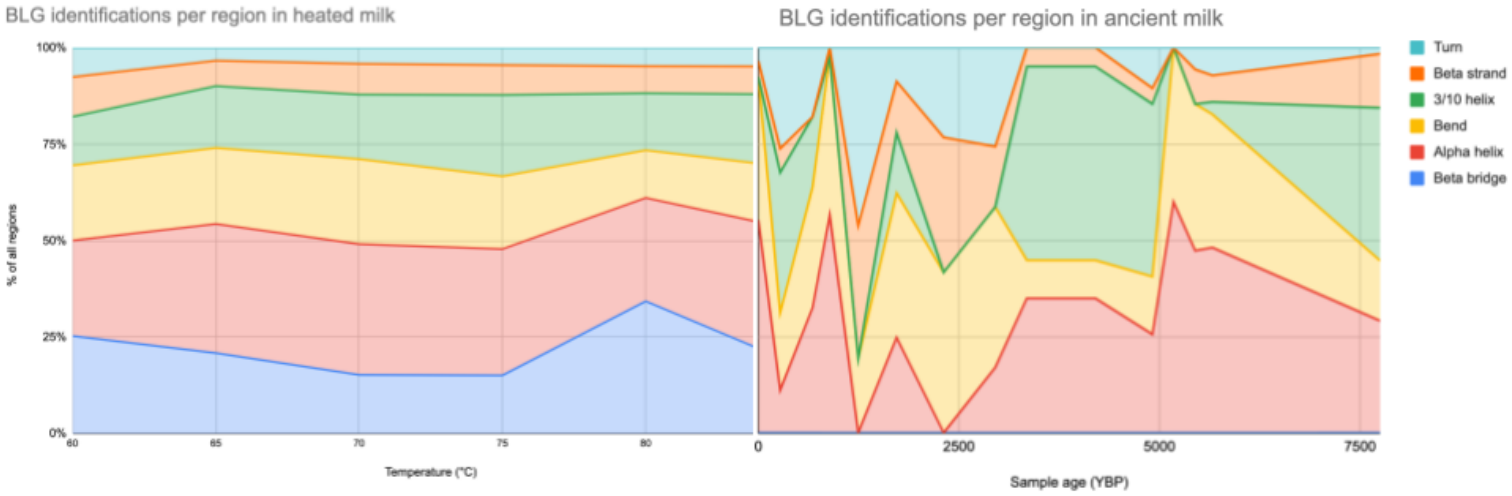


Appendix B, Figure 3: Proportional coverage at each region of BLG in milk powder

Appendix B, Figure 3 shows that the majority of peptides recovered from skim milk powder are from the beta bridge region, and that bend and alpha helical regions are also well represented. Beta bridges are also strongly represented at all heating levels in the heated milk dataset (Appendix B, Figure 4). There looks to be a slight decline in beta bridge as the heating levels increase towards 75°C, while the coverage of 3/10 helix structures slightly increases. However, the beta bridge coverage increases again at 80°C, so this pattern may not be representative of real preferential preservation.

Lastly, Appendix B, Figure 4 also shows the coverage in each region of the BLG identified in the ancient milk samples. None of the samples, even the modern dental patients, have any coverage in the beta bridge region, which is in a huge contrast to the modern milk. This implies that

neither dental calculus nor ceramic vessels bind this region, and it would not be found in genuine ancient artefacts where the proteins have bound to the surface. It could be argued that the coverage of the turn region decreases with time, and that the alpha helical regions increase, but these patterns are not robust enough to differentiate between genuine and contaminant milk.



Appendix B, Figure 4: Proportional coverage at each region of BLG

Appendix B, File 1

>sp|P04653|CASA1_SHEEP Alpha-S1-casein OS=Ovis aries OX=9940 GN=CSN1S1 PE=1 SV=3
MKLLILTCLVAVALARPKHPIKHQGLSSEVLNENLLRFVVAPFPEVFRKENINELSKDIG
SESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKITMPLW
>sp|P02669|CASK_SHEEP Kappa-casein OS=Ovis aries OX=9940 GN=CSN3 PE=1 SV=2
MMKSFFLVVTILALTLPFLGAQEQNQEQRICCEKDERFFDDKIAKYIPIQYVLSRYPSYG
LNYYQQRPVALINNQFLPYPYYAKPVAVRSPAQTLQWQVLPNAVPAKSCQDQPTAMARHP
HPHLSFMAIPPKKDQDKTEIPAINTIASAEPTVHSTPTTEAVVNAVDNPEASSESIASAP
ETNTAQVTSTEV
>sp|P11839|CASB_SHEEP Beta-casein OS=Ovis aries OX=9940 GN=CSN2 PE=1 SV=3
MKVLILACLVALALAREQEELNVVGETVESLSSSEESITHINKKIEKFQSEEQQQTEDEL
QDKIHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPKVKETMVPK
HKEMPFPKYPVEPFTESQSLTLTDVEKLHLPLPLVQSWMHQPPQPLPPTVMFPPQSVLSL
SQPKVLPVPQKAVPQRDMPIQAFLLYQEPVLGPVRGPFPILV
>sp|P04654|CASA2_SHEEP Alpha-S2-casein OS=Ovis aries OX=9940 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHKMEHVSSSEEPINISQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNADEEEYSIRSSSEESAEVAPEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL
KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL
>sp|P67976|LACB_SHEEP Beta-lactoglobulin-1/B OS=Ovis aries OX=9940 PE=1 SV=1
MKCLLLALGLALACGVQAIIVTQTMKGLDIQKVAGTWHSLAMAASDISLLDAQSAPLRVY
VEELKPTPEGNLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKY
LLFCMENSAEPEQSLACQCLVRTPEVDNEALEKFDKALKALPMHIRLAFNPTQLEGQCHV
>sp|P02758|LACB1_HORSE Beta-lactoglobulin-1 OS=Equus caballus OX=9796 GN=LGB1 PE=1 SV=3
MKCLLLALGLALMCGIQATNIPQTMQDLDLQEVAGKWHSVAMAASDISLLDSESAPLRVY

IEKLRPTPEDNLEIILREGENKGCAEKKIFAEKTESPAEFKINYLDEDTVFALDTDYKNY
LFLCMKNAATPGQSLVCQYLARTQMVDEEIMEKFRRALQPLPGRVQIVPDLTRMAERCRI
>sp|P07380|LACB2_HORSE Beta-lactoglobulin-2 OS=Equus caballus
OX=9796 GN=LGB2 PE=1 SV=3
MKCLLLALGLSLMCGNQATDIPQTMQDLDLQEVAGRWHSVAMVASDISLLDSESVPLRVY
VEELRPTPEGNLEIILREGANHACVERNIVAQKTEDPAVFTVNYQGERKISVLDTDYAHY
MFFCVGPPLPSAEHGMVCQYLARTQKVDEEVMEKFSRALQPLPGRVQIVQDPSGGQERCG
F
>sp|Q9GKK3|CASB_HORSE Beta-casein OS=Equus caballus OX=9796
GN=CSN2 PE=1 SV=3
MKILILACLVALALAREKEELNVSSETVESLSSNEPDSSSEESITHINKEKLQKFKHEGQ
QQREVERQDKISRFVQPQPVVYPYAEPVPYAVVPQSILPLAQPPILPFLQPEIMEVSQAK
ETILPKRKVMPFLKSPIVPFSERQILNPTNGENLRLPVHLIQPFMHQVPQSLLQTLMLPS
QPVLSPPQSKVAPFPQPVVPYPQRDTPVQAFLLYQDPRLGPTGELDPATQPIVAVHNPVI
V
>sp|P82187|CASK_HORSE Kappa-casein OS=Equus caballus OX=9796
GN=CSN3 PE=1 SV=2
MKSFFLVVNILALTLPFLGAEVQNQEQPTCHKNDERFFDLKTVKYIPIYYVLNSSPRYEP
IYYQHRLALLINNQHMPYQYYARPAAVRPHVQIPQWQVLPNIYPSTVVRHPCPHPSFIAI
PPKKLQEITVIPKINTIATVEPTPIPTPEPTVNNAVIPDASSEFIIASTPETTTVPVTSP
VVQKL
>tr|Q8SPR1|Q8SPR1_HORSE Alpha s1 casein OS=Equus caballus OX=9796
PE=2 SV=1
MKLLILTCLVAVALARPKLPHRQPEIIQNEQDSREKVLKERKFPSFALEYINELNRQREL
LKEKQKDEHKEYLIEDPEQQESSSTSSSEEVVPINTEQKRIPREDMLYQHTLEQLRRLSK
YNQLQLQAIHAQEQLIRMKENSQRKPMRVVNQEQAYFYLEPFQPSYQLDVYPYAAWFHPA
QIMQHVAYSPFHDTAKLIASENSEKTDIIPEW
>tr|D2KAS0|D2KAS0_HORSE Alpha-S2-casein variant B OS=Equus
caballus OX=9796 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKFKQEKYVVIPTSKESICSTSCEEAT
RNINEMESAKFPTEREEKEVEEKHHLKQLNKINQFYEKLNFLQYLQALRQPRIVLTPWDQ
TKTGDSPFIPIVNTEQLFTSEEIPKKTVDMESTEVVTEKTELTEEEKNYLKLLYYEKFTL
PQYFKIVRQHQTTMDPRSHRKTNSYQIIPVLRYF
>tr|A0A0C5DH76|A0A0C5DH76_HORSE Alpha-S2-casein variant A OS=Equus
caballus OX=9796 GN=CSN1S2 PE=2 SV=1

MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKFKQEKYVVIPTSKESICSTSCEEAT
RNINEMESAKFPTEVYSSSSSSEESAKFPTEREEKEVEEKHHLKQLNKINQFYEKLNFLQ
YLQALRQPRIVLTPWDQTKTGDSPFIPIVNTEQLFTSEEIPKKTVDMESTEVVTEKTELT
EEEKNYLKLLYYEKFTLPQYFKIVRQHQTTMDPRSHRKTNSYQIIPVLRYF
>sp|P02670|CASK_CAPHI Kappa-casein OS=Capra hircus OX=9925 GN=CSN3 PE=1 SV=2
MMKSFFLVVTILALTLPFLGAQEQNQEQPICCEKDERFFDDKIAKYIPIQYVLSRYPSYG
LNYYQQRPVALINNQFLPYPYYAKPVAVRSPAQTLQWQVLPNTVPAKSCQDQPTTLARHP
HPHLSFMAIPPKKDQDKTEVPAINTIASAEPTVHSTPTTEAIVNTVDNPEASSESIASAS
ETNTAQVTSTEV
>sp|P18626|CASA1_CAPHI Alpha-S1-casein OS=Capra hircus OX=9925 GN=CSN1S1 PE=1 SV=2
MKLLILTCLVAVALARPKHPINHRGLSPEVPNENLLRFVVAPFPEVFRKENINELSKDIG
SESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKTTMPLW
>sp|P33049|CASA2_CAPHI Alpha-S2-casein OS=Capra hircus OX=9925 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHKMEHVSSSEEPINIFQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNANEEEYSIRSSSEESAEVAPEEIKITVDDKHYQKALNEINQFYQKFPQYLQYPYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL
KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL
>sp|P33048|CASB_CAPHI Beta-casein OS=Capra hircus OX=9925 GN=CSN2 PE=2 SV=1
MKVLILACLVALAIAREQEELNVVGETVESLSSSEESITHINKKIEKFQSEEQQQTEDEL
QDKIHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPKVKETMVPK
HKEMPFPKYPVEPFTESQSLTLTDVEKLHLPLPLVQSWMHQPPQPLSPTVMFPPQSVLSL
SQPKVLPVPQKAVPQRDMPIQAFLLYQEPVLGPVRGPFPILV
>sp|P02754|LACB_BOVIN Beta-lactoglobulin OS=Bos taurus OX=9913 GN=LGB PE=1 SV=3
MKCLLLALALTCGAQALIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVE
ELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLL
FCMENSAEPEQSLACQCLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEEQCHI
>tr|B2YKY6|B2YKY6_CAPHI Beta-lactoglobulin (Fragment) OS=Capra hircus OX=9925 GN=BLG PE=3 SV=1

LNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQCLVRTPEVDKEALEKFDKALKALPMH
IRLAFNPTQLE
>sp|P02662|CASA1_BOVIN Alpha-S1-casein OS=Bos taurus OX=9913
GN=CSN1S1 PE=1 SV=2
MKLLILTCLVAVALARPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIG
SESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYK
VPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWY
YVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW
>sp|P02666|CASB_BOVIN Beta-casein OS=Bos taurus OX=9913 GN=CSN2
PE=1 SV=2
MKVLILACLVALALARELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDEL
QDKIHPFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK
HKEMPFPKYPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQPHQPLPPTVMFPPQSVLSL
SQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFPIIV
>sp|P02668|CASK_BOVIN Kappa-casein OS=Bos taurus OX=9913 GN=CSN3
PE=1 SV=1
MMKSFFLVVTILALTLPFLGAQEQNQEQPIRCEKDERFFSDKIAKYIPIQYVLSRYPSYG
LNYYQQKPVALINNQFLPYPYYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMARHP
HPHLSFMAIPPKKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPEI
NTVQVTSTAV
>sp|P02663|CASA2_BOVIN Alpha-S2-casein OS=Bos taurus OX=9913
GN=CSN1S2 PE=1 SV=2
MKFFIFTCLLAVALAKNTMEHVSSSEESIISQETYKQEKNMAINPSKENLCSTFCKEVVR
NANEEEYSIGSSSEESAEVATEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPIV
LNPWDQVKRNAVPITPTLNREQLSTSEENSKKTVDMESTEVFTKKTKLTEEEKNRLNFLK
KISQRYQKFALPQYLKTVYQHQKAMKPWIQPKTKVIPYVRYL

Appendix B, File 1

>sp|P04653|CASA1_SHEEP Alpha-S1-casein OS=Ovis aries OX=9940 GN=CSN1S1 PE=1 SV=3
MKLLILTCLVAVALARPKHPIKHQGLSSEVLNENLLRFVVAPFPEVFRKENINELSKDIG
SESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKITMPLW
>sp|P02669|CASK_SHEEP Kappa-casein OS=Ovis aries OX=9940 GN=CSN3 PE=1 SV=2
MMKSFFLVVTILALTLPFLGAQEQNQEQRICCEKDERFFDDKIAKYIPIQYVLSRYPSYG
LNYYQQRPVALINNQFLPYPYYAKPVAVRSPAQTLQWQVLPNAVPAKSCQDQPTAMARHP
HPHLSFMAIPPKKDQDKTEIPAINTIASAEPTVHSTPTTEAVVNAVDNPEASSESIASAP
ETNTAQVTSTEV
>sp|P11839|CASB_SHEEP Beta-casein OS=Ovis aries OX=9940 GN=CSN2 PE=1 SV=3
MKVLILACLVALALAREQEELNVVGETVESLSSSEESITHINKKIEKFQSEEQQQTEDEL
QDKIHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPKVKETMVPK
HKEMPFPKYPVEPFTESQSLTLTDVEKLHLPLPLVQSWMHQPPQPLPPTVMFPPQSVLSL
SQPKVLPVPQKAVPQRDMPIQAFLLYQEPVLGPVRGPFPILV
>sp|P04654|CASA2_SHEEP Alpha-S2-casein OS=Ovis aries OX=9940 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHKMEHVSSSEEPINISQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNADEEEYSIRSSSEESAEVAPEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL
KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL
>sp|P67976|LACB_SHEEP Beta-lactoglobulin-1/B OS=Ovis aries OX=9940 PE=1 SV=1
MKCLLLALGLALACGVQAIIVTQTMKGLDIQKVAGTWHSLAMAASDISLLDAQSAPLRVY
VEELKPTPEGNLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKY
LLFCMENSAEPEQSLACQCLVRTPEVDNEALEKFDKALKALPMHIRLAFNPTQLEGQCHV
>sp|P02758|LACB1_HORSE Beta-lactoglobulin-1 OS=Equus caballus OX=9796 GN=LGB1 PE=1 SV=3
MKCLLLALGLALMCGIQATNIPQTMQDLDLQEVAGKWHSVAMAASDISLLDSESAPLRVY

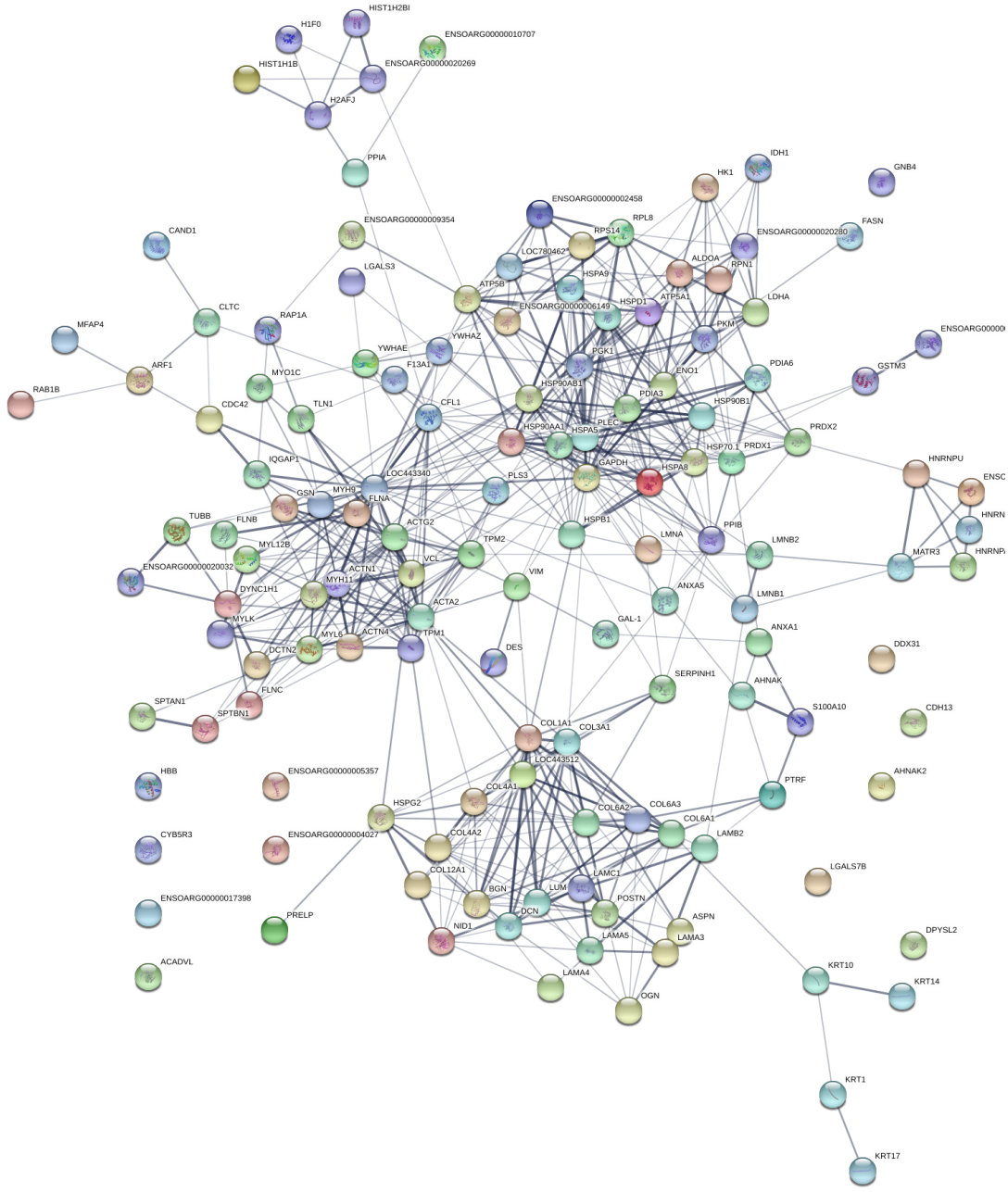
IEKLRPTPEDNLEIILREGENKGCAEKKIFAEKTESPAEFKINYLDEDTVFALDTDYKNY
LFLCMKNAATPGQSLVCQYLARTQMVDEEIMEKFRRALQPLPGRVQIVPDLTRMAERCRI
>sp|P07380|LACB2_HORSE Beta-lactoglobulin-2 OS=Equus caballus
OX=9796 GN=LGB2 PE=1 SV=3
MKCLLLALGLSLMCGNQATDIPQTMQDLDLQEVAGRWHSVAMVASDISLLDSESVPLRVY
VEELRPTPEGNLEIILREGANHACVERNIVAQKTEDPAVFTVNYQGERKISVLDTDYAHY
MFFCVGPPLPSAEHGMVCQYLARTQKVDEEVMEKFSRALQPLPGRVQIVQDPSGGQERCG
F
>sp|Q9GKK3|CASB_HORSE Beta-casein OS=Equus caballus OX=9796
GN=CSN2 PE=1 SV=3
MKILILACLVALALAREKEELNVSSETVESLSSNEPDSSSEESITHINKEKLQKFKHEGQ
QQREVERQDKISRFVQPQPVVYPYAEPVPYAVVPQSILPLAQPPILPFLQPEIMEVSQAK
ETILPKRKVMPFLKSPIVPFSERQILNPTNGENLRLPVHLIQPFMHQVPQSLLQTLMLPS
QPVLSPPQSKVAPFPQPVVPYPQRDTPVQAFLLYQDPRLGPTGELDPATQPIVAVHNPVI
V
>sp|P82187|CASK_HORSE Kappa-casein OS=Equus caballus OX=9796
GN=CSN3 PE=1 SV=2
MKSFFLVVNILALTLPFLGAEVQNQEQPTCHKNDERFFDLKTVKYIPIYYVLNSSPRYEP
IYYQHRLALLINNQHMPYQYYARPAAVRPHVQIPQWQVLPNIYPSTVVRHPCPHPSFIAI
PPKKLQEITVIPKINTIATVEPTPIPTPEPTVNNAVIPDASSEFIIASTPETTTVPVTSP
VVQKL
>tr|Q8SPR1|Q8SPR1_HORSE Alpha s1 casein OS=Equus caballus OX=9796
PE=2 SV=1
MKLLILTCLVAVALARPKLPHRQPEIIQNEQDSREKVLKERKFPSFALEYINELNRQREL
LKEKQKDEHKEYLIEDPEQQESSSTSSSEEVVPINTEQKRIPREDMLYQHTLEQLRRLSK
YNQLQLQAIHAQEQLIRMKENSQRKPMRVVNQEQAYFYLEPFQPSYQLDVYPYAAWFHPA
QIMQHVAYSPFHDTAKLIASENSEKTDIIPEW
>tr|D2KAS0|D2KAS0_HORSE Alpha-S2-casein variant B OS=Equus
caballus OX=9796 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKFKQEKYVVIPTSKESICSTSCEEAT
RNINEMESAKFPTEREEKEVEEKHHLKQLNKINQFYEKLNFLQYLQALRQPRIVLTPWDQ
TKTGDSPFIPIVNTEQLFTSEEIPKKTVDMESTEVVTEKTELTEEEKNYLKLLYYEKFTL
PQYFKIVRQHQTTMDPRSHRKTNSYQIIPVLRYF
>tr|A0A0C5DH76|A0A0C5DH76_HORSE Alpha-S2-casein variant A OS=Equus
caballus OX=9796 GN=CSN1S2 PE=2 SV=1

MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKFKQEKYVVIPTSKESICSTSCEEAT
RNINEMESAKFPTEVYSSSSSSEESAKFPTEREEKEVEEKHHLKQLNKINQFYEKLNFLQ
YLQALRQPRIVLTPWDQTKTGDSPFIPIVNTEQLFTSEEIPKKTVDMESTEVVTEKTELT
EEEKNYLKLLYYEKFTLPQYFKIVRQHQTTMDPRSHRKTNSYQIIPVLRYF
>sp|P02670|CASK_CAPHI Kappa-casein OS=Capra hircus OX=9925 GN=CSN3 PE=1 SV=2
MMKSFFLVVTILALTLPFLGAQEQNQEQPICCEKDERFFDDKIAKYIPIQYVLSRYPSYG
LNYYQQRPVALINNQFLPYPYYAKPVAVRSPAQTLQWQVLPNTVPAKSCQDQPTTLARHP
HPHLSFMAIPPKKDQDKTEVPAINTIASAEPTVHSTPTTEAIVNTVDNPEASSESIASAS
ETNTAQVTSTEV
>sp|P18626|CASA1_CAPHI Alpha-S1-casein OS=Capra hircus OX=9925 GN=CSN1S1 PE=1 SV=2
MKLLILTCLVAVALARPKHPINHRGLSPEVPNENLLRFVVAPFPEVFRKENINELSKDIG
SESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN
VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY
YLPLGTQYTDAPSFSDIPNPIGSENSGKTTMPLW
>sp|P33049|CASA2_CAPHI Alpha-S2-casein OS=Capra hircus OX=9925 GN=CSN1S2 PE=2 SV=1
MKFFIFTCLLAVALAKHKMEHVSSSEEPINIFQEIYKQEKNMAIHPRKEKLCTTSCEEVV
RNANEEEYSIRSSSEESAEVAPEEIKITVDDKHYQKALNEINQFYQKFPQYLQYPYQGPI
VLNPWDQVKRNAGPFTPTVNREQLSTSEENSKKTIDMESTEVFTKKTKLTEEEKNRLNFL
KKISQYYQKFAWPQYLKTVDQHQKAMKPWTQPKTNAIPYVRYL
>sp|P33048|CASB_CAPHI Beta-casein OS=Capra hircus OX=9925 GN=CSN2 PE=2 SV=1
MKVLILACLVALAIAREQEELNVVGETVESLSSSEESITHINKKIEKFQSEEQQQTEDEL
QDKIHPFAQAQSLVYPFTGPIPNSLPQNILPLTQTPVVVPPFLQPEIMGVPKVKETMVPK
HKEMPFPKYPVEPFTESQSLTLTDVEKLHLPLPLVQSWMHQPPQPLSPTVMFPPQSVLSL
SQPKVLPVPQKAVPQRDMPIQAFLLYQEPVLGPVRGPFPILV
>sp|P02754|LACB_BOVIN Beta-lactoglobulin OS=Bos taurus OX=9913 GN=LGB PE=1 SV=3
MKCLLLALALTCGAQALIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVE
ELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLL
FCMENSAEPEQSLACQCLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEEQCHI
>tr|B2YKY6|B2YKY6_CAPHI Beta-lactoglobulin (Fragment) OS=Capra hircus OX=9925 GN=BLG PE=3 SV=1

LNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQCLVRTPEVDKEALEKFDKALKALPMH
IRLAFNPTQLE
>sp|P02662|CASA1_BOVIN Alpha-S1-casein OS=Bos taurus OX=9913
GN=CSN1S1 PE=1 SV=2
MKLLILTCLVAVALARPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIG
SESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYK
VPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWY
YVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW
>sp|P02666|CASB_BOVIN Beta-casein OS=Bos taurus OX=9913 GN=CSN2
PE=1 SV=2
MKVLILACLVALALARELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDEL
QDKIHPFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK
HKEMPFPKYPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQPHQPLPPTVMFPPQSVLSL
SQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFPIIV
>sp|P02668|CASK_BOVIN Kappa-casein OS=Bos taurus OX=9913 GN=CSN3
PE=1 SV=1
MMKSFFLVVTILALTLPFLGAQEQNQEQPIRCEKDERFFSDKIAKYIPIQYVLSRYPSYG
LNYYQQKPVALINNQFLPYPYYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMARHP
HPHLSFMAIPPKKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPEI
NTVQVTSTAV
>sp|P02663|CASA2_BOVIN Alpha-S2-casein OS=Bos taurus OX=9913
GN=CSN1S2 PE=1 SV=2
MKFFIFTCLLAVALAKNTMEHVSSSEESIISQETYKQEKNMAINPSKENLCSTFCKEVVR
NANEEEYSIGSSSEESAEVATEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPIV
LNPWDQVKRNAVPITPTLNREQLSTSEENSKKTVDMESTEVFTKKTKLTEEEKNRLNFLK
KISQRYQKFALPQYLKTVYQHQKAMKPWIQPKTKVIPYVRYL

Appendix C, Figure 1

In order to investigate the connection between the proteins identified in all skins, a STRING analysis (Szklarczyk et al. 2018) was conducted, as shown in Appendix C, Figure 1. It shows the functional relationships between all proteins that were identified by at least two unique peptides in all skins before liming. It is clear that there are three main clusters of proteins - structural proteins, cell motility proteins, and proteins related to heat shock proteins.



Appendix C, Figure 1: STRING diagram of confidently identified proteins before liming

Appendix C, File 1

>tr|W5P481|W5P481_SHEEP Uncharacterized protein OS=Ovis aries OX=9940 GN=COL1A1 PE=4 SV=1

MFSFVDLRLLLLLAATALLTHGQEEGQEEGQEEDIPPVTCVQNGLRYHDRDVWKPVPCQI
CVCDNGNVLCDDVICDELKDCPNAKVPTPPRPVSPHPPPPPPPPTTTKQRGGRGGPCPRP
GAHRPSHSPPQGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYDEKST
GISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKNGDDG
EAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPGSPG
ENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAVGAKG
EAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPG
ARGPSGPQGPSGPPGPKGNSGEPGAPGSKGDTGAKGEPGPTGIQGPPGPAGEEGKRGARG
EPGPAGLPGPPGERGGPGSRGFPGSDGVAGPKGPAGERGAPGPAGPKGSPGEAGRPGEAG
LPGAKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPG
KAGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGEQGPAGSPGFQGLPGPAGPPG
EAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPPGPAGPRGANGAPGNDGAKGDAG
APGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGAPGKDGVRGLTGPIGPPG
PAGAPGDKGETGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAKGEPGDAG
AKGDAGPPGPAGPAGPPGPIGNVGAPGPKGARGSAGPPGATGFPGAAGRVGPPGPSGNAG
PPGPPGPAGKEGSKGPRGETGPAGRAGEVGPPGPPGPAGEKGAPGADGPAGAPGTPGPQG
IAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESG
REGAPGAEGSPGRDGAPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAG
PAGPIGPVGARGPAGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPPGPPVSMLSPSPAP
SPASSLQGPPGSAGTPGKDGLNGLPGPIGPPGPRGRTGDAGPAVSPPTLSAHGPPALKAP
SPRSRYDLSFLPQPPQEKAHDGGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEG
SRKNPARTCRDLKMCHPDWKSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVPQK
NWYISKNPKDKRHVWYGESMTGGFQVREGGQGSDPADVAIQLTFLRLMSTEASQNITYHC
KNSVAYMDQQTGNLKKALLLQGSNEIEIRAEGNSRFTYSVTYDGCTSHTGAWGKTVIEYK
TTKTSRLPIIDVAPLDVGAPDQEFGFDIGSVCFL

>tr|W5NTT7|W5NTT7_SHEEP Fibrillar collagen NC1 domain-containing protein OS=Ovis aries OX=9940 GN=COL1A2 PE=4 SV=1

MLSFVDTRTLLLLAVTSCLATCQSLQEATARKGPSGDRGPRGERGPPGPPGRDGDDGIPG
PPGPPGPPGPPGLGGNFAAQFDGKGGGPGPMGLMGPRGPPGASGAPGPQGFQGPPGEPGE
PGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGARGFPGTPGLPGFKGIRGH

NGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDGSVGP
VGPAGPIGSAGPPGFPGAPGPKGELGPVGNPGPAGPAGPRGEVGLPGLSGPVGPPGNPGA
NGLPGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGESGNKGE
PGAVGQPGPPGPSGEEGKRGSTGEIGPAGPPGPPGLRGNPGSRGLPGADGRAGVMGPAGS
RGATGPAGVRGPNGDSGRPGEPGLMGPRGFPGSPGNIGPAGKEGPAGLPGIDGRPGPIGP
AGARGEPGNIGFPGPKGPTGDPGKAGEKGHAGLAGPRGAPGPDGNNGAQGPPGLQGVQGG
KGEQGPAGPPGFQGLPGPAGTAGEAGKPGERGIPGEFGLPGPAGARGERGPPGESGAAGP
TGPIGSRGPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERGAAGIPGGKGEKGETGL
RGDVGSPGRDGARGAPGAVGAPGPAGANGDRGEAGPAGPAGPAGPRGSPGERGEVGPAGP
NGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPAGSRGDGGP
PGATGFPGAAGRTGPPGPAGISGPPGPPGPAGKEGLRGPRGDQGPVGRTGEPGAAGPPGF
VGEKGPSGEPGTAGPPGTPGPQGLLGAPGFLGLPGSRGERGLPGVAGSVGEPGPLGIAGP
PGARGPPGNVGNPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNAGPVGAAGA
PGPQGPVGPTGKHGSRGEPGPVGAVGPAGAVGPRGPSGPQGIRGDKGEPGDKGPRGLPGL
KGHNGLQGLPGLAGHHGDQGAPGAVGPAGPRGPAGPTGPAGKDGRTGQPGAVGPAGIRGS
QGSQGPAGPPGPPGPPGPPGPSGGGYDFGFDGDFYRADQPRSPASLRPKDYEVDATLKSL
NNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVYCDFSTGET
CIRAQPEDIPVKNWYRNSKAKKHVWVGETINGGTQFEYNVEGVTTKEMATQLAFMRLLAN
HASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDGCSKKTN
EWKKTIIEYKTNKPSRLPILDIAPLDIGGADQEIRLNIGPVCFK

Appendix C, Table 1

Protein	Position	Row (QStagger)	Position (QStagger)	AA (From Fasta)	AA (From MQ)	Region	N-1	N	Hydrophobicity scale value (kcal/mol)-moving towards water	Collision diameter of side chain(Å)	Longest Dimension (Å)	Side Chain Volume (Å^3)	Linear combination of hydrophobicity and collision radius	Linear combination of hydrophobicity and side chain volume	SAPs
A2	1	-1	-1	M	X	Telopeptide	-100%	-100%		5.7353	7.9317	81.657	-6.4053	-82.327	
A2	2	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	3	-1	-1	S	X	Telopeptide	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	4	-1	-1	F	X	Telopeptide	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A2	5	-1	-1	V	X	Telopeptide	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	6	-1	-1	D	X	Telopeptide	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A2	7	-1	-1	T	X	Telopeptide	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	8	-1	-1	R	X	Telopeptide	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A2	9	-1	-1	T	X	Telopeptide	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	10	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	11	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	12	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	13	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	14	-1	-1	A	X	Telopeptide	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	15	-1	-1	V	X	Telopeptide	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	16	-1	-1	T	X	Telopeptide	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	17	-1	-1	S	X	Telopeptide	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	18	-1	-1	C	X	Telopeptide	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	19	-1	-1	L	X	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	20	-1	-1	A	X	Telopeptide	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	21	-1	-1	T	X	Telopeptide	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	22	-1	-1	C	X	Telopeptide	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	23	-1	-1	Q	Q	Telopeptide	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	24	-1	-1	S	S	Telopeptide	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	25	-1	-1	L	L	Telopeptide	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	26	-1	-1	Q	Q	Telopeptide	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	27	-1	-1	E	E	Telopeptide	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	28	-1	-1	A	A	Telopeptide	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	29	-1	-1	T	T	Telopeptide	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	30	-1	-1	A	A	Telopeptide	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	31	-1	-1	R	R	Telopeptide	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A2	32	-1	-1	K	K	Telopeptide	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	33	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	34	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A2	35	-1	-1	S	S	Telopeptide	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	36	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	37	-1	-1	D	D	Telopeptide	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	

A2	38	-1	-1	R	R	Telopeptide	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	
A2	39	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	40	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	41	-1	-1	R	R	Telopeptide	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	
A2	42	-1	-1	G	X	Telopeptide	43%	25%	-1.15						
A2	43	-1	-1	E	X	Telopeptide	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112	
A2	44	-1	-1	R	X	Telopeptide	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	
A2	45	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	46	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	47	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	48	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	49	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	50	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	51	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	52	-1	-1	R	R	Telopeptide	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	
A2	53	-1	-1	D	X	Telopeptide	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	
A2	54	-1	-1	G	X	Telopeptide	43%	25%	-1.15						
A2	55	-1	-1	D	X	Telopeptide	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	
A2	56	-1	-1	D	X	Telopeptide	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	
A2	57	-1	-1	G	X	Telopeptide	43%	25%	-1.15						
A2	58	-1	-1	I	X	Telopeptide	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222	
A2	59	-1	-1	P	X	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	60	-1	-1	G	X	Telopeptide	43%	25%	-1.15						
A2	61	-1	-1	P	X	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	62	-1	-1	P	X	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	63	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	64	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	65	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	66	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	67	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	68	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	69	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	70	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	71	-1	-1	P	P	Telopeptide	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	
A2	72	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	73	-1	-1	L	L	Telopeptide	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936	
A2	74	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	75	-1	-1	G	G	Telopeptide	43%	25%	-1.15						
A2	76	-1	-1	N	N	Telopeptide	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627	
A2	77	-1	-1	F	F	Telopeptide	-85%	-81%	1.71	5.96373	8.4757	83.928	-7.6737	-85.638	
A2	78	-1	-1	A	A	Telopeptide	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	
A2	79	-1	-1	A	A	Telopeptide	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	
A2	80	-1	-1	Q	Q	N-terminal	44%	-19%	-0.77	5.51489	7.3776	74.788	-4.7448	-74.018	1
A2	81	-1	-1	F	F	N-terminal	-85%	-81%	1.71	5.96373	8.4757	83.928	-7.6737	-85.638	4
A2	82	-1	-1	D	D	N-terminal	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	4
A2	83	-1	-1	G	G	N-terminal	43%	25%	-1.15						2
A2	84	-1	-1	K	K	N-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011	5
A2	85	-1	-1	G	G	N-terminal	43%	25%	-1.15						1
A2	86	-1	-1	G	G	N-terminal	43%	25%	-1.15						2
A2	87	2	0	G	G	Overlap	43%	25%	-1.15						4
A2	88	2	1	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	3
A2	89	2	2	G	G	Overlap	43%	25%	-1.15						8
A2	90	2	3	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	3
A2	91	2	4	M	M	Overlap	-100%	-100%	0.67	5.73531	7.9317	81.657	-6.4053	-82.327	1
A2	92	2	5	G	G	Overlap	43%	25%	-1.15						1

A2	93	2	6	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	94	2	7	M	M	Overlap	-100%	-100%	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	2
A2	95	2	8	G	G	Overlap	43%	25%	-1.15						1
A2	96	2	9	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	97	2	10	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	98	2	11	G	G	Overlap	43%	25%	-1.15						1
A2	99	2	12	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	100	2	13	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	101	2	14	G	G	Overlap	43%	25%	-1.15						1
A2	102	2	15	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	103	2	16	S	S	Overlap	43%	262%	-0.46	4.2262	4.8442		-3.7662		2
A2	104	2	17	G	G	Overlap	43%	25%	-1.15						1
A2	105	2	18	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	106	2	19	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	5
A2	107	2	20	G	G	Overlap	43%	25%	-1.15						1
A2	108	2	21	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	109	2	22	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	110	2	23	G	G	Overlap	43%	25%	-1.15						1
A2	111	2	24	F	F	Overlap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	3
A2	112	2	25	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	2
A2	113	2	26	G	G	Overlap	43%	25%	-1.15						1
A2	114	2	27	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	115	2	28	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	116	2	29	G	G	Overlap	43%	25%	-1.15						1
A2	117	2	30	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	5
A2	118	2	31	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	119	2	32	G	G	Overlap	43%	25%	-1.15						1
A2	120	2	33	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A2	121	2	34	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	122	2	35	G	G	Overlap	43%	25%	-1.15						1
A2	123	2	36	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	124	2	37	T	T	Overlap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	125	2	38	G	G	Overlap	43%	25%	-1.15						1
A2	126	2	39	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	127	2	40	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	128	2	41	G	G	Overlap	43%	25%	-1.15						1
A2	129	2	42	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	130	2	43	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	4
A2	131	2	44	G	G	Overlap	43%	25%	-1.15						1
A2	132	2	45	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	133	2	46	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	134	2	47	G	G	Overlap	43%	25%	-1.15						1
A2	135	2	48	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	136	2	49	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	137	2	50	G	G	Overlap	43%	25%	-1.15						1
A2	138	2	51	K	K	Overlap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	3
A2	139	2	52	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	140	2	53	G	G	Overlap	43%	25%	-1.15						1
A2	141	2	54	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A2	142	2	55	D	D	Overlap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	3
A2	143	2	56	G	G	Overlap	43%	25%	-1.15						1
A2	144	2	57	H	H	Overlap	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	1
A2	145	2	58	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	146	2	59	G	G	Overlap	43%	25%	-1.15						1
A2	147	2	60	K	K	Overlap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2

A2	148	2	61	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	149	2	62	G	G	Overlap	43%	25%	-1.15						1
A2	150	2	63	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	151	2	64	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	152	2	65	G	G	Overlap	43%	25%	-1.15						1
A2	153	2	66	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A2	154	2	67	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A2	155	2	68	G	G	Overlap	43%	25%	-1.15						1
A2	156	2	69	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A2	157	2	70	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A2	158	2	71	G	G	Overlap	43%	25%	-1.15						1
A2	159	2	72	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	160	2	73	Q	Q	Overlap	44%	-19%	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	3
A2	161	2	74	G	G	Overlap	43%	25%	-1.15						1
A2	162	2	75	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	163	2	76	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	164	2	77	G	G	Overlap	43%	25%	-1.15						1
A2	165	2	78	F	F	Overlap	-85%	-81%	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A2	166	2	79	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	167	2	80	G	G	Gap	43%	25%	-1.15						1
A2	168	2	81	T	T	Gap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A2	169	2	82	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	170	2	83	G	G	Gap	43%	25%	-1.15						1
A2	171	2	84	L	L	Gap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	1
A2	172	2	85	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	173	2	86	G	G	Gap	43%	25%	-1.15						1
A2	174	2	87	F	F	Gap	-85%	-81%	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	2
A2	175	2	88	K	K	Gap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A2	176	2	89	G	G	Gap	43%	25%	-1.15						1
A2	177	2	90	I	I	Gap	-51%	-7%	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	2
A2	178	2	91	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	179	2	92	G	G	Gap	43%	25%	-1.15						1
A2	180	2	93	H	H	Gap	25%	-49%	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338	6
A2	181	2	94	N	N	Gap	44%	-100%	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	3
A2	182	2	95	G	G	Gap	43%	25%	-1.15						1
A2	183	2	96	L	L	Gap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	3
A2	184	2	97	D	D	Gap	54%	-13%	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	3
A2	185	2	98	G	G	Gap	43%	25%	-1.15						1
A2	186	2	99	L	L	Gap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	4
A2	187	2	100	K	K	Gap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A2	188	2	101	G	G	Gap	43%	25%	-1.15						1
A2	189	2	102	Q	Q	Gap	44%	-19%	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	3
A2	190	2	103	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	191	2	104	G	G	Gap	43%	25%	-1.15						1
A2	192	2	105	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	193	2	106	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	194	2	107	G	G	Gap	43%	25%	-1.15						1
A2	195	2	108	V	V	Gap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	3
A2	196	2	109	K	K	Gap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	3
A2	197	2	110	G	G	Gap	43%	25%	-1.15						1
A2	198	2	111	E	E	Gap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	6
A2	199	2	112	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	200	2	113	G	G	Gap	43%	25%	-1.15						1
A2	201	2	114	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	202	2	115	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1

A2	203	2	116	G	G	Gap	43%	25%	-1.15						1
A2	204	2	117	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	2
A2	205	2	118	N	N	Gap	44%	-100%	-0.85	5.10027	5.3919-	59.477	-4.2502	-58.627	2
A2	206	2	119	G	G	Gap	43%	25%	-1.15						1
A2	207	2	120	T	T	Gap	31%	9%	-0.25	4.9029'	6.2805[	51.458	-4.6529	-51.208	1
A2	208	2	121	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	209	2	122	G	G	Gap	43%	25%	-1.15						1
A2	210	2	123	Q	Q	Gap	44%	-19%	-0.77	5.5148[	7.3776(	74.788	-4.7448	-74.018	3
A2	211	2	124	T	T	Gap	31%	9%	-0.25	4.9029'	6.2805[	51.458	-4.6529	-51.208	1
A2	212	2	125	G	G	Gap	43%	25%	-1.15						1
A2	213	2	126	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	2
A2	214	2	127	R	R	Gap	50%	-30%	-1.81	6.4003'	9.7848[	57.551	-4.5903	-55.741	7
A2	215	2	128	G	X	Gap	43%	25%	-1.15						1
A2	216	2	129	L	X	Gap	-73%	-78%	1.25	5.7075[	5.7170'	72.686	-6.9575	-73.936	1
A2	217	2	130	P	X	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	218	2	131	G	X	Gap	43%	25%	-1.15						1
A2	219	2	132	E	X	Gap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A2	220	2	133	R	X	Gap	50%	-30%	-1.81	6.4003'	9.7848[	57.551	-4.5903	-55.741	1
A2	221	2	134	G	G	Gap	43%	25%	-1.15						1
A2	222	2	135	R	R	Gap	50%	-30%	-1.81	6.4003'	9.7848[	57.551	-4.5903	-55.741	1
A2	223	2	136	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578[	60.07	-5.7587	-60.53	1
A2	224	2	137	G	G	Gap	43%	25%	-1.15						1
A2	225	2	138	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	1
A2	226	2	139	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	2
A2	227	2	140	G	G	Gap	43%	25%	-1.15						1
A2	228	2	141	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	2
A2	229	2	142	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	5
A2	230	2	143	G	G	Gap	43%	25%	-1.15						1
A2	231	2	144	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	1
A2	232	2	145	R	R	Gap	50%	-30%	-1.81	6.4003'	9.7848[	57.551	-4.5903	-55.741	5
A2	233	2	146	G	G	Gap	43%	25%	-1.15						1
A2	234	2	147	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		2
A2	235	2	148	D	D	Gap	54%	-13%	-3.64	4.7470[	5.6096[	55.197	-1.1070	-51.557	1
A2	236	2	149	G	G	Gap	43%	25%	-1.15						1
A2	237	2	150	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		2
A2	238	2	151	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578[	60.07	-5.7587	-60.53	1
A2	239	2	152	G	G	Gap	43%	25%	-1.15						1
A2	240	2	153	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	241	2	154	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578[	60.07	-5.7587	-60.53	2
A2	242	2	155	G	G	Gap	43%	25%	-1.15						1
A2	243	2	156	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	244	2	157	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	3
A2	245	2	158	G	G	Gap	43%	25%	-1.15						1
A2	246	2	159	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	247	2	160	I	I	Gap	-51%	-7%	1.12	5.8093-	6.2971-	75.102	-6.9293	-76.222	1
A2	248	2	161	G	G	Gap	43%	25%	-1.15						1
A2	249	2	162	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		1
A2	250	2	163	A	A	Gap	5%	-43%	-0.5	3.9244[	3.9245(	31.474	-3.4244	-30.974	3
A2	251	2	164	G	G	Gap	43%	25%	-1.15						1
A2	252	2	165	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	3
A2	253	2	166	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	254	2	167	G	G	Gap	43%	25%	-1.15						1
A2	255	2	168	F	F	Gap	-85%	-81%	1.71	5.9637(	8.4757[	83.928	-7.6737	-85.638	1
A2	256	2	169	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186[	80.704	-5.4368	-80.564	1
A2	257	2	170	G	G	Gap	43%	25%	-1.15						1

A2	258	2	171	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	259	2	172	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	260	2	173	G	G	Gap	43%	25%	-1.15						1
A2	261	2	174	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	262	2	175	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A2	263	2	176	G	G	Gap	43%	25%	-1.15						1
A2	264	2	177	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	265	2	178	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	266	2	179	G	G	Gap	43%	25%	-1.15						1
A2	267	2	180	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	268	2	181	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A2	269	2	182	G	G	Gap	43%	25%	-1.15						1
A2	270	2	183	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	3
A2	271	2	184	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	272	2	185	G	G	Gap	43%	25%	-1.15						1
A2	273	2	186	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	274	2	187	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	275	2	188	G	G	Gap	43%	25%	-1.15						1
A2	276	2	189	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	5
A2	277	2	190	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	278	2	191	G	G	Gap	43%	25%	-1.15						1
A2	279	2	192	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	280	2	193	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	281	2	194	G	G	Gap	43%	25%	-1.15						1
A2	282	2	195	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	3
A2	283	2	196	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	1
A2	284	2	197	G	G	Gap	43%	25%	-1.15						1
A2	285	2	198	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	286	2	199	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	6
A2	287	2	200	G	G	Gap	43%	25%	-1.15						1
A2	288	2	201	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	289	2	202	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		3
A2	290	2	203	G	G	Gap	43%	25%	-1.15						1
A2	291	2	204	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	5
A2	292	2	205	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A2	293	2	206	G	G	Gap	43%	25%	-1.15						1
A2	294	2	207	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	295	2	208	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	296	2	209	G	G	Gap	43%	25%	-1.15						1
A2	297	2	210	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	298	2	211	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	299	2	212	G	G	Gap	43%	25%	-1.15						1
A2	300	2	213	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	301	2	214	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	302	2	215	G	G	Gap	43%	25%	-1.15						1
A2	303	2	216	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	4
A2	304	2	217	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	305	2	218	G	G	Gap	43%	25%	-1.15						1
A2	306	2	219	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	307	2	220	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	5
A2	308	2	221	G	G	Gap	43%	25%	-1.15						1
A2	309	2	222	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	310	2	223	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	311	2	224	G	G	Gap	43%	25%	-1.15						1
A2	312	2	225	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1

A2	313	2	226	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	314	2	227	G	G	Gap	43%	25%	-1.15						1
A2	315	2	228	V	V	Gap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	2
A2	316	2	229	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	317	2	230	G	G	Gap	43%	25%	-1.15						1
A2	318	2	231	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	319	2	232	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	320	2	233	G	G	Gap	43%	25%	-1.15						1
A2	321	2	234	L	L	Gap	-73%	-78%	1.25	5.70759	5.71707	72.686	-6.9575	-73.936	1
A2	322	2	235	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	323	5	2	G	G	Overlap	43%	25%	-1.15						1
A2	324	5	3	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	325	5	4	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	326	5	5	G	G	Overlap	43%	25%	-1.15						1
A2	327	5	6	I	I	Overlap	-51%	-7%	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	1
A2	328	5	7	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	329	5	8	G	G	Overlap	43%	25%	-1.15						1
A2	330	5	9	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	331	5	10	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	3
A2	332	5	11	G	G	Overlap	43%	25%	-1.15						1
A2	333	5	12	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	334	5	13	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A2	335	5	14	G	G	Overlap	43%	25%	-1.15						1
A2	336	5	15	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A2	337	5	16	T	T	Overlap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	4
A2	338	5	17	G	G	Overlap	43%	25%	-1.15						1
A2	339	5	18	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A2	340	5	19	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	3
A2	341	5	20	G	G	Overlap	43%	25%	-1.15						1
A2	342	5	21	L	L	Overlap	-73%	-78%	1.25	5.70759	5.71707	72.686	-6.9575	-73.936	2
A2	343	5	22	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A2	344	5	23	G	G	Overlap	43%	25%	-1.15						1
A2	345	5	24	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A2	346	5	25	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	347	5	26	G	G	Overlap	43%	25%	-1.15						1
A2	348	5	27	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	349	5	28	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	350	5	29	G	G	Overlap	43%	25%	-1.15						1
A2	351	5	30	S	S	Overlap	43%	262%	-0.46	4.22627	4.84426		-3.7662		2
A2	352	5	31	K	K	Overlap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	2
A2	353	5	32	G	G	Overlap	43%	25%	-1.15						1
A2	354	5	33	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	5
A2	355	5	34	S	S	Overlap	43%	262%	-0.46	4.22627	4.84426		-3.7662		1
A2	356	5	35	G	G	Overlap	43%	25%	-1.15						1
A2	357	5	36	N	N	Overlap	44%	-100%	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	2
A2	358	5	37	K	K	Overlap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	3
A2	359	5	38	G	G	Overlap	43%	25%	-1.15						1
A2	360	5	39	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	4
A2	361	5	40	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	362	5	41	G	G	Overlap	43%	25%	-1.15						1
A2	363	5	42	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	364	5	43	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A2	365	5	44	G	G	Overlap	43%	25%	-1.15						1
A2	366	5	45	Q	Q	Overlap	44%	-19%	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	2
A2	367	5	46	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3

A2	368	5	47	G	G	Overlap	43%	25%	-1.15						1
A2	369	5	48	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	370	5	49	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	371	5	50	G	G	Overlap	43%	25%	-1.15						1
A2	372	5	51	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	373	5	52	S	S	Overlap	43%	262%	-0.46	4.2262	4.8442		-3.7662		2
A2	374	5	53	G	G	Overlap	43%	25%	-1.15						1
A2	375	5	54	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	376	5	55	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	3
A2	377	5	56	G	G	Overlap	43%	25%	-1.15						1
A2	378	5	57	K	K	Overlap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A2	379	5	58	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	380	5	59	G	G	Overlap	43%	25%	-1.15						1
A2	381	5	60	S	S	Overlap	43%	262%	-0.46	4.2262	4.8442		-3.7662		1
A2	382	5	61	T	T	Overlap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	383	5	62	G	G	Overlap	43%	25%	-1.15						1
A2	384	5	63	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	5
A2	385	5	64	I	I	Overlap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	5
A2	386	5	65	G	G	Overlap	43%	25%	-1.15						1
A2	387	5	66	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	388	5	67	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	6
A2	389	5	68	G	G	Overlap	43%	25%	-1.15						1
A2	390	5	69	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	391	5	70	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	392	5	71	G	G	Overlap	43%	25%	-1.15						1
A2	393	5	72	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	394	5	73	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	7
A2	395	5	74	G	G	Overlap	43%	25%	-1.15						1
A2	396	5	75	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	397	5	76	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	3
A2	398	5	77	G	G	Overlap	43%	25%	-1.15						1
A2	399	5	78	N	N	Overlap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	2
A2	400	5	79	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	401	5	80	G	G	Gap	43%	25%	-1.15						1
A2	402	5	81	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		5
A2	403	5	82	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	404	5	83	G	G	Gap	43%	25%	-1.15						1
A2	405	5	84	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	406	5	85	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	407	5	86	G	G	Gap	43%	25%	-1.15						1
A2	408	5	87	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	409	5	88	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	1
A2	410	5	89	G	G	Gap	43%	25%	-1.15						1
A2	411	5	90	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	412	5	91	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	413	5	92	G	G	Gap	43%	25%	-1.15						1
A2	414	5	93	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A2	415	5	94	M	M	Gap	-100%	-100%	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	4
A2	416	5	95	G	G	Gap	43%	25%	-1.15						1
A2	417	5	96	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	418	5	97	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	419	5	98	G	G	Gap	43%	25%	-1.15						1
A2	420	5	99	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		2
A2	421	5	100	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	3
A2	422	5	101	G	G	Gap	43%	25%	-1.15						1

A2	423	5	102	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A2	424	5	103	T	T	Gap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A2	425	5	104	G	G	Gap	43%	25%	-1.15						1
A2	426	5	105	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	6
A2	427	5	106	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A2	428	5	107	G	G	Gap	43%	25%	-1.15						1
A2	429	5	108	V	V	Gap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A2	430	5	109	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	3
A2	431	5	110	G	G	Gap	43%	25%	-1.15						1
A2	432	5	111	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	6
A2	433	5	112	N	N	Gap	44%	-100%	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	2
A2	434	5	113	G	G	Gap	43%	25%	-1.15						1
A2	435	5	114	D	D	Gap	54%	-13%	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A2	436	5	115	S	S	Gap	43%	262%	-0.46	4.22627	4.84426		-3.7662		3
A2	437	5	116	G	G	Gap	43%	25%	-1.15						1
A2	438	5	117	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	439	5	118	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A2	440	5	119	G	G	Gap	43%	25%	-1.15						1
A2	441	5	120	E	E	Gap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A2	442	5	121	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	443	5	122	G	G	Gap	43%	25%	-1.15						1
A2	444	5	123	L	L	Gap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A2	445	5	124	M	M	Gap	-100%	-100%	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	1
A2	446	5	125	G	G	Gap	43%	25%	-1.15						1
A2	447	5	126	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	448	5	127	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A2	449	5	128	G	G	Gap	43%	25%	-1.15						1
A2	450	5	129	F	F	Gap	-85%	-81%	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	2
A2	451	5	130	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	452	5	131	G	G	Gap	43%	25%	-1.15						1
A2	453	5	132	S	S	Gap	43%	262%	-0.46	4.22627	4.84426		-3.7662		3
A2	454	5	133	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	455	5	134	G	G	Gap	43%	25%	-1.15						1
A2	456	5	135	N	N	Gap	44%	-100%	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	4
A2	457	5	136	I	I	Gap	-51%	-7%	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	1
A2	458	5	137	G	G	Gap	43%	25%	-1.15						1
A2	459	5	138	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	460	5	139	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	6
A2	461	5	140	G	G	Gap	43%	25%	-1.15						1
A2	462	5	141	K	K	Gap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A2	463	5	142	E	E	Gap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	4
A2	464	5	143	G	G	Gap	43%	25%	-1.15						1
A2	465	5	144	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	466	5	145	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	467	5	146	G	G	Gap	43%	25%	-1.15						1
A2	468	5	147	L	L	Gap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A2	469	5	148	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	6
A2	470	5	149	G	G	Gap	43%	25%	-1.15						1
A2	471	5	150	I	I	Gap	-51%	-7%	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	3
A2	472	5	151	D	D	Gap	54%	-13%	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A2	473	5	152	G	G	Gap	43%	25%	-1.15						1
A2	474	5	153	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	6
A2	475	5	154	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	476	5	155	G	G	Gap	43%	25%	-1.15						1
A2	477	5	156	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1

A2	478	5	157	I	I	Gap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	3
A2	479	5	158	G	G	Gap	43%	25%	-1.15						1
A2	480	5	159	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	481	5	160	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	482	5	161	G	G	Gap	43%	25%	-1.15						1
A2	483	5	162	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	484	5	163	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	3
A2	485	5	164	G	G	Gap	43%	25%	-1.15						1
A2	486	5	165	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	5
A2	487	5	166	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	488	5	167	G	G	Gap	43%	25%	-1.15						1
A2	489	5	168	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	490	5	169	I	I	Gap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	2
A2	491	5	170	G	G	Gap	43%	25%	-1.15						1
A2	492	5	171	F	F	Gap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	2
A2	493	5	172	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	494	5	173	G	G	Gap	43%	25%	-1.15						1
A2	495	5	174	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	496	5	175	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A2	497	5	176	G	G	Gap	43%	25%	-1.15						1
A2	498	5	177	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	499	5	178	T	T	Gap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	500	5	179	G	G	Gap	43%	25%	-1.15						1
A2	501	5	180	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	2
A2	502	5	181	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	503	5	182	G	G	Gap	43%	25%	-1.15						1
A2	504	5	183	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	505	5	184	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	506	5	185	G	G	Gap	43%	25%	-1.15						1
A2	507	5	186	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A2	508	5	187	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	9
A2	509	5	188	G	G	Gap	43%	25%	-1.15						1
A2	510	5	189	H	H	Gap	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	2
A2	511	5	190	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	512	5	191	G	G	Gap	43%	25%	-1.15						1
A2	513	5	192	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	5
A2	514	5	193	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	515	5	194	G	G	Gap	43%	25%	-1.15						1
A2	516	5	195	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	517	5	196	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	518	5	197	G	G	Gap	43%	25%	-1.15						1
A2	519	5	198	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A2	520	5	199	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	521	5	200	G	G	Gap	43%	25%	-1.15						1
A2	522	5	201	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	523	5	202	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	2
A2	524	5	203	G	G	Gap	43%	25%	-1.15						1
A2	525	5	204	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	526	5	205	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	2
A2	527	5	206	G	G	Gap	43%	25%	-1.15						1
A2	528	5	207	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	529	5	208	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	2
A2	530	5	209	G	G	Gap	43%	25%	-1.15						1
A2	531	5	210	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	532	5	211	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2

A2	533	5	212	G	G	Gap	43%	25%	-1.15						1
A2	534	5	213	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A2	535	5	214	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	536	5	215	G	G	Gap	43%	25%	-1.15						1
A2	537	5	216	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	6
A2	538	5	217	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	3
A2	539	5	218	G	G	Gap	43%	25%	-1.15						1
A2	540	5	219	G	G	Gap	43%	25%	-1.15						3
A2	541	5	220	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	542	5	221	G	G	Gap	43%	25%	-1.15						2
A2	543	5	222	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	544	5	223	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	545	5	224	G	G	Gap	43%	25%	-1.15						1
A2	546	5	225	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	547	5	226	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	548	5	227	G	G	Gap	43%	25%	-1.15						1
A2	549	5	228	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	550	5	229	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	551	5	230	G	G	Gap	43%	25%	-1.15						1
A2	552	5	231	F	F	Gap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	1
A2	553	5	232	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	554	5	233	G	G	Gap	43%	25%	-1.15						1
A2	555	5	234	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	556	5	235	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	557	8	2	G	G	Overlap	43%	25%	-1.15						1
A2	558	8	3	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	559	8	4	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	560	8	5	G	G	Overlap	43%	25%	-1.15						1
A2	561	8	6	T	T	Overlap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	562	8	7	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	563	8	8	G	G	Overlap	43%	25%	-1.15						1
A2	564	8	9	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	4
A2	565	8	10	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	566	8	11	G	G	Overlap	43%	25%	-1.15						1
A2	567	8	12	K	K	Overlap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A2	568	8	13	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	5
A2	569	8	14	G	G	Overlap	43%	25%	-1.15						1
A2	570	8	15	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A2	571	8	16	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	3
A2	572	8	17	G	G	Overlap	43%	25%	-1.15						1
A2	573	8	18	I	I	Overlap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	1
A2	574	8	19	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	575	8	20	G	G	Overlap	43%	25%	-1.15						1
A2	576	8	21	E	E	Overlap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	4
A2	577	8	22	F	F	Overlap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	3
A2	578	8	23	G	G	Overlap	43%	25%	-1.15						1
A2	579	8	24	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	580	8	25	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	581	8	26	G	G	Overlap	43%	25%	-1.15						1
A2	582	8	27	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	583	8	28	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	584	8	29	G	G	Overlap	43%	25%	-1.15						1
A2	585	8	30	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	586	8	31	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	587	8	32	G	G	Overlap	43%	25%	-1.15						1

A2	588	8	33	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	5
A2	589	8	34	R	R	Overlap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	1
A2	590	8	35	G	G	Overlap	43%	25%	-1.15						1
A2	591	8	36	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	592	8	37	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	593	8	38	G	G	Overlap	43%	25%	-1.15						2
A2	594	8	39	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	4
A2	595	8	40	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		1
A2	596	8	41	G	G	Overlap	43%	25%	-1.15						1
A2	597	8	42	A	A	Overlap	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	3
A2	598	8	43	A	A	Overlap	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	1
A2	599	8	44	G	G	Overlap	43%	25%	-1.15						1
A2	600	8	45	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A2	601	8	46	T	T	Overlap	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	3
A2	602	8	47	G	G	Overlap	43%	25%	-1.15						1
A2	603	8	48	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	604	8	49	I	I	Overlap	-51%	-7%	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	5
A2	605	8	50	G	G	Overlap	43%	25%	-1.15						1
A2	606	8	51	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		3
A2	607	8	52	R	R	Overlap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	5
A2	608	8	53	G	G	Overlap	43%	25%	-1.15						1
A2	609	8	54	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A2	610	8	55	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		1
A2	611	8	56	G	G	Overlap	43%	25%	-1.15						1
A2	612	8	57	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	613	8	58	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A2	614	8	59	G	G	Overlap	43%	25%	-1.15						1
A2	615	8	60	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A2	616	8	61	D	D	Overlap	54%	-13%	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	2
A2	617	8	62	G	G	Overlap	43%	25%	-1.15						1
A2	618	8	63	N	N	Overlap	44%	-100%	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	2
A2	619	8	64	K	K	Overlap	-67%	40%	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	2
A2	620	8	65	G	G	Overlap	43%	25%	-1.15						1
A2	621	8	66	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A2	622	8	67	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	623	8	68	G	G	Overlap	43%	25%	-1.15						1
A2	624	8	69	V	V	Overlap	-100%	-46%	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	1
A2	625	8	70	V	V	Overlap	-100%	-46%	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	2
A2	626	8	71	G	G	Overlap	43%	25%	-1.15						1
A2	627	8	72	A	A	Overlap	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	5
A2	628	8	73	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	5
A2	629	8	74	G	G	Overlap	43%	25%	-1.15						1
A2	630	8	75	T	T	Overlap	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	1
A2	631	8	76	A	A	Overlap	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	3
A2	632	8	77	G	G	Overlap	43%	25%	-1.15						1
A2	633	8	78	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	4
A2	634	8	79	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		3
A2	635	8	80	G	G	Gap	43%	25%	-1.15						1
A2	636	8	81	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A2	637	8	82	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		4
A2	638	8	83	G	G	Gap	43%	25%	-1.15						1
A2	639	8	84	L	L	Gap	-73%	-78%	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936	3
A2	640	8	85	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	4
A2	641	8	86	G	G	Gap	43%	25%	-1.15						1
A2	642	8	87	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	4

A2	643	8	88	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	1
A2	644	8	89	G	G	Gap	43%	25%	-1.15						1
A2	645	8	90	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	646	8	91	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	647	8	92	G	G	Gap	43%	25%	-1.15						1
A2	648	8	93	I	I	Gap	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222	4
A2	649	8	94	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	5
A2	650	8	95	G	G	Gap	43%	25%	-1.15						1
A2	651	8	96	G	G	Gap	43%	25%	-1.15						4
A2	652	8	97	K	K	Gap	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011	1
A2	653	8	98	G	G	Gap	43%	25%	-1.15						1
A2	654	8	99	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112	2
A2	655	8	100	K	K	Gap	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011	1
A2	656	8	101	G	G	Gap	43%	25%	-1.15						1
A2	657	8	102	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112	2
A2	658	8	103	T	T	Gap	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208	1
A2	659	8	104	G	G	Gap	43%	25%	-1.15						1
A2	660	8	105	L	L	Gap	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936	3
A2	661	8	106	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	4
A2	662	8	107	G	G	Gap	43%	25%	-1.15						1
A2	663	8	108	D	D	Gap	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	5
A2	664	8	109	V	V	Gap	-100%	-46%	0.46	5.29870	6.9578	60.07	-5.7587	-60.53	1
A2	665	8	110	G	G	Gap	43%	25%	-1.15						1
A2	666	8	111	S	S	Gap	43%	262%	-0.46	4.22627	4.8442		-3.7662		2
A2	667	8	112	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	7
A2	668	8	113	G	G	Gap	43%	25%	-1.15						1
A2	669	8	114	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	4
A2	670	8	115	D	D	Gap	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	4
A2	671	8	116	G	G	Gap	43%	25%	-1.15						1
A2	672	8	117	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	673	8	118	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	1
A2	674	8	119	G	G	Gap	43%	25%	-1.15						1
A2	675	8	120	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	676	8	121	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	1
A2	677	8	122	G	G	Gap	43%	25%	-1.15						1
A2	678	8	123	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	5
A2	679	8	124	V	V	Gap	-100%	-46%	0.46	5.29870	6.9578	60.07	-5.7587	-60.53	1
A2	680	8	125	G	G	Gap	43%	25%	-1.15						1
A2	681	8	126	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	682	8	127	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	3
A2	683	8	128	G	G	Gap	43%	25%	-1.15						1
A2	684	8	129	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	2
A2	685	8	130	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	686	8	131	G	G	Gap	43%	25%	-1.15						1
A2	687	8	132	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	1
A2	688	8	133	N	N	Gap	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627	4
A2	689	8	134	G	G	Gap	43%	25%	-1.15						1
A2	690	8	135	D	D	Gap	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557	3
A2	691	8	136	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	4
A2	692	8	137	G	G	Gap	43%	25%	-1.15						1
A2	693	8	138	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112	3
A2	694	8	139	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	2
A2	695	8	140	G	G	Gap	43%	25%	-1.15						1
A2	696	8	141	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564	1
A2	697	8	142	A	A	Gap	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974	4

A2	698	8	143	G	G	Gap	43%	25%	-1.15						1
A2	699	8	144	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	700	8	145	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	701	8	146	G	G	Gap	43%	25%	-1.15						1
A2	702	8	147	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	703	8	148	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	704	8	149	G	G	Gap	43%	25%	-1.15						1
A2	705	8	150	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	706	8	151	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	5
A2	707	8	152	G	G	Gap	43%	25%	-1.15						1
A2	708	8	153	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		4
A2	709	8	154	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	710	8	155	G	G	Gap	43%	25%	-1.15						1
A2	711	8	156	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	6
A2	712	8	157	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	713	8	158	G	G	Gap	43%	25%	-1.15						1
A2	714	8	159	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	715	8	160	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	1
A2	716	8	161	G	G	Gap	43%	25%	-1.15						1
A2	717	8	162	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	718	8	163	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	719	8	164	G	G	Gap	43%	25%	-1.15						1
A2	720	8	165	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	721	8	166	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	2
A2	722	8	167	G	G	Gap	43%	25%	-1.15						1
A2	723	8	168	F	F	Gap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	1
A2	724	8	169	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	725	8	170	G	G	Gap	43%	25%	-1.15						1
A2	726	8	171	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	727	8	172	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	728	8	173	G	G	Gap	43%	25%	-1.15						1
A2	729	8	174	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	730	8	175	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	731	8	176	G	G	Gap	43%	25%	-1.15						1
A2	732	8	177	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	3
A2	733	8	178	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	734	8	179	G	G	Gap	43%	25%	-1.15						1
A2	735	8	180	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	736	8	181	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	737	8	182	G	G	Gap	43%	25%	-1.15						1
A2	738	8	183	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	739	8	184	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	740	8	185	G	X	Gap	43%	25%	-1.15						1
A2	741	8	186	T	X	Gap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	742	8	187	K	X	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	743	8	188	G	G	Gap	43%	25%	-1.15						1
A2	744	8	189	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	745	8	190	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	746	8	191	G	G	Gap	43%	25%	-1.15						1
A2	747	8	192	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A2	748	8	193	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	749	8	194	G	G	Gap	43%	25%	-1.15						1
A2	750	8	195	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	751	8	196	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A2	752	8	197	G	G	Gap	43%	25%	-1.15						1

A2	753	8	198	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	5
A2	754	8	199	T	T	Gap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	3
A2	755	8	200	G	G	Gap	43%	25%	-1.15						1
A2	756	8	201	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	757	8	202	V	V	Gap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	7
A2	758	8	203	G	G	Gap	43%	25%	-1.15						1
A2	759	8	204	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A2	760	8	205	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A2	761	8	206	G	G	Gap	43%	25%	-1.15						1
A2	762	8	207	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	763	8	208	S	S	Gap	43%	262%	-0.46	4.22627	4.84426		-3.7662		3
A2	764	8	209	G	G	Gap	43%	25%	-1.15						1
A2	765	8	210	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	766	8	211	N	N	Gap	44%	-100%	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	3
A2	767	8	212	G	G	Gap	43%	25%	-1.15						1
A2	768	8	213	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	769	8	214	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	770	8	215	G	G	Gap	43%	25%	-1.15						1
A2	771	8	216	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	772	8	217	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A2	773	8	218	G	G	Gap	43%	25%	-1.15						1
A2	774	8	219	S	S	Gap	43%	262%	-0.46	4.22627	4.84426		-3.7662		1
A2	775	8	220	R	R	Gap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	5
A2	776	8	221	G	G	Gap	43%	25%	-1.15						1
A2	777	8	222	D	D	Gap	54%	-13%	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	6
A2	778	8	223	G	G	Gap	43%	25%	-1.15						1
A2	779	8	224	G	G	Gap	43%	25%	-1.15						1
A2	780	8	225	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	781	8	226	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	782	8	227	G	G	Gap	43%	25%	-1.15						1
A2	783	8	228	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	784	8	229	T	T	Gap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A2	785	8	230	G	G	Gap	43%	25%	-1.15						1
A2	786	8	231	F	F	Gap	-85%	-81%	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	3
A2	787	8	232	P	P	Gap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	788	8	233	G	G	Gap	43%	25%	-1.15						1
A2	789	8	234	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	790	8	235	A	A	Gap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	791	11	2	G	G	Overlap	43%	25%	-1.15						1
A2	792	11	3	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	793	11	4	T	T	Overlap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	2
A2	794	11	5	G	G	Overlap	43%	25%	-1.15						1
A2	795	11	6	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	796	11	7	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	797	11	8	G	G	Overlap	43%	25%	-1.15						1
A2	798	11	9	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	799	11	10	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A2	800	11	11	G	G	Overlap	43%	25%	-1.15						1
A2	801	11	12	I	I	Overlap	-51%	-7%	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	1
A2	802	11	13	S	S	Overlap	43%	262%	-0.46	4.22627	4.84426		-3.7662		2
A2	803	11	14	G	G	Overlap	43%	25%	-1.15						1
A2	804	11	15	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	805	11	16	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	806	11	17	G	G	Overlap	43%	25%	-1.15						1
A2	807	11	18	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1

A2	808	11	19	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	809	11	20	G	G	Overlap	43%	25%	-1.15						1
A2	810	11	21	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A2	811	11	22	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A2	812	11	23	G	G	Overlap	43%	25%	-1.15						1
A2	813	11	24	K	K	Overlap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	2
A2	814	11	25	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	3
A2	815	11	26	G	G	Overlap	43%	25%	-1.15						1
A2	816	11	27	L	L	Overlap	-73%	-78%	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	1
A2	817	11	28	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A2	818	11	29	G	G	Overlap	43%	25%	-1.15						1
A2	819	11	30	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A2	820	11	31	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	821	11	32	G	G	Overlap	43%	25%	-1.15						1
A2	822	11	33	D	D	Overlap	54%	-13%	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A2	823	11	34	Q	Q	Overlap	44%	-19%	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A2	824	11	35	G	G	Overlap	43%	25%	-1.15						1
A2	825	11	36	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	826	11	37	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	2
A2	827	11	38	G	G	Overlap	43%	25%	-1.15						1
A2	828	11	39	R	R	Overlap	50%	-30%	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A2	829	11	40	T	T	Overlap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	4
A2	830	11	41	G	G	Overlap	43%	25%	-1.15						1
A2	831	11	42	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A2	832	11	43	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	833	11	44	G	G	Overlap	43%	25%	-1.15						1
A2	834	11	45	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A2	835	11	46	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	6
A2	836	11	47	G	G	Overlap	43%	25%	-1.15						1
A2	837	11	48	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A2	838	11	49	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	6
A2	839	11	50	G	G	Overlap	43%	25%	-1.15						1
A2	840	11	51	F	F	Overlap	-85%	-81%	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	2
A2	841	11	52	V	V	Overlap	-100%	-46%	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	3
A2	842	11	53	G	G	Overlap	43%	25%	-1.15						1
A2	843	11	54	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A2	844	11	55	K	K	Overlap	-67%	40%	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	5
A2	845	11	56	G	G	Overlap	43%	25%	-1.15						1
A2	846	11	57	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	847	11	58	S	S	Overlap	43%	262%	-0.46	4.22627	4.84426		-3.7662		1
A2	848	11	59	G	G	Overlap	43%	25%	-1.15						1
A2	849	11	60	E	E	Overlap	-53%	-14%	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A2	850	11	61	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A2	851	11	62	G	G	Overlap	43%	25%	-1.15						1
A2	852	11	63	T	T	Overlap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A2	853	11	64	A	A	Overlap	5%	-43%	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A2	854	11	65	G	G	Overlap	43%	25%	-1.15						1
A2	855	11	66	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	856	11	67	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A2	857	11	68	G	G	Overlap	43%	25%	-1.15						1
A2	858	11	69	T	T	Overlap	31%	9%	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	3
A2	859	11	70	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A2	860	11	71	G	G	Overlap	43%	25%	-1.15						1
A2	861	11	72	P	P	Overlap	-10%	-7%	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A2	862	11	73	Q	Q	Overlap	44%	-19%	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	4

A2	863	11	74	G	G	Overlap	43%	25%	-1.15						1
A2	864	11	75	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	865	11	76	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	866	11	77	G	G	Overlap	43%	25%	-1.15						1
A2	867	11	78	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	868	11	79	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	869	11	80	G	G	Gap	43%	25%	-1.15						1
A2	870	11	81	F	F	Gap	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	6
A2	871	11	82	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	872	11	83	G	G	Gap	43%	25%	-1.15						1
A2	873	11	84	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	874	11	85	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	875	11	86	G	G	Gap	43%	25%	-1.15						1
A2	876	11	87	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		2
A2	877	11	88	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	878	11	89	G	X	Gap	43%	25%	-1.15						1
A2	879	11	90	E	X	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	3
A2	880	11	91	R	X	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	881	11	92	G	G	Gap	43%	25%	-1.15						1
A2	882	11	93	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	1
A2	883	11	94	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	884	11	95	G	G	Gap	43%	25%	-1.15						1
A2	885	11	96	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A2	886	11	97	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	887	11	98	G	G	Gap	43%	25%	-1.15						1
A2	888	11	99	S	S	Gap	43%	262%	-0.46	4.2262	4.8442		-3.7662		4
A2	889	11	100	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	4
A2	890	11	101	G	G	Gap	43%	25%	-1.15						1
A2	891	11	102	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	5
A2	892	11	103	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	6
A2	893	11	104	G	G	Gap	43%	25%	-1.15						1
A2	894	11	105	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	895	11	106	L	L	Gap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A2	896	11	107	G	G	Gap	43%	25%	-1.15						1
A2	897	11	108	I	I	Gap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	1
A2	898	11	109	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	899	11	110	G	G	Gap	43%	25%	-1.15						1
A2	900	11	111	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	901	11	112	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	902	11	113	G	G	Gap	43%	25%	-1.15						1
A2	903	11	114	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	904	11	115	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	4
A2	905	11	116	G	G	Gap	43%	25%	-1.15						1
A2	906	11	117	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	907	11	118	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	908	11	119	G	G	Gap	43%	25%	-1.15						1
A2	909	11	120	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	910	11	121	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A2	911	11	122	G	G	Gap	43%	25%	-1.15						1
A2	912	11	123	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	4
A2	913	11	124	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	914	11	125	G	G	Gap	43%	25%	-1.15						1
A2	915	11	126	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A2	916	11	127	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	917	11	128	G	G	Gap	43%	25%	-1.15						1

A2	918	11	129	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	919	11	130	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	920	11	131	G	G	Gap	43%	25%	-1.15						1
A2	921	11	132	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A2	922	11	133	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	923	11	134	G	G	Gap	43%	25%	-1.15						1
A2	924	11	135	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	925	11	136	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	2
A2	926	11	137	G	G	Gap	43%	25%	-1.15						1
A2	927	11	138	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	928	11	139	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	929	11	140	G	G	Gap	43%	25%	-1.15						1
A2	930	11	141	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	1
A2	931	11	142	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	1
A2	932	11	143	G	G	Gap	43%	25%	-1.15						1
A2	933	11	144	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	934	11	145	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	935	11	146	G	G	Gap	43%	25%	-1.15						1
A2	936	11	147	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	937	11	148	D	D	Gap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	1
A2	938	11	149	G	G	Gap	43%	25%	-1.15						1
A2	939	11	150	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	940	11	151	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	941	11	152	G	G	Gap	43%	25%	-1.15						1
A2	942	11	153	H	H	Gap	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	4
A2	943	11	154	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A2	944	11	155	G	G	Gap	43%	25%	-1.15						1
A2	945	11	156	E	E	Gap	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	4
A2	946	11	157	R	R	Gap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	947	11	158	G	G	Gap	43%	25%	-1.15						1
A2	948	11	159	Y	Y	Gap	1%	53%	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	2
A2	949	11	160	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	950	11	161	G	G	Gap	43%	25%	-1.15						1
A2	951	11	162	N	N	Gap	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	3
A2	952	11	163	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	953	11	164	G	G	Gap	43%	25%	-1.15						1
A2	954	11	165	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A2	955	11	166	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	6
A2	956	11	167	G	G	Gap	43%	25%	-1.15						1
A2	957	11	168	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	958	11	169	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A2	959	11	170	G	G	Gap	43%	25%	-1.15						1
A2	960	11	171	A	A	Gap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A2	961	11	172	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A2	962	11	173	G	G	Gap	43%	25%	-1.15						1
A2	963	11	174	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	964	11	175	Q	Q	Gap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	965	11	176	G	G	Gap	43%	25%	-1.15						1
A2	966	11	177	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	967	11	178	V	V	Gap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A2	968	11	179	G	G	Gap	43%	25%	-1.15						1
A2	969	11	180	P	P	Gap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	6
A2	970	11	181	T	T	Gap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	2
A2	971	11	182	G	G	Gap	43%	25%	-1.15						1
A2	972	11	183	K	K	Gap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1

A2	973	11	184	H	H	Gap	25%	-49%	-2.33	5.7063(	7.1763-	80.668	-3.3763	-78.338	5
A2	974	11	185	G	G	Gap	43%	25%	-1.15						1
A2	975	11	186	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		1
A2	976	11	187	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	4
A2	977	11	188	G	G	Gap	43%	25%	-1.15						1
A2	978	11	189	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	4
A2	979	11	190	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	980	11	191	G	G	Gap	43%	25%	-1.15						1
A2	981	11	192	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	982	11	193	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	3
A2	983	11	194	G	G	Gap	43%	25%	-1.15						1
A2	984	11	195	A	A	Gap	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	2
A2	985	11	196	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	3
A2	986	11	197	G	G	Gap	43%	25%	-1.15						1
A2	987	11	198	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	5
A2	988	11	199	A	A	Gap	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	4
A2	989	11	200	G	G	Gap	43%	25%	-1.15						1
A2	990	11	201	A	A	Gap	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	3
A2	991	11	202	V	V	Gap	-100%	-46%	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	4
A2	992	11	203	G	G	Gap	43%	25%	-1.15						1
A2	993	11	204	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A2	994	11	205	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	4
A2	995	11	206	G	G	Gap	43%	25%	-1.15						1
A2	996	11	207	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A2	997	11	208	S	S	Gap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		1
A2	998	11	209	G	G	Gap	43%	25%	-1.15						1
A2	999	11	210	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A2	1000	11	211	Q	Q	Gap	44%	-19%	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	4
A2	1001	11	212	G	G	Gap	43%	25%	-1.15						1
A2	1002	11	213	I	I	Gap	-51%	-7%	1.12	5.8093-	6.2971-	75.102	-6.9293	-76.222	4
A2	1003	11	214	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	3
A2	1004	11	215	G	G	Gap	43%	25%	-1.15						1
A2	1005	11	216	D	D	Gap	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	6
A2	1006	11	217	K	K	Gap	-67%	40%	-2.8	5.5897-	8.1697(	72.811	-2.7897	-70.011	1
A2	1007	11	218	G	G	Gap	43%	25%	-1.15						1
A2	1008	11	219	E	E	Gap	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	2
A2	1009	11	220	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	1010	11	221	G	G	Gap	43%	25%	-1.15						1
A2	1011	11	222	D	D	Gap	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	3
A2	1012	11	223	K	K	Gap	-67%	40%	-2.8	5.5897-	8.1697(	72.811	-2.7897	-70.011	5
A2	1013	11	224	G	G	Gap	43%	25%	-1.15						1
A2	1014	11	225	P	P	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A2	1015	11	226	R	R	Gap	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	1
A2	1016	11	227	G	X	Gap	43%	25%	-1.15						1
A2	1017	11	228	L	X	Gap	-73%	-78%	1.25	5.7075(	5.7170-	72.686	-6.9575	-73.936	4
A2	1018	11	229	P	X	Gap	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A2	1019	11	230	G	X	Gap	43%	25%	-1.15						1
A2	1020	11	231	L	X	Gap	-73%	-78%	1.25	5.7075(	5.7170-	72.686	-6.9575	-73.936	4
A2	1021	11	232	K	X	Gap	-67%	40%	-2.8	5.5897-	8.1697(	72.811	-2.7897	-70.011	4
A2	1022	11	233	G	G	Gap	43%	25%	-1.15						1
A2	1023	11	234	H	H	Gap	25%	-49%	-2.33	5.7063(	7.1763-	80.668	-3.3763	-78.338	4
A2	1024	11	235	N	N	Gap	44%	-100%	-0.85	5.1002-	5.3919-	59.477	-4.2502	-58.627	1
A2	1025	14	2	G	G	Overlap	43%	25%	-1.15						1
A2	1026	14	3	L	L	Overlap	-73%	-78%	1.25	5.7075(	5.7170-	72.686	-6.9575	-73.936	2
A2	1027	14	4	Q	Q	Overlap	44%	-19%	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	6

A2	1028	14	5	G	G	Overlap	43%	25%	-1.15						1
A2	1029	14	6	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A2	1030	14	7	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	1031	14	8	G	G	Overlap	43%	25%	-1.15						1
A2	1032	14	9	L	L	Overlap	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A2	1033	14	10	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	1034	14	11	G	G	Overlap	43%	25%	-1.15						1
A2	1035	14	12	H	H	Overlap	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	5
A2	1036	14	13	H	H	Overlap	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	2
A2	1037	14	14	G	G	Overlap	43%	25%	-1.15						1
A2	1038	14	15	D	D	Overlap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	5
A2	1039	14	16	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	2
A2	1040	14	17	G	G	Overlap	43%	25%	-1.15						1
A2	1041	14	18	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A2	1042	14	19	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	1043	14	20	G	G	Overlap	43%	25%	-1.15						1
A2	1044	14	21	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A2	1045	14	22	V	V	Overlap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A2	1046	14	23	G	G	Overlap	43%	25%	-1.15						1
A2	1047	14	24	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	6
A2	1048	14	25	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	6
A2	1049	14	26	G	G	Overlap	43%	25%	-1.15						1
A2	1050	14	27	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	1051	14	28	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	1052	14	29	G	G	Overlap	43%	25%	-1.15						1
A2	1053	14	30	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	1054	14	31	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A2	1055	14	32	G	G	Overlap	43%	25%	-1.15						1
A2	1056	14	33	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A2	1057	14	34	T	T	Overlap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	5
A2	1058	14	35	G	G	Overlap	43%	25%	-1.15						1
A2	1059	14	36	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A2	1060	14	37	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	1061	14	38	G	G	Overlap	43%	25%	-1.15						1
A2	1062	14	39	K	K	Overlap	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A2	1063	14	40	D	D	Overlap	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	6
A2	1064	14	41	G	G	Overlap	43%	25%	-1.15						1
A2	1065	14	42	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A2	1066	14	43	T	T	Overlap	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	1
A2	1067	14	44	G	G	Overlap	43%	25%	-1.15						1
A2	1068	14	45	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	4
A2	1069	14	46	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	7
A2	1070	14	47	G	G	Overlap	43%	25%	-1.15						1
A2	1071	14	48	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A2	1072	14	49	V	V	Overlap	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A2	1073	14	50	G	G	Overlap	43%	25%	-1.15						1
A2	1074	14	51	P	P	Overlap	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	5
A2	1075	14	52	A	A	Overlap	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A2	1076	14	53	G	G	Overlap	43%	25%	-1.15						1
A2	1077	14	54	I	I	Overlap	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	1
A2	1078	14	55	R	R	Overlap	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A2	1079	14	56	G	G	Overlap	43%	25%	-1.15						1
A2	1080	14	57	S	S	Overlap	43%	262%	-0.46	4.2262	4.8442		-3.7662		4
A2	1081	14	58	Q	Q	Overlap	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A2	1082	14	59	G	G	Overlap	43%	25%	-1.15						1

A2	1083	14	60	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		5
A2	1084	14	61	Q	Q	Overlap	44%	-19%	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	3
A2	1085	14	62	G	G	Overlap	43%	25%	-1.15						1
A2	1086	14	63	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	4
A2	1087	14	64	A	A	Overlap	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	2
A2	1088	14	65	G	G	Overlap	43%	25%	-1.15						1
A2	1089	14	66	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	2
A2	1090	14	67	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	5
A2	1091	14	68	G	G	Overlap	43%	25%	-1.15						1
A2	1092	14	69	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	1
A2	1093	14	70	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	3
A2	1094	14	71	G	G	Overlap	43%	25%	-1.15						1
A2	1095	14	72	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	1
A2	1096	14	73	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	3
A2	1097	14	74	G	G	Overlap	43%	25%	-1.15						1
A2	1098	14	75	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	3
A2	1099	14	76	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	2
A2	1100	14	77	G	G	Overlap	43%	25%	-1.15						1
A2	1101	14	78	P	P	Overlap	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	1
A2	1102	14	79	S	S	Overlap	43%	262%	-0.46	4.22627	4.8442(		-3.7662		2
A2	1103	14	80	G	G	Overlap	43%	25%	-1.15						1
A2	1104	14	81	G	X	C-terminal	43%	25%	-1.15						
A2	1105	14	82	G	G	C-terminal	43%	25%	-1.15						
A2	1106	14	83	Y	Y	C-terminal	1%	53%	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A2	1107	14	84	D	D	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1108	14	85	F	F	C-terminal	-85%	-81%	1.71	5.9637(	8.4757(	83.928	-7.6737	-85.638	
A2	1109	14	86	G	G	C-terminal	43%	25%	-1.15						
A2	1110	14	87	F	F	C-terminal	-85%	-81%	1.71	5.9637(	8.4757(	83.928	-7.6737	-85.638	
A2	1111	14	88	D	D	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1112	14	89	G	G	C-terminal	43%	25%	-1.15						
A2	1113	14	90	D	D	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1114	14	91	F	F	C-terminal	-85%	-81%	1.71	5.9637(	8.4757(	83.928	-7.6737	-85.638	
A2	1115	14	92	Y	Y	C-terminal	1%	53%	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A2	1116	14	93	R	R	C-terminal	50%	-30%	-1.81	6.4003(	9.7848(	57.551	-4.5903	-55.741	
A2	1117	14	94	A	X	C-terminal	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A2	1118	14	95	D	X	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1119	14	96	Q	X	C-terminal	44%	-19%	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	
A2	1120	14	97	P	X	C-terminal	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A2	1121	14	98	R	X	C-terminal	50%	-30%	-1.81	6.4003(	9.7848(	57.551	-4.5903	-55.741	
A2	1122	14	99	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1123	14	100	P	X	C-terminal	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A2	1124	14	101	A	X	C-terminal	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A2	1125	14	102	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1126	14	103	L	X	C-terminal	-73%	-78%	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A2	1127	14	104	R	X	C-terminal	50%	-30%	-1.81	6.4003(	9.7848(	57.551	-4.5903	-55.741	
A2	1128	14	105	P	X	C-terminal	-10%	-7%	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A2	1129	14	106	K	X	C-terminal	-67%	40%	-2.8	5.5897(	8.1697(	72.811	-2.7897	-70.011	
A2	1130	14	107	D	X	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1131	14	108	Y	X	C-terminal	1%	53%	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A2	1132	14	109	E	X	C-terminal	-53%	-14%	-3.63	5.3136(	6.5667(	69.742	-1.6836	-66.112	
A2	1133	14	110	V	X	C-terminal	-100%	-46%	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	
A2	1134	14	111	D	X	C-terminal	54%	-13%	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A2	1135	14	112	A	X	C-terminal	5%	-43%	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A2	1136	14	113	T	X	C-terminal	31%	9%	-0.25	4.9029(	6.2805(	51.458	-4.6529	-51.208	
A2	1137	14	114	L	X	C-terminal	-73%	-78%	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	

A2	1138	14	115	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A2	1139	14	116	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1140	14	117	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170'	72.686	-6.9575	-73.936	
A2	1141	14	118	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A2	1142	14	119	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A2	1143	14	120	Q	X	C-terminal	44%	-19%	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A2	1144	14	121	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A2	1145	14	122	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A2	1146	14	123	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A2	1147	14	124	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170'	72.686	-6.9575	-73.936	
A2	1148	14	125	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170'	72.686	-6.9575	-73.936	
A2	1149	14	126	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A2	1150	14	127	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	
A2	1151	14	128	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A2	1152	14	129	G	X	C-terminal	43%	25%	-1.15						
A2	1153	14	130	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1154	14	131	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A2	1155	14	132	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A2	1156	14	133	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A2	1157	14	134	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	
A2	1158	14	135	A	X	C-terminal	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A2	1159	14	136	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A2	1160	14	137	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A2	1161	14	138	C	X	C-terminal	131%	-95%	0.02	4.54756	6.0951(		-4.5675		
A2	1162	14	139	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A2	1163	14	140	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	
A2	1164	14	141	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170'	72.686	-6.9575	-73.936	
A2	1165	14	142	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A2	1166	14	143	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170'	72.686	-6.9575	-73.936	
A2	1167	14	144	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1168	14	145	H	X	C-terminal	25%	-49%	-2.33	5.70630	7.1763(	80.668	-3.3763	-78.338	
A2	1169	14	146	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	
A2	1170	14	147	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A2	1171	14	148	W	X	C-terminal	-56%	-79%	2.09	6.50540	9.4442(	81.36	-8.5954	-83.45	
A2	1172	14	149	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1173	14	150	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442(		-3.7662		
A2	1174	14	151	G	X	C-terminal	43%	25%	-1.15						
A2	1175	14	152	Y	X	C-terminal	1%	53%	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A2	1176	14	153	Y	X	C-terminal	1%	53%	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A2	1177	14	154	W	X	C-terminal	-56%	-79%	2.09	6.50540	9.4442(	81.36	-8.5954	-83.45	
A2	1178	14	155	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A2	1179	14	156	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	
A2	1180	14	157	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	
A2	1181	14	158	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A2	1182	14	159	Q	X	C-terminal	44%	-19%	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A2	1183	14	160	G	X	C-terminal	43%	25%	-1.15						
A2	1184	14	161	C	X	C-terminal	131%	-95%	0.02	4.54756	6.0951(		-4.5675		
A2	1185	14	162	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A2	1186	14	163	M	X	C-terminal	-100%	-100%	0.67	5.73531	7.9317(	81.657	-6.4053	-82.327	
A2	1187	14	164	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	
A2	1188	14	165	A	X	C-terminal	5%	-43%	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A2	1189	14	166	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A2	1190	14	167	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A2	1191	14	168	V	X	C-terminal	-100%	-46%	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	
A2	1192	14	169	Y	X	C-terminal	1%	53%	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	

A2	1193	14	170	C	X	C-terminal	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	1194	14	171	D	X	C-terminal	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A2	1195	14	172	F	X	C-terminal	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A2	1196	14	173	S	X	C-terminal	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	1197	14	174	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1198	14	175	G	X	C-terminal	43%	25%	-1.15						
A2	1199	14	176	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1200	14	177	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1201	14	178	C	X	C-terminal	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	1202	14	179	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1203	14	180	R	X	C-terminal	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A2	1204	14	181	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1205	14	182	Q	X	C-terminal	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	1206	14	183	P	X	C-terminal	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A2	1207	14	184	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1208	14	185	D	X	C-terminal	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A2	1209	14	186	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1210	14	187	P	X	C-terminal	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A2	1211	14	188	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1212	14	189	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1213	14	190	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1214	14	191	W	X	C-terminal	-56%	-79%	2.09	6.5054	9.4442	81.36	-8.5954	-83.45	
A2	1215	14	192	Y	X	C-terminal	1%	53%	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A2	1216	14	193	R	X	C-terminal	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A2	1217	14	194	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1218	14	195	S	X	C-terminal	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	1219	14	196	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1220	14	197	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1221	14	198	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1222	14	199	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1223	14	200	H	X	C-terminal	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A2	1224	14	201	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1225	14	202	W	X	C-terminal	-56%	-79%	2.09	6.5054	9.4442	81.36	-8.5954	-83.45	
A2	1226	14	203	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1227	14	204	G	X	C-terminal	43%	25%	-1.15						
A2	1228	14	205	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1229	14	206	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1230	14	207	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1231	14	208	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1232	14	209	G	X	C-terminal	43%	25%	-1.15						
A2	1233	14	210	G	X	C-terminal	43%	25%	-1.15						
A2	1234	14	211	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1235	14	212	Q	X	C-terminal	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	1236	14	213	F	X	C-terminal	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A2	1237	14	214	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1238	14	215	Y	X	C-terminal	1%	53%	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A2	1239	14	216	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1240	14	217	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1241	14	218	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1242	14	219	G	X	C-terminal	43%	25%	-1.15						
A2	1243	14	220	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1244	14	221	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1245	14	222	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1246	14	223	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1247	14	224	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	

A2	1248	14	225	M	X	C-terminal	-100%	-100%	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	
A2	1249	14	226	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1250	14	227	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1251	14	228	Q	X	C-terminal	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	1252	14	229	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1253	14	230	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1254	14	231	F	X	C-terminal	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A2	1255	14	232	M	X	C-terminal	-100%	-100%	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	
A2	1256	14	233	R	X	C-terminal	50%	-30%	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A2	1257	14	234	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1258	14	235	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1259	14	236	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1260	14	237	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1261	14	238	H	X	C-terminal	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A2	1262	14	239	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1263	14	240	S	X	C-terminal	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	1264	14	241	Q	X	C-terminal	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	1265	14	242	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1266	14	243	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1267	14	244	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1268	14	245	Y	X	C-terminal	1%	53%	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A2	1269	14	246	H	X	C-terminal	25%	-49%	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A2	1270	14	247	C	X	C-terminal	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	1271	14	248	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1272	14	249	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1273	14	250	S	X	C-terminal	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	1274	14	251	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1275	14	252	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1276	14	253	Y	X	C-terminal	1%	53%	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A2	1277	14	254	M	X	C-terminal	-100%	-100%	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	
A2	1278	14	255	D	X	C-terminal	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A2	1279	14	256	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1280	14	257	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1281	14	258	T	X	C-terminal	31%	9%	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A2	1282	14	259	G	X	C-terminal	43%	25%	-1.15						
A2	1283	14	260	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1284	14	261	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1285	14	262	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1286	14	263	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A2	1287	14	264	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1288	14	265	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1289	14	266	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1290	14	267	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1291	14	268	Q	X	C-terminal	44%	-19%	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A2	1292	14	269	G	X	C-terminal	43%	25%	-1.15						
A2	1293	14	270	S	X	C-terminal	43%	262%	-0.46	4.2262	4.8442		-3.7662		
A2	1294	14	271	N	X	C-terminal	44%	-100%	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A2	1295	14	272	D	X	C-terminal	54%	-13%	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A2	1296	14	273	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1297	14	274	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1298	14	275	L	X	C-terminal	-73%	-78%	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A2	1299	14	276	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1300	14	277	A	X	C-terminal	5%	-43%	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A2	1301	14	278	E	X	C-terminal	-53%	-14%	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A2	1302	14	279	G	X	C-terminal	43%	25%	-1.15						

A2	1303	14	280	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627
A2	1304	14	281	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442		-3.7662	
A2	1305	14	282	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741
A2	1306	14	283	F	X	C-terminal	-85%	-81%	1.71	5.96373	8.4757	83.928	-7.6737	-85.638
A2	1307	14	284	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208
A2	1308	14	285	Y	X	C-terminal	1%	53%	0.71	6.18798	8.0649	91.178	-6.8979	-91.888
A2	1309	14	286	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208
A2	1310	14	287	V	X	C-terminal	-100%	-46%	0.46	5.29870	6.9578	60.07	-5.7587	-60.53
A2	1311	14	288	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936
A2	1312	14	289	V	X	C-terminal	-100%	-46%	0.46	5.29870	6.9578	60.07	-5.7587	-60.53
A2	1313	14	290	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557
A2	1314	14	291	G	X	C-terminal	43%	25%	-1.15					
A2	1315	14	292	C	X	C-terminal	131%	-95%	0.02	4.54756	6.0951		-4.5675	
A2	1316	14	293	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442		-3.7662	
A2	1317	14	294	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1318	14	295	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1319	14	296	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208
A2	1320	14	297	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627
A2	1321	14	298	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112
A2	1322	14	299	W	X	C-terminal	-56%	-79%	2.09	6.50540	9.4442	81.36	-8.5954	-83.45
A2	1323	14	300	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1324	14	301	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1325	14	302	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208
A2	1326	14	303	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1327	14	304	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1328	14	305	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112
A2	1329	14	306	Y	X	C-terminal	1%	53%	0.71	6.18798	8.0649	91.178	-6.8979	-91.888
A2	1330	14	307	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1331	14	308	T	X	C-terminal	31%	9%	-0.25	4.90291	6.2805	51.458	-4.6529	-51.208
A2	1332	14	309	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627
A2	1333	14	310	K	X	C-terminal	-67%	40%	-2.8	5.58972	8.1697	72.811	-2.7897	-70.011
A2	1334	14	311	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564
A2	1335	14	312	S	X	C-terminal	43%	262%	-0.46	4.22627	4.8442		-3.7662	
A2	1336	14	313	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741
A2	1337	14	314	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936
A2	1338	14	315	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564
A2	1339	14	316	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1340	14	317	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936
A2	1341	14	318	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557
A2	1342	14	319	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1343	14	320	A	X	C-terminal	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974
A2	1344	14	321	P	X	C-terminal	-10%	-7%	-0.14	5.57682	5.6186	80.704	-5.4368	-80.564
A2	1345	14	322	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936
A2	1346	14	323	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557
A2	1347	14	324	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1348	14	325	G	X	C-terminal	43%	25%	-1.15					
A2	1349	14	326	G	X	C-terminal	43%	25%	-1.15					
A2	1350	14	327	A	X	C-terminal	5%	-43%	-0.5	3.92448	3.9245	31.474	-3.4244	-30.974
A2	1351	14	328	D	X	C-terminal	54%	-13%	-3.64	4.74709	5.6096	55.197	-1.1070	-51.557
A2	1352	14	329	Q	X	C-terminal	44%	-19%	-0.77	5.51489	7.3776	74.788	-4.7448	-74.018
A2	1353	14	330	E	X	C-terminal	-53%	-14%	-3.63	5.31362	6.5667	69.742	-1.6836	-66.112
A2	1354	14	331	I	X	C-terminal	-51%	-7%	1.12	5.80934	6.2971	75.102	-6.9293	-76.222
A2	1355	14	332	R	X	C-terminal	50%	-30%	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741
A2	1356	14	333	L	X	C-terminal	-73%	-78%	1.25	5.70759	5.7170	72.686	-6.9575	-73.936
A2	1357	14	334	N	X	C-terminal	44%	-100%	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627

A2	1358	14	335	I	X	C-terminal	-51%	-7%	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A2	1359	14	336	G	X	C-terminal	43%	25%	-1.15						
A2	1360	14	337	P	X	C-terminal	-10%	-7%	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A2	1361	14	338	V	X	C-terminal	-100%	-46%	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A2	1362	14	339	C	X	C-terminal	131%	-95%	0.02	4.5475	6.0951		-4.5675		
A2	1363	14	340	F	X	C-terminal	-85%	-81%	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A2	1364	14	341	K	X	C-terminal	-67%	40%	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1	-1	-1	M	X	Telopeptide	-1	-1	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	
A1	2	-1	-1	F	X	Telopeptide	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	3	-1	-1	S	X	Telopeptide	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	4	-1	-1	F	X	Telopeptide	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	5	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	6	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	7	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	8	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	9	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	10	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	11	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	12	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	13	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	14	-1	-1	A	X	Telopeptide	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	15	-1	-1	A	X	Telopeptide	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	16	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	17	-1	-1	A	X	Telopeptide	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	18	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	19	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	20	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	21	-1	-1	H	X	Telopeptide	0.25	-0.49	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A1	22	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15						
A1	23	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	24	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	25	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	26	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15						
A1	27	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	28	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	29	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	30	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15						
A1	31	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	32	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	33	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	34	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	35	-1	-1	I	X	Telopeptide	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	36	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	37	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	38	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	39	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	40	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.5475	6.0951		-4.5675		
A1	41	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	42	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	43	-1	-1	N	X	Telopeptide	0.44	-1	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A1	44	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15						
A1	45	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	46	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	47	-1	-1	Y	X	Telopeptide	0.01	0.53	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A1	48	-1	-1	H	X	Telopeptide	0.25	-0.49	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	

A1	49	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	50	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	51	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	52	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	53	-1	-1	W	X	Telopeptide	-0.56	-0.79	2.09	6.50540	9.44420	81.36	-8.5954	-83.45
A1	54	-1	-1	K	X	Telopeptide	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011
A1	55	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	56	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	57	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	58	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	59	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018
A1	60	-1	-1	I	X	Telopeptide	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222
A1	61	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	62	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	63	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	64	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	65	-1	-1	N	X	Telopeptide	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627
A1	66	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	67	-1	-1	N	X	Telopeptide	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627
A1	68	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	69	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936
A1	70	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	71	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	72	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	73	-1	-1	V	X	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	74	-1	-1	I	X	Telopeptide	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222
A1	75	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	76	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	77	-1	-1	E	X	Telopeptide	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112
A1	78	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936
A1	79	-1	-1	K	X	Telopeptide	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011
A1	80	-1	-1	D	X	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	81	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	82	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	83	-1	-1	N	X	Telopeptide	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627
A1	84	-1	-1	A	X	Telopeptide	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974
A1	85	-1	-1	K	X	Telopeptide	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011
A1	86	-1	-1	V	V	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	87	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	88	-1	-1	T	T	Telopeptide	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208
A1	89	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	90	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	91	-1	-1	R	R	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	92	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	93	-1	-1	V	V	Telopeptide	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53
A1	94	-1	-1	S	S	Telopeptide	0.43	2.62	-0.46	4.22627	4.84426		-3.7662	
A1	95	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	96	-1	-1	H	X	Telopeptide	0.25	-0.49	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338
A1	97	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	98	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	99	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	100	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	101	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	102	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	103	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564

A1	104	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	105	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208
A1	106	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208
A1	107	-1	-1	T	X	Telopeptide	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208
A1	108	-1	-1	K	X	Telopeptide	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011
A1	109	-1	-1	Q	X	Telopeptide	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018
A1	110	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	111	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	112	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	113	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	114	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	115	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	116	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	117	-1	-1	C	X	Telopeptide	1.31	-0.95	0.02	4.54756	6.09519		-4.5675	
A1	118	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	119	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	120	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	121	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	122	-1	-1	A	X	Telopeptide	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974
A1	123	-1	-1	H	X	Telopeptide	0.25	-0.49	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338
A1	124	-1	-1	R	X	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	125	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	126	-1	-1	S	X	Telopeptide	0.43	2.62	-0.46	4.22627	4.84426		-3.7662	
A1	127	-1	-1	H	X	Telopeptide	0.25	-0.49	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338
A1	128	-1	-1	S	S	Telopeptide	0.43	2.62	-0.46	4.22627	4.84426		-3.7662	
A1	129	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	130	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	131	-1	-1	Q	Q	Telopeptide	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018
A1	132	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	133	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	134	-1	-1	A	A	Telopeptide	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974
A1	135	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	136	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	137	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	138	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	139	-1	-1	R	R	Telopeptide	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741
A1	140	-1	-1	D	D	Telopeptide	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557
A1	141	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	142	-1	-1	I	I	Telopeptide	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222
A1	143	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	144	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	145	-1	-1	Q	Q	Telopeptide	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018
A1	146	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	147	-1	-1	G	X	Telopeptide	0.43	0.25	-1.15					
A1	148	-1	-1	L	X	Telopeptide	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936
A1	149	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	150	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	151	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	152	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	153	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	154	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	155	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	156	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15					
A1	157	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564
A1	158	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564

A1	159	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15						
A1	160	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	161	-1	-1	P	P	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	162	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15						
A1	163	-1	-1	L	L	Telopeptide	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	
A1	164	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15						
A1	165	-1	-1	G	G	Telopeptide	0.43	0.25	-1.15						
A1	166	-1	-1	N	N	Telopeptide	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	
A1	167	-1	-1	F	F	Telopeptide	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	
A1	168	-1	-1	A	A	Telopeptide	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	
A1	169	-1	-1	P	X	Telopeptide	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	170	-1	-1	Q	Q	N-term	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	171	-1	-1	L	L	N-term	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	3
A1	172	-1	-1	S	S	N-term	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	173	-1	-1	Y	Y	N-term	0.01	0.53	0.71	6.18798	8.06490	91.178	-6.8979	-91.888	1
A1	174	-1	-1	G	G	N-term	0.43	0.25	-1.15						1
A1	175	-1	-1	Y	Y	N-term	0.01	0.53	0.71	6.18798	8.06490	91.178	-6.8979	-91.888	1
A1	176	-1	-1	D	D	N-term	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	177	-1	-1	E	E	N-term	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	178	-1	-1	K	K	N-term	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	179	-1	-1	S	S	N-term	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	180	-1	-1	T	T	N-term	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	6
A1	181	-1	-1	G	G	N-term	0.43	0.25	-1.15						4
A1	182	-1	-1	I	I	N-term	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	6
A1	183	-1	-1	S	S	N-term	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	184	-1	-1	V	V	N-term	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	2
A1	185	-1	-1	P	P	N-term	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	186	0	0	G	G	Overlap	0.43	0.25	-1.15						1
A1	187	0	1	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	188	0	2	M	M	Overlap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	1
A1	189	0	3	G	G	Overlap	0.43	0.25	-1.15						1
A1	190	0	4	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	191	0	5	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	192	0	6	G	G	Overlap	0.43	0.25	-1.15						1
A1	193	0	7	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	194	0	8	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	195	0	9	G	G	Overlap	0.43	0.25	-1.15						1
A1	196	0	10	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	197	0	11	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	198	0	12	G	G	Overlap	0.43	0.25	-1.15						1
A1	199	0	13	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	200	0	14	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	201	0	15	G	G	Overlap	0.43	0.25	-1.15						1
A1	202	0	16	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	203	0	17	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	204	0	18	G	G	Overlap	0.43	0.25	-1.15						1
A1	205	0	19	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	206	0	20	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	207	0	21	G	G	Overlap	0.43	0.25	-1.15						1
A1	208	0	22	F	F	Overlap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	209	0	23	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	210	0	24	G	G	Overlap	0.43	0.25	-1.15						1
A1	211	0	25	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	212	0	26	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	213	0	27	G	G	Overlap	0.43	0.25	-1.15						1

A1	214	0	28	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	215	0	29	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	216	0	30	G	G	Overlap	0.43	0.25	-1.15						1
A1	217	0	31	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	218	0	32	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	219	0	33	G	G	Overlap	0.43	0.25	-1.15						1
A1	220	0	34	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	221	0	35	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		1
A1	222	0	36	G	G	Overlap	0.43	0.25	-1.15						1
A1	223	0	37	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	224	0	38	M	M	Overlap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	1
A1	225	0	39	G	G	Overlap	0.43	0.25	-1.15						1
A1	226	0	40	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	227	0	41	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	228	0	42	G	G	Overlap	0.43	0.25	-1.15						1
A1	229	0	43	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	230	0	44	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	231	0	45	G	G	Overlap	0.43	0.25	-1.15						1
A1	232	0	46	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	233	0	47	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	234	0	48	G	G	Overlap	0.43	0.25	-1.15						1
A1	235	0	49	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	236	0	50	N	N	Overlap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	237	0	51	G	G	Overlap	0.43	0.25	-1.15						1
A1	238	0	52	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	239	0	53	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	240	0	54	G	G	Overlap	0.43	0.25	-1.15						1
A1	241	0	55	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	242	0	56	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	243	0	57	G	G	Overlap	0.43	0.25	-1.15						1
A1	244	0	58	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	245	0	59	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	246	0	60	G	G	Overlap	0.43	0.25	-1.15						1
A1	247	0	61	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	248	0	62	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	249	0	63	G	G	Overlap	0.43	0.25	-1.15						1
A1	250	0	64	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	2
A1	251	0	65	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	252	0	66	G	G	Overlap	0.43	0.25	-1.15						1
A1	253	0	67	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	254	0	68	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	255	0	69	G	G	Overlap	0.43	0.25	-1.15						1
A1	256	0	70	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	257	0	71	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	1
A1	258	0	72	G	G	Overlap	0.43	0.25	-1.15						1
A1	259	0	73	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	260	0	74	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	261	0	75	G	G	Overlap	0.43	0.25	-1.15						2
A1	262	0	76	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	263	0	77	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	264	0	78	G	G	Overlap	0.43	0.25	-1.15						1
A1	265	0	79	T	T	Overlap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A1	266	0	80	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	267	0	81	G	G	Gap	0.43	0.25	-1.15						1
A1	268	0	82	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2

A1	269	0	83	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	270	0	84	G	G	Gap	0.43	0.25	-1.15						1
A1	271	0	85	M	M	Gap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	2
A1	272	0	86	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	273	0	87	G	G	Gap	0.43	0.25	-1.15						1
A1	274	0	88	H	H	Gap	0.25	-0.49	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338	1
A1	275	0	89	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	276	0	90	G	G	Gap	0.43	0.25	-1.15						1
A1	277	0	91	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	278	0	92	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		1
A1	279	0	93	G	G	Gap	0.43	0.25	-1.15						1
A1	280	0	94	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	281	0	95	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	282	0	96	G	G	Gap	0.43	0.25	-1.15						1
A1	283	0	97	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	284	0	98	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	285	0	99	G	G	Gap	0.43	0.25	-1.15						1
A1	286	0	100	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	287	0	101	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	288	0	102	G	G	Gap	0.43	0.25	-1.15						1
A1	289	0	103	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	290	0	104	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	291	0	105	G	G	Gap	0.43	0.25	-1.15						1
A1	292	0	106	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	293	0	107	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	294	0	108	G	G	Gap	0.43	0.25	-1.15						1
A1	295	0	109	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	296	0	110	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	297	0	111	G	G	Gap	0.43	0.25	-1.15						1
A1	298	0	112	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		1
A1	299	0	113	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	300	0	114	G	G	Gap	0.43	0.25	-1.15						1
A1	301	0	115	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	302	0	116	N	N	Gap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	303	0	117	G	G	Gap	0.43	0.25	-1.15						1
A1	304	0	118	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	305	0	119	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	306	0	120	G	G	Gap	0.43	0.25	-1.15						1
A1	307	0	121	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	308	0	122	M	M	Gap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	1
A1	309	0	123	G	G	Gap	0.43	0.25	-1.15						1
A1	310	0	124	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	311	0	125	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	312	0	126	G	G	Gap	0.43	0.25	-1.15						1
A1	313	0	127	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	314	0	128	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	315	0	129	G	G	Gap	0.43	0.25	-1.15						1
A1	316	0	130	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	317	0	131	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	318	0	132	G	G	Gap	0.43	0.25	-1.15						1
A1	319	0	133	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	320	0	134	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	321	0	135	G	G	Gap	0.43	0.25	-1.15						1
A1	322	0	136	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	323	0	137	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4

A1	324	0	138	G	G	Gap	0.43	0.25	-1.15						1
A1	325	0	139	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	326	0	140	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	327	0	141	G	G	Gap	0.43	0.25	-1.15						1
A1	328	0	142	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	329	0	143	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	330	0	144	G	G	Gap	0.43	0.25	-1.15						1
A1	331	0	145	N	N	Gap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	332	0	146	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	333	0	147	G	G	Gap	0.43	0.25	-1.15						1
A1	334	0	148	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	335	0	149	T	T	Gap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	4
A1	336	0	150	G	G	Gap	0.43	0.25	-1.15						1
A1	337	0	151	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	338	0	152	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	339	0	153	G	G	Gap	0.43	0.25	-1.15						1
A1	340	0	154	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	341	0	155	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	342	0	156	G	G	Gap	0.43	0.25	-1.15						1
A1	343	0	157	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	344	0	158	T	T	Gap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	2
A1	345	0	159	G	G	Gap	0.43	0.25	-1.15						1
A1	346	0	160	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	347	0	161	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	348	0	162	G	G	Gap	0.43	0.25	-1.15						1
A1	349	0	163	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	350	0	164	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	351	0	165	G	G	Gap	0.43	0.25	-1.15						1
A1	352	0	166	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	353	0	167	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	354	0	168	G	G	Gap	0.43	0.25	-1.15						1
A1	355	0	169	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	356	0	170	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	3
A1	357	0	171	G	G	Gap	0.43	0.25	-1.15						1
A1	358	0	172	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	359	0	173	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	360	0	174	G	G	Gap	0.43	0.25	-1.15						1
A1	361	0	175	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	362	0	176	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	363	0	177	G	G	Gap	0.43	0.25	-1.15						1
A1	364	0	178	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	365	0	179	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	366	0	180	G	G	Gap	0.43	0.25	-1.15						1
A1	367	0	181	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	368	0	182	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	369	0	183	G	G	Gap	0.43	0.25	-1.15						1
A1	370	0	184	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		1
A1	371	0	185	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	372	0	186	G	G	Gap	0.43	0.25	-1.15						1
A1	373	0	187	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	374	0	188	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	375	0	189	G	G	Gap	0.43	0.25	-1.15						1
A1	376	0	190	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	3
A1	377	0	191	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	378	0	192	G	G	Gap	0.43	0.25	-1.15						1

A1	379	0	193	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	380	0	194	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	381	0	195	G	G	Gap	0.43	0.25	-1.15						1
A1	382	0	196	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	383	0	197	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	384	0	198	G	G	Gap	0.43	0.25	-1.15						1
A1	385	0	199	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	386	0	200	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	387	0	201	G	G	Gap	0.43	0.25	-1.15						1
A1	388	0	202	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	389	0	203	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	390	0	204	G	G	Gap	0.43	0.25	-1.15						1
A1	391	0	205	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	392	0	206	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	393	0	207	G	G	Gap	0.43	0.25	-1.15						1
A1	394	0	208	N	N	Gap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	395	0	209	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	396	0	210	G	G	Gap	0.43	0.25	-1.15						1
A1	397	0	211	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	398	0	212	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	399	0	213	G	G	Gap	0.43	0.25	-1.15						1
A1	400	0	214	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	1
A1	401	0	215	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	402	0	216	G	G	Gap	0.43	0.25	-1.15						1
A1	403	0	217	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	404	0	218	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	405	0	219	G	G	Gap	0.43	0.25	-1.15						1
A1	406	0	220	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	407	0	221	N	N	Gap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	2
A1	408	0	222	G	G	Gap	0.43	0.25	-1.15						1
A1	409	0	223	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	410	0	224	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	411	0	225	G	G	Gap	0.43	0.25	-1.15						1
A1	412	0	226	I	I	Gap	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	1
A1	413	0	227	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	414	0	228	G	G	Gap	0.43	0.25	-1.15						1
A1	415	0	229	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	416	0	230	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	417	0	231	G	G	Gap	0.43	0.25	-1.15						1
A1	418	0	232	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	419	0	233	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	420	3	0	G	G	Overlap	0.43	0.25	-1.15						1
A1	421	3	1	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	422	3	2	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	423	3	3	G	G	Overlap	0.43	0.25	-1.15						1
A1	424	3	4	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	425	3	5	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		2
A1	426	3	6	G	G	Overlap	0.43	0.25	-1.15						1
A1	427	3	7	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	428	3	8	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	1
A1	429	3	9	G	G	Overlap	0.43	0.25	-1.15						1
A1	430	3	10	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	431	3	11	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		2
A1	432	3	12	G	G	Overlap	0.43	0.25	-1.15						1
A1	433	3	13	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4

A1	434	3	14	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	435	3	15	G	G	Overlap	0.43	0.25	-1.15						1
A1	436	3	16	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	437	3	17	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	438	3	18	G	G	Overlap	0.43	0.25	-1.15						1
A1	439	3	19	N	N	Overlap	0.44	-1	-0.85	5.10027	5.3919	59.477	-4.2502	-58.627	1
A1	440	3	20	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	441	3	21	G	G	Overlap	0.43	0.25	-1.15						1
A1	442	3	22	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	443	3	23	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	444	3	24	G	G	Overlap	0.43	0.25	-1.15						1
A1	445	3	25	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	446	3	26	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	447	3	27	G	G	Overlap	0.43	0.25	-1.15						1
A1	448	3	28	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	449	3	29	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	450	3	30	G	G	Overlap	0.43	0.25	-1.15						1
A1	451	3	31	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	452	3	32	T	T	Overlap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	3
A1	453	3	33	G	G	Overlap	0.43	0.25	-1.15						1
A1	454	3	34	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	455	3	35	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	456	3	36	G	G	Overlap	0.43	0.25	-1.15						1
A1	457	3	37	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	2
A1	458	3	38	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	459	3	39	G	G	Overlap	0.43	0.25	-1.15						1
A1	460	3	40	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	461	3	41	T	T	Overlap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	5
A1	462	3	42	G	G	Overlap	0.43	0.25	-1.15						1
A1	463	3	43	I	I	Overlap	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	3
A1	464	3	44	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	1
A1	465	3	45	G	G	Overlap	0.43	0.25	-1.15						1
A1	466	3	46	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	467	3	47	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	468	3	48	G	G	Overlap	0.43	0.25	-1.15						1
A1	469	3	49	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	470	3	50	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	471	3	51	G	G	Overlap	0.43	0.25	-1.15						1
A1	472	3	52	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	473	3	53	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	474	3	54	G	G	Overlap	0.43	0.25	-1.15						1
A1	475	3	55	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	1
A1	476	3	56	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	1
A1	477	3	57	G	G	Overlap	0.43	0.25	-1.15						1
A1	478	3	58	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	479	3	59	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.7848	57.551	-4.5903	-55.741	1
A1	480	3	60	G	G	Overlap	0.43	0.25	-1.15						1
A1	481	3	61	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	482	3	62	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	483	3	63	G	G	Overlap	0.43	0.25	-1.15						1
A1	484	3	64	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	485	3	65	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	486	3	66	G	G	Overlap	0.43	0.25	-1.15						1
A1	487	3	67	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.7170	72.686	-6.9575	-73.936	2
A1	488	3	68	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1

A1	489	3	69	G	G	Overlap	0.43	0.25	-1.15						1
A1	490	3	70	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	491	3	71	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	492	3	72	G	G	Overlap	0.43	0.25	-1.15						1
A1	493	3	73	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	494	3	74	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	495	3	75	G	G	Overlap	0.43	0.25	-1.15						1
A1	496	3	76	G	G	Overlap	0.43	0.25	-1.15						2
A1	497	3	77	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	498	3	78	G	G	Overlap	0.43	0.25	-1.15						1
A1	499	3	79	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		3
A1	500	3	80	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	501	3	81	G	G	Gap	0.43	0.25	-1.15						1
A1	502	3	82	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	503	3	83	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	504	3	84	G	G	Gap	0.43	0.25	-1.15						1
A1	505	3	85	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	506	3	86	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	507	3	87	G	G	Gap	0.43	0.25	-1.15						1
A1	508	3	88	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	4
A1	509	3	89	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	510	3	90	G	G	Gap	0.43	0.25	-1.15						1
A1	511	3	91	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	512	3	92	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	513	3	93	G	G	Gap	0.43	0.25	-1.15						1
A1	514	3	94	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	515	3	95	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A1	516	3	96	G	G	Gap	0.43	0.25	-1.15						1
A1	517	3	97	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	518	3	98	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	519	3	99	G	G	Gap	0.43	0.25	-1.15						1
A1	520	3	100	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	521	3	101	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	522	3	102	G	G	Gap	0.43	0.25	-1.15						1
A1	523	3	103	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	524	3	104	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	525	3	105	G	G	Gap	0.43	0.25	-1.15						1
A1	526	3	106	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	527	3	107	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	528	3	108	G	G	Gap	0.43	0.25	-1.15						1
A1	529	3	109	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		3
A1	530	3	110	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	531	3	111	G	G	Gap	0.43	0.25	-1.15						1
A1	532	3	112	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	533	3	113	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	534	3	114	G	G	Gap	0.43	0.25	-1.15						1
A1	535	3	115	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	536	3	116	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	537	3	117	G	G	Gap	0.43	0.25	-1.15						1
A1	538	3	118	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	1
A1	539	3	119	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	540	3	120	G	G	Gap	0.43	0.25	-1.15						1
A1	541	3	121	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71707	72.686	-6.9575	-73.936	2
A1	542	3	122	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	543	3	123	G	G	Gap	0.43	0.25	-1.15						1

A1	544	3	124	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A1	545	3	125	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A1	546	3	126	G	G	Gap	0.43	0.25	-1.15						1
A1	547	3	127	L	L	Gap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A1	548	3	128	T	T	Gap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	2
A1	549	3	129	G	G	Gap	0.43	0.25	-1.15						1
A1	550	3	130	S	S	Gap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		1
A1	551	3	131	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	552	3	132	G	G	Gap	0.43	0.25	-1.15						1
A1	553	3	133	S	S	Gap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		2
A1	554	3	134	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	555	3	135	G	G	Gap	0.43	0.25	-1.15						1
A1	556	3	136	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	557	3	137	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	1
A1	558	3	138	G	G	Gap	0.43	0.25	-1.15						1
A1	559	3	139	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A1	560	3	140	T	T	Gap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	3
A1	561	3	141	G	G	Gap	0.43	0.25	-1.15						1
A1	562	3	142	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	563	3	143	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A1	564	3	144	G	G	Gap	0.43	0.25	-1.15						1
A1	565	3	145	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	566	3	146	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	1
A1	567	3	147	G	G	Gap	0.43	0.25	-1.15						1
A1	568	3	148	Q	Q	Gap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A1	569	3	149	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	2
A1	570	3	150	G	G	Gap	0.43	0.25	-1.15						1
A1	571	3	151	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A1	572	3	152	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	573	3	153	G	G	Gap	0.43	0.25	-1.15						1
A1	574	3	154	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	575	3	155	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	576	3	156	G	G	Gap	0.43	0.25	-1.15						1
A1	577	3	157	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	578	3	158	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	579	3	159	G	G	Gap	0.43	0.25	-1.15						1
A1	580	3	160	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	581	3	161	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	1
A1	582	3	162	G	G	Gap	0.43	0.25	-1.15						1
A1	583	3	163	Q	Q	Gap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	1
A1	584	3	164	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	585	3	165	G	G	Gap	0.43	0.25	-1.15						1
A1	586	3	166	V	V	Gap	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	1
A1	587	3	167	M	M	Gap	-1	-1	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	1
A1	588	3	168	G	G	Gap	0.43	0.25	-1.15						1
A1	589	3	169	F	F	Gap	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	1
A1	590	3	170	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	591	3	171	G	G	Gap	0.43	0.25	-1.15						1
A1	592	3	172	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	593	3	173	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A1	594	3	174	G	G	Gap	0.43	0.25	-1.15						1
A1	595	3	175	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	596	3	176	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	597	3	177	G	G	Gap	0.43	0.25	-1.15						1
A1	598	3	178	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2

A1	599	3	179	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	600	3	180	G	G	Gap	0.43	0.25	-1.15						1
A1	601	3	181	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	2
A1	602	3	182	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A1	603	3	183	G	G	Gap	0.43	0.25	-1.15						2
A1	604	3	184	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A1	605	3	185	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	606	3	186	G	G	Gap	0.43	0.25	-1.15						1
A1	607	3	187	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	6
A1	608	3	188	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	609	3	189	G	G	Gap	0.43	0.25	-1.15						3
A1	610	3	190	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	611	3	191	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	612	3	192	G	G	Gap	0.43	0.25	-1.15						1
A1	613	3	193	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	614	3	194	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	5
A1	615	3	195	G	G	Gap	0.43	0.25	-1.15						1
A1	616	3	196	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	617	3	197	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	618	3	198	G	G	Gap	0.43	0.25	-1.15						1
A1	619	3	199	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	3
A1	620	3	200	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	621	3	201	G	G	Gap	0.43	0.25	-1.15						1
A1	622	3	202	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	623	3	203	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	624	3	204	G	G	Gap	0.43	0.25	-1.15						1
A1	625	3	205	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	626	3	206	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	2
A1	627	3	207	G	G	Gap	0.43	0.25	-1.15						2
A1	628	3	208	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	629	3	209	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	630	3	210	G	G	Gap	0.43	0.25	-1.15						2
A1	631	3	211	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	632	3	212	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	633	3	213	G	G	Gap	0.43	0.25	-1.15						1
A1	634	3	214	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	635	3	215	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	636	3	216	G	G	Gap	0.43	0.25	-1.15						1
A1	637	3	217	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	638	3	218	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	639	3	219	G	G	Gap	0.43	0.25	-1.15						1
A1	640	3	220	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	641	3	221	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	642	3	222	G	G	Gap	0.43	0.25	-1.15						1
A1	643	3	223	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	644	3	224	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	645	3	225	G	G	Gap	0.43	0.25	-1.15						1
A1	646	3	226	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	647	3	227	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	648	3	228	G	G	Gap	0.43	0.25	-1.15						1
A1	649	3	229	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	650	3	230	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	651	3	231	G	G	Gap	0.43	0.25	-1.15						1
A1	652	3	232	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71707	72.686	-6.9575	-73.936	2
A1	653	3	233	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1

A1	654	6	0	G	G	Overlap	0.43	0.25	-1.15						1
A1	655	6	1	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	656	6	2	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	657	6	3	G	G	Overlap	0.43	0.25	-1.15						1
A1	658	6	4	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	659	6	5	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	660	6	6	G	G	Overlap	0.43	0.25	-1.15						1
A1	661	6	7	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	662	6	8	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	663	6	9	G	G	Overlap	0.43	0.25	-1.15						1
A1	664	6	10	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	665	6	11	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	666	6	12	G	G	Overlap	0.43	0.25	-1.15						1
A1	667	6	13	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A1	668	6	14	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	669	6	15	G	G	Overlap	0.43	0.25	-1.15						1
A1	670	6	16	V	V	Overlap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	5
A1	671	6	17	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	672	6	18	G	G	Overlap	0.43	0.25	-1.15						1
A1	673	6	19	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	674	6	20	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	4
A1	675	6	21	G	G	Overlap	0.43	0.25	-1.15						1
A1	676	6	22	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	677	6	23	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	678	6	24	G	G	Overlap	0.43	0.25	-1.15						1
A1	679	6	25	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	680	6	26	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	681	6	27	G	G	Overlap	0.43	0.25	-1.15						1
A1	682	6	28	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	683	6	29	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	684	6	30	G	G	Overlap	0.43	0.25	-1.15						1
A1	685	6	31	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	686	6	32	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	687	6	33	G	G	Overlap	0.43	0.25	-1.15						1
A1	688	6	34	F	F	Overlap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	689	6	35	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	690	6	36	G	G	Overlap	0.43	0.25	-1.15						1
A1	691	6	37	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	692	6	38	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	693	6	39	G	G	Overlap	0.43	0.25	-1.15						1
A1	694	6	40	V	V	Overlap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A1	695	6	41	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	696	6	42	G	G	Overlap	0.43	0.25	-1.15						1
A1	697	6	43	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	698	6	44	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	699	6	45	G	G	Overlap	0.43	0.25	-1.15						1
A1	700	6	46	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	701	6	47	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	702	6	48	G	G	Overlap	0.43	0.25	-1.15						1
A1	703	6	49	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	704	6	50	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	705	6	51	G	G	Overlap	0.43	0.25	-1.15						1
A1	706	6	52	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	707	6	53	N	N	Overlap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	708	6	54	G	G	Overlap	0.43	0.25	-1.15						1

A1	709	6	55	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	710	6	56	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	711	6	57	G	G	Overlap	0.43	0.25	-1.15						1
A1	712	6	58	N	N	Overlap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	1
A1	713	6	59	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	714	6	60	G	G	Overlap	0.43	0.25	-1.15						1
A1	715	6	61	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	716	6	62	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	717	6	63	G	G	Overlap	0.43	0.25	-1.15						1
A1	718	6	64	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	719	6	65	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	720	6	66	G	G	Overlap	0.43	0.25	-1.15						1
A1	721	6	67	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	722	6	68	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	723	6	69	G	G	Overlap	0.43	0.25	-1.15						1
A1	724	6	70	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	725	6	71	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	726	6	72	G	G	Overlap	0.43	0.25	-1.15						1
A1	727	6	73	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		3
A1	728	6	74	Q	Q	Overlap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	729	6	75	G	G	Overlap	0.43	0.25	-1.15						1
A1	730	6	76	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	731	6	77	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	732	6	78	G	G	Overlap	0.43	0.25	-1.15						1
A1	733	6	79	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	734	6	80	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	735	6	81	G	G	Gap	0.43	0.25	-1.15						1
A1	736	6	82	M	M	Gap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	1
A1	737	6	83	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	738	6	84	G	G	Gap	0.43	0.25	-1.15						1
A1	739	6	85	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	740	6	86	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	741	6	87	G	G	Gap	0.43	0.25	-1.15						1
A1	742	6	88	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	743	6	89	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	744	6	90	G	G	Gap	0.43	0.25	-1.15						1
A1	745	6	91	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	2
A1	746	6	92	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	747	6	93	G	G	Gap	0.43	0.25	-1.15						1
A1	748	6	94	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	749	6	95	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	750	6	96	G	G	Gap	0.43	0.25	-1.15						1
A1	751	6	97	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	752	6	98	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	753	6	99	G	G	Gap	0.43	0.25	-1.15						1
A1	754	6	100	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	755	6	101	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A1	756	6	102	G	G	Gap	0.43	0.25	-1.15						1
A1	757	6	103	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	758	6	104	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	759	6	105	G	G	Gap	0.43	0.25	-1.15						1
A1	760	6	106	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	761	6	107	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	762	6	108	G	G	Gap	0.43	0.25	-1.15						1
A1	763	6	109	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4

A1	764	6	110	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	765	6	111	G	G	Gap	0.43	0.25	-1.15						1
A1	766	6	112	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	767	6	113	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	768	6	114	G	G	Gap	0.43	0.25	-1.15						1
A1	769	6	115	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	6
A1	770	6	116	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	771	6	117	G	G	Gap	0.43	0.25	-1.15						1
A1	772	6	118	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	3
A1	773	6	119	T	T	Gap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A1	774	6	120	G	G	Gap	0.43	0.25	-1.15						1
A1	775	6	121	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	776	6	122	I	I	Gap	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	1
A1	777	6	123	G	G	Gap	0.43	0.25	-1.15						1
A1	778	6	124	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	779	6	125	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	780	6	126	G	G	Gap	0.43	0.25	-1.15						1
A1	781	6	127	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	782	6	128	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	783	6	129	G	G	Gap	0.43	0.25	-1.15						1
A1	784	6	130	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	785	6	131	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	786	6	132	G	G	Gap	0.43	0.25	-1.15						1
A1	787	6	133	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	788	6	134	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	789	6	135	G	G	Gap	0.43	0.25	-1.15						1
A1	790	6	136	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	791	6	137	T	T	Gap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	3
A1	792	6	138	G	G	Gap	0.43	0.25	-1.15						1
A1	793	6	139	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	794	6	140	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		3
A1	795	6	141	G	G	Gap	0.43	0.25	-1.15						1
A1	796	6	142	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	797	6	143	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	798	6	144	G	G	Gap	0.43	0.25	-1.15						1
A1	799	6	145	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	800	6	146	T	T	Gap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	1
A1	801	6	147	G	G	Gap	0.43	0.25	-1.15						1
A1	802	6	148	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	803	6	149	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	804	6	150	G	G	Gap	0.43	0.25	-1.15						1
A1	805	6	151	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	806	6	152	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	807	6	153	G	G	Gap	0.43	0.25	-1.15						1
A1	808	6	154	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	809	6	155	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	810	6	156	G	G	Gap	0.43	0.25	-1.15						1
A1	811	6	157	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	812	6	158	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	813	6	159	G	G	Gap	0.43	0.25	-1.15						1
A1	814	6	160	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	815	6	161	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	816	6	162	G	G	Gap	0.43	0.25	-1.15						1
A1	817	6	163	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	818	6	164	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1

A1	819	6	165	G	G	Gap	0.43	0.25	-1.15						1
A1	820	6	166	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.47575	83.928	-7.6737	-85.638	1
A1	821	6	167	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	822	6	168	G	G	Gap	0.43	0.25	-1.15						1
A1	823	6	169	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	824	6	170	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	825	6	171	G	G	Gap	0.43	0.25	-1.15						1
A1	826	6	172	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	827	6	173	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	1
A1	828	6	174	G	G	Gap	0.43	0.25	-1.15						1
A1	829	6	175	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	1
A1	830	6	176	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	831	6	177	G	G	Gap	0.43	0.25	-1.15						1
A1	832	6	178	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	833	6	179	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	834	6	180	G	G	Gap	0.43	0.25	-1.15						1
A1	835	6	181	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	836	6	182	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	837	6	183	G	G	Gap	0.43	0.25	-1.15						1
A1	838	6	184	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	839	6	185	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	840	6	186	G	G	Gap	0.43	0.25	-1.15						1
A1	841	6	187	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	842	6	188	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	843	6	189	G	G	Gap	0.43	0.25	-1.15						1
A1	844	6	190	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	845	6	191	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	846	6	192	G	G	Gap	0.43	0.25	-1.15						1
A1	847	6	193	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	848	6	194	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	849	6	195	G	G	Gap	0.43	0.25	-1.15						1
A1	850	6	196	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	851	6	197	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	1
A1	852	6	198	G	G	Gap	0.43	0.25	-1.15						1
A1	853	6	199	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	854	6	200	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	855	6	201	G	G	Gap	0.43	0.25	-1.15						1
A1	856	6	202	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	857	6	203	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	858	6	204	G	G	Gap	0.43	0.25	-1.15						1
A1	859	6	205	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	860	6	206	I	I	Gap	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	5
A1	861	6	207	G	G	Gap	0.43	0.25	-1.15						1
A1	862	6	208	N	N	Gap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	3
A1	863	6	209	V	V	Gap	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	1
A1	864	6	210	G	G	Gap	0.43	0.25	-1.15						1
A1	865	6	211	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	2
A1	866	6	212	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	867	6	213	G	G	Gap	0.43	0.25	-1.15						1
A1	868	6	214	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	4
A1	869	6	215	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	1
A1	870	6	216	G	G	Gap	0.43	0.25	-1.15						1
A1	871	6	217	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	872	6	218	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	3
A1	873	6	219	G	G	Gap	0.43	0.25	-1.15						1

A1	874	6	220	S	S	Gap	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		4
A1	875	6	221	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	3
A1	876	6	222	G	G	Gap	0.43	0.25	-1.15						1
A1	877	6	223	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A1	878	6	224	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A1	879	6	225	G	G	Gap	0.43	0.25	-1.15						1
A1	880	6	226	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	1
A1	881	6	227	T	T	Gap	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	1
A1	882	6	228	G	G	Gap	0.43	0.25	-1.15						1
A1	883	6	229	F	F	Gap	-0.85	-0.81	1.71	5.96373	8.4757(	83.928	-7.6737	-85.638	1
A1	884	6	230	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A1	885	6	231	G	G	Gap	0.43	0.25	-1.15						1
A1	886	6	232	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	1
A1	887	6	233	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	1
A1	888	9	0	G	G	Overlap	0.43	0.25	-1.15						1
A1	889	9	1	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	2
A1	890	9	2	V	V	Overlap	-1	-0.46	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	2
A1	891	9	3	G	G	Overlap	0.43	0.25	-1.15						2
A1	892	9	4	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A1	893	9	5	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A1	894	9	6	G	G	Overlap	0.43	0.25	-1.15						3
A1	895	9	7	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	3
A1	896	9	8	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		3
A1	897	9	9	G	G	Overlap	0.43	0.25	-1.15						1
A1	898	9	10	N	N	Overlap	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	2
A1	899	9	11	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	6
A1	900	9	12	G	G	Overlap	0.43	0.25	-1.15						1
A1	901	9	13	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	4
A1	902	9	14	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A1	903	9	15	G	G	Overlap	0.43	0.25	-1.15						1
A1	904	9	16	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	1
A1	905	9	17	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A1	906	9	18	G	G	Overlap	0.43	0.25	-1.15						1
A1	907	9	19	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A1	908	9	20	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	5
A1	909	9	21	G	G	Overlap	0.43	0.25	-1.15						1
A1	910	9	22	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	2
A1	911	9	23	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	3
A1	912	9	24	G	G	Overlap	0.43	0.25	-1.15						1
A1	913	9	25	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		4
A1	914	9	26	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	2
A1	915	9	27	G	G	Overlap	0.43	0.25	-1.15						1
A1	916	9	28	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	5
A1	917	9	29	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	3
A1	918	9	30	G	G	Overlap	0.43	0.25	-1.15						1
A1	919	9	31	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	3
A1	920	9	32	T	T	Overlap	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	2
A1	921	9	33	G	G	Overlap	0.43	0.25	-1.15						1
A1	922	9	34	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564	2
A1	923	9	35	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	4
A1	924	9	36	G	G	Overlap	0.43	0.25	-1.15						1
A1	925	9	37	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	3
A1	926	9	38	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	4
A1	927	9	39	G	G	Overlap	0.43	0.25	-1.15						1
A1	928	9	40	E	E	Overlap	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	2

A1	929	9	41	V	V	Overlap	-1	-0.46	0.46	5.2987(	6.9578	60.07	-5.7587	-60.53	5
A1	930	9	42	G	G	Overlap	0.43	0.25	-1.15						1
A1	931	9	43	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	932	9	44	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	933	9	45	G	G	Overlap	0.43	0.25	-1.15						1
A1	934	9	46	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	935	9	47	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	936	9	48	G	G	Overlap	0.43	0.25	-1.15						1
A1	937	9	49	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	938	9	50	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	6
A1	939	9	51	G	G	Overlap	0.43	0.25	-1.15						1
A1	940	9	52	E	E	Overlap	-0.53	-0.14	-3.63	5.3136	6.5667(	69.742	-1.6836	-66.112	2
A1	941	9	53	K	K	Overlap	-0.67	0.4	-2.8	5.5897	8.1697(	72.811	-2.7897	-70.011	2
A1	942	9	54	G	G	Overlap	0.43	0.25	-1.15						1
A1	943	9	55	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	944	9	56	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	945	9	57	G	G	Overlap	0.43	0.25	-1.15						1
A1	946	9	58	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	947	9	59	D	D	Overlap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	3
A1	948	9	60	G	G	Overlap	0.43	0.25	-1.15						1
A1	949	9	61	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	950	9	62	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A1	951	9	63	G	G	Overlap	0.43	0.25	-1.15						1
A1	952	9	64	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A1	953	9	65	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	954	9	66	G	G	Overlap	0.43	0.25	-1.15						1
A1	955	9	67	T	T	Overlap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	4
A1	956	9	68	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	957	9	69	G	G	Overlap	0.43	0.25	-1.15						1
A1	958	9	70	P	P	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	959	9	71	Q	Q	Overlap	0.44	-0.19	-0.77	5.5148	7.3776(	74.788	-4.7448	-74.018	2
A1	960	9	72	G	G	Overlap	0.43	0.25	-1.15						1
A1	961	9	73	I	I	Overlap	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	3
A1	962	9	74	A	A	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A1	963	9	75	G	G	Overlap	0.43	0.25	-1.15						1
A1	964	9	76	Q	Q	Overlap	0.44	-0.19	-0.77	5.5148	7.3776(	74.788	-4.7448	-74.018	2
A1	965	9	77	R	R	Overlap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	966	9	78	G	G	Overlap	0.43	0.25	-1.15						1
A1	967	9	79	V	V	Overlap	-1	-0.46	0.46	5.2987(	6.9578	60.07	-5.7587	-60.53	3
A1	968	9	80	V	V	Gap	-1	-0.46	0.46	5.2987(	6.9578	60.07	-5.7587	-60.53	2
A1	969	9	81	G	G	Gap	0.43	0.25	-1.15						1
A1	970	9	82	L	L	Gap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A1	971	9	83	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	972	9	84	G	G	Gap	0.43	0.25	-1.15						1
A1	973	9	85	Q	Q	Gap	0.44	-0.19	-0.77	5.5148	7.3776(	74.788	-4.7448	-74.018	2
A1	974	9	86	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	975	9	87	G	G	Gap	0.43	0.25	-1.15						1
A1	976	9	88	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667(	69.742	-1.6836	-66.112	2
A1	977	9	89	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	978	9	90	G	G	Gap	0.43	0.25	-1.15						1
A1	979	9	91	F	F	Gap	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	2
A1	980	9	92	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	981	9	93	G	G	Gap	0.43	0.25	-1.15						1
A1	982	9	94	L	L	Gap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A1	983	9	95	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2

A1	984	9	96	G	G	Gap	0.43	0.25	-1.15						1
A1	985	9	97	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	986	9	98	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		2
A1	987	9	99	G	G	Gap	0.43	0.25	-1.15						1
A1	988	9	100	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A1	989	9	101	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	990	9	102	G	G	Gap	0.43	0.25	-1.15						1
A1	991	9	103	K	K	Gap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	2
A1	992	9	104	Q	Q	Gap	0.44	-0.19	-0.77	5.51489	7.37760	74.788	-4.7448	-74.018	2
A1	993	9	105	G	G	Gap	0.43	0.25	-1.15						1
A1	994	9	106	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	995	9	107	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		3
A1	996	9	108	G	G	Gap	0.43	0.25	-1.15						1
A1	997	9	109	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	7
A1	998	9	110	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		6
A1	999	9	111	G	G	Gap	0.43	0.25	-1.15						1
A1	1000	9	112	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	3
A1	1001	9	113	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	1
A1	1002	9	114	G	G	Gap	0.43	0.25	-1.15						1
A1	1003	9	115	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	1004	9	116	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1005	9	117	G	G	Gap	0.43	0.25	-1.15						1
A1	1006	9	118	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	1007	9	119	M	M	Gap	-1	-1	0.67	5.73531	7.93174	81.657	-6.4053	-82.327	4
A1	1008	9	120	G	G	Gap	0.43	0.25	-1.15						1
A1	1009	9	121	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	1
A1	1010	9	122	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1011	9	123	G	G	Gap	0.43	0.25	-1.15						1
A1	1012	9	124	L	L	Gap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	4
A1	1013	9	125	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	1014	9	126	G	G	Gap	0.43	0.25	-1.15						1
A1	1015	9	127	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1016	9	128	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	1017	9	129	G	G	Gap	0.43	0.25	-1.15						1
A1	1018	9	130	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	1
A1	1019	9	131	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		4
A1	1020	9	132	G	G	Gap	0.43	0.25	-1.15						1
A1	1021	9	133	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	1022	9	134	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	2
A1	1023	9	135	G	G	Gap	0.43	0.25	-1.15						1
A1	1024	9	136	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	1025	9	137	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1026	9	138	G	G	Gap	0.43	0.25	-1.15						1
A1	1027	9	139	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	4
A1	1028	9	140	E	E	Gap	-0.53	-0.14	-3.63	5.31362	6.56670	69.742	-1.6836	-66.112	3
A1	1029	9	141	G	G	Gap	0.43	0.25	-1.15						1
A1	1030	9	142	S	S	Gap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		4
A1	1031	9	143	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1032	9	144	G	G	Gap	0.43	0.25	-1.15						1
A1	1033	9	145	R	R	Gap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	1034	9	146	D	D	Gap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	1035	9	147	G	G	Gap	0.43	0.25	-1.15						1
A1	1036	9	148	A	A	Gap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	1037	9	149	P	P	Gap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1038	9	150	G	G	Gap	0.43	0.25	-1.15						1

A1	1039	9	151	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	1040	9	152	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A1	1041	9	153	G	G	Gap	0.43	0.25	-1.15						1
A1	1042	9	154	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	3
A1	1043	9	155	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1044	9	156	G	G	Gap	0.43	0.25	-1.15						1
A1	1045	9	157	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	1
A1	1046	9	158	T	T	Gap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	5
A1	1047	9	159	G	G	Gap	0.43	0.25	-1.15						1
A1	1048	9	160	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1049	9	161	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A1	1050	9	162	G	G	Gap	0.43	0.25	-1.15						1
A1	1051	9	163	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1052	9	164	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1053	9	165	G	G	Gap	0.43	0.25	-1.15						1
A1	1054	9	166	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	1055	9	167	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1056	9	168	G	G	Gap	0.43	0.25	-1.15						1
A1	1057	9	169	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	1058	9	170	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1059	9	171	G	G	Gap	0.43	0.25	-1.15						1
A1	1060	9	172	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	1061	9	173	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1062	9	174	G	G	Gap	0.43	0.25	-1.15						1
A1	1063	9	175	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1064	9	176	V	V	Gap	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A1	1065	9	177	G	G	Gap	0.43	0.25	-1.15						1
A1	1066	9	178	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1067	9	179	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	1068	9	180	G	G	Gap	0.43	0.25	-1.15						1
A1	1069	9	181	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	1
A1	1070	9	182	S	S	Gap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		5
A1	1071	9	183	G	G	Gap	0.43	0.25	-1.15						1
A1	1072	9	184	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	3
A1	1073	9	185	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1074	9	186	G	G	Gap	0.43	0.25	-1.15						1
A1	1075	9	187	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A1	1076	9	188	T	T	Gap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	4
A1	1077	9	189	G	G	Gap	0.43	0.25	-1.15						1
A1	1078	9	190	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1079	9	191	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	4
A1	1080	9	192	G	G	Gap	0.43	0.25	-1.15						1
A1	1081	9	193	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1082	9	194	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	1083	9	195	G	G	Gap	0.43	0.25	-1.15						1
A1	1084	9	196	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1085	9	197	I	I	Gap	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	5
A1	1086	9	198	G	G	Gap	0.43	0.25	-1.15						1
A1	1087	9	199	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1088	9	200	V	V	Gap	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	3
A1	1089	9	201	G	G	Gap	0.43	0.25	-1.15						1
A1	1090	9	202	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A1	1091	9	203	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1092	9	204	G	G	Gap	0.43	0.25	-1.15						1
A1	1093	9	205	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1

A1	1094	9	206	A	A	Gap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	5
A1	1095	9	207	G	G	Gap	0.43	0.25	-1.15						1
A1	1096	9	208	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	1
A1	1097	9	209	Q	Q	Gap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	2
A1	1098	9	210	G	G	Gap	0.43	0.25	-1.15						1
A1	1099	9	211	P	P	Gap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1100	9	212	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1101	9	213	G	G	Gap	0.43	0.25	-1.15						1
A1	1102	9	214	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	2
A1	1103	9	215	K	K	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A1	1104	9	216	G	G	Gap	0.43	0.25	-1.15						1
A1	1105	9	217	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A1	1106	9	218	T	T	Gap	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	3
A1	1107	9	219	G	G	Gap	0.43	0.25	-1.15						1
A1	1108	9	220	E	E	Gap	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	2
A1	1109	9	221	Q	Q	Gap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	3
A1	1110	9	222	G	G	Gap	0.43	0.25	-1.15						2
A1	1111	9	223	D	D	Gap	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	3
A1	1112	9	224	R	R	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1113	9	225	G	X	Gap	0.43	0.25	-1.15						2
A1	1114	9	226	I	X	Gap	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	4
A1	1115	9	227	K	X	Gap	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	2
A1	1116	9	228	G	X	Gap	0.43	0.25	-1.15						2
A1	1117	9	229	H	X	Gap	0.25	-0.49	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	2
A1	1118	9	230	R	X	Gap	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	2
A1	1119	9	231	G	X	Gap	0.43	0.25	-1.15						2
A1	1120	9	232	F	X	Gap	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	4
A1	1121	9	233	S	X	Gap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		2
A1	1122	12	0	G	X	Overlap	0.43	0.25	-1.15						2
A1	1123	12	1	L	X	Overlap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	3
A1	1124	12	2	Q	X	Overlap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	2
A1	1125	12	3	G	X	Overlap	0.43	0.25	-1.15						2
A1	1126	12	4	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	1127	12	5	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	1128	12	6	G	X	Overlap	0.43	0.25	-1.15						2
A1	1129	12	7	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	4
A1	1130	12	8	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	1131	12	9	V	X	Overlap	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	2
A1	1132	12	10	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		4
A1	1133	12	11	M	X	Overlap	-1	-1	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	3
A1	1134	12	12	L	X	Overlap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A1	1135	12	13	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		3
A1	1136	12	14	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	3
A1	1137	12	15	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		2
A1	1138	12	16	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1139	12	17	A	X	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	3
A1	1140	12	18	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1141	12	19	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		3
A1	1142	12	20	P	X	Overlap	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	2
A1	1143	12	21	A	X	Overlap	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	2
A1	1144	12	22	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		2
A1	1145	12	23	S	X	Overlap	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		2
A1	1146	12	24	L	X	Overlap	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	2
A1	1147	12	25	Q	X	Overlap	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	3
A1	1148	12	26	G	G	Overlap	0.43	0.25	-1.15						2

A1	1149	12	27	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1150	12	28	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1151	12	29	G	G	Overlap	0.43	0.25	-1.15						2
A1	1152	12	30	S	S	Overlap	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		4
A1	1153	12	31	A	A	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	1154	12	32	G	G	Overlap	0.43	0.25	-1.15						2
A1	1155	12	33	T	T	Overlap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	5
A1	1156	12	34	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	1157	12	35	G	G	Overlap	0.43	0.25	-1.15						2
A1	1158	12	36	K	K	Overlap	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	3
A1	1159	12	37	D	D	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	2
A1	1160	12	38	G	G	Overlap	0.43	0.25	-1.15						2
A1	1161	12	39	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	5
A1	1162	12	40	N	N	Overlap	0.44	-1	-0.85	5.10027	5.39194	59.477	-4.2502	-58.627	3
A1	1163	12	41	G	G	Overlap	0.43	0.25	-1.15						2
A1	1164	12	42	L	L	Overlap	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	4
A1	1165	12	43	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	3
A1	1166	12	44	G	G	Overlap	0.43	0.25	-1.15						2
A1	1167	12	45	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1168	12	46	I	I	Overlap	-0.51	-0.07	1.12	5.80934	6.29714	75.102	-6.9293	-76.222	3
A1	1169	12	47	G	G	Overlap	0.43	0.25	-1.15						2
A1	1170	12	48	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1171	12	49	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1172	12	50	G	G	Overlap	0.43	0.25	-1.15						2
A1	1173	12	51	P	P	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1174	12	52	R	R	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	2
A1	1175	12	53	G	X	Overlap	0.43	0.25	-1.15						2
A1	1176	12	54	R	X	Overlap	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	3
A1	1177	12	55	T	X	Overlap	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	3
A1	1178	12	56	G	X	Overlap	0.43	0.25	-1.15						2
A1	1179	12	57	D	X	Overlap	0.54	-0.13	-3.64	4.74709	5.60968	55.197	-1.1070	-51.557	3
A1	1180	12	58	A	X	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	5
A1	1181	12	59	G	X	Overlap	0.43	0.25	-1.15						2
A1	1182	12	60	P	X	Overlap	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	2
A1	1183	12	61	A	X	Overlap	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	3
A1	1184	12	62	V	X	C terminal	-1	-0.46	0.46	5.29870	6.95785	60.07	-5.7587	-60.53	
A1	1185	12	63	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		
A1	1186	12	64	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1187	12	65	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1188	12	66	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.28055	51.458	-4.6529	-51.208	
A1	1189	12	67	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	
A1	1190	12	68	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		
A1	1191	12	69	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	
A1	1192	12	70	H	X	C terminal	0.25	-0.49	-2.33	5.70630	7.17634	80.668	-3.3763	-78.338	
A1	1193	12	71	G	X	C terminal	0.43	0.25	-1.15						
A1	1194	12	72	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1195	12	73	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1196	12	74	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	
A1	1197	12	75	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.71701	72.686	-6.9575	-73.936	
A1	1198	12	76	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.16970	72.811	-2.7897	-70.011	
A1	1199	12	77	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.92453	31.474	-3.4244	-30.974	
A1	1200	12	78	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1201	12	79	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.84426		-3.7662		
A1	1202	12	80	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.61868	80.704	-5.4368	-80.564	
A1	1203	12	81	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.78488	57.551	-4.5903	-55.741	

A1	1204	12	82	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1205	12	83	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1206	12	84	Y	X	C terminal	0.01	0.53	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A1	1207	12	85	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1208	12	86	L	X	C terminal	-0.73	-0.78	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A1	1209	12	87	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1210	12	88	F	X	C terminal	-0.85	-0.81	1.71	5.9637(	8.4757(	83.928	-7.6737	-85.638	
A1	1211	12	89	L	X	C terminal	-0.73	-0.78	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A1	1212	12	90	P	X	C terminal	-0.1	-0.07	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A1	1213	12	91	Q	X	C terminal	0.44	-0.19	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	
A1	1214	12	92	P	X	C terminal	-0.1	-0.07	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A1	1215	12	93	P	X	C terminal	-0.1	-0.07	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	
A1	1216	12	94	Q	X	C terminal	0.44	-0.19	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	
A1	1217	12	95	E	X	C terminal	-0.53	-0.14	-3.63	5.3136(	6.5667(	69.742	-1.6836	-66.112	
A1	1218	12	96	K	X	C terminal	-0.67	0.4	-2.8	5.5897(	8.1697(	72.811	-2.7897	-70.011	
A1	1219	12	97	A	A	C terminal	0.05	-0.43	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A1	1220	12	98	H	H	C terminal	0.25	-0.49	-2.33	5.7063(	7.1763(	80.668	-3.3763	-78.338	
A1	1221	12	99	D	D	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1222	12	100	G	G	C terminal	0.43	0.25	-1.15						
A1	1223	12	101	G	G	C terminal	0.43	0.25	-1.15						
A1	1224	12	102	R	R	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1225	12	103	Y	Y	C terminal	0.01	0.53	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A1	1226	12	104	Y	Y	C terminal	0.01	0.53	0.71	6.1879(	8.0649(	91.178	-6.8979	-91.888	
A1	1227	12	105	R	R	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1228	12	106	A	X	C terminal	0.05	-0.43	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A1	1229	12	107	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1230	12	108	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1231	12	109	A	X	C terminal	0.05	-0.43	-0.5	3.9244(	3.9245(	31.474	-3.4244	-30.974	
A1	1232	12	110	N	X	C terminal	0.44	-1	-0.85	5.1002(	5.3919(	59.477	-4.2502	-58.627	
A1	1233	12	111	V	X	C terminal	-1	-0.46	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	
A1	1234	12	112	V	X	C terminal	-1	-0.46	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	
A1	1235	12	113	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1236	12	114	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1237	12	115	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1238	12	116	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1239	12	117	L	X	C terminal	-0.73	-0.78	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A1	1240	12	118	E	X	C terminal	-0.53	-0.14	-3.63	5.3136(	6.5667(	69.742	-1.6836	-66.112	
A1	1241	12	119	V	X	C terminal	-1	-0.46	0.46	5.2987(	6.9578(	60.07	-5.7587	-60.53	
A1	1242	12	120	D	X	C terminal	0.54	-0.13	-3.64	4.7470(	5.6096(	55.197	-1.1070	-51.557	
A1	1243	12	121	T	X	C terminal	0.31	0.09	-0.25	4.9029(	6.2805(	51.458	-4.6529	-51.208	
A1	1244	12	122	T	X	C terminal	0.31	0.09	-0.25	4.9029(	6.2805(	51.458	-4.6529	-51.208	
A1	1245	12	123	L	X	C terminal	-0.73	-0.78	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A1	1246	12	124	K	X	C terminal	-0.67	0.4	-2.8	5.5897(	8.1697(	72.811	-2.7897	-70.011	
A1	1247	12	125	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1248	12	126	L	X	C terminal	-0.73	-0.78	1.25	5.7075(	5.7170(	72.686	-6.9575	-73.936	
A1	1249	12	127	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1250	12	128	Q	X	C terminal	0.44	-0.19	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	
A1	1251	12	129	Q	X	C terminal	0.44	-0.19	-0.77	5.5148(	7.3776(	74.788	-4.7448	-74.018	
A1	1252	12	130	I	X	C terminal	-0.51	-0.07	1.12	5.8093(	6.2971(	75.102	-6.9293	-76.222	
A1	1253	12	131	E	X	C terminal	-0.53	-0.14	-3.63	5.3136(	6.5667(	69.742	-1.6836	-66.112	
A1	1254	12	132	N	X	C terminal	0.44	-1	-0.85	5.1002(	5.3919(	59.477	-4.2502	-58.627	
A1	1255	12	133	I	X	C terminal	-0.51	-0.07	1.12	5.8093(	6.2971(	75.102	-6.9293	-76.222	
A1	1256	12	134	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1257	12	135	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1258	12	136	P	X	C terminal	-0.1	-0.07	-0.14	5.5768(	5.6186(	80.704	-5.4368	-80.564	

A1	1259	12	137	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112
A1	1260	12	138	G	X	C terminal	0.43	0.25	-1.15					
A1	1261	12	139	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662	
A1	1262	12	140	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741
A1	1263	12	141	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011
A1	1264	12	142	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627
A1	1265	12	143	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564
A1	1266	12	144	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974
A1	1267	12	145	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741
A1	1268	12	146	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208
A1	1269	12	147	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675	
A1	1270	12	148	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741
A1	1271	12	149	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557
A1	1272	12	150	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936
A1	1273	12	151	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011
A1	1274	12	152	M	X	C terminal	-1	-1	0.67	5.73531	7.9317(	81.657	-6.4053	-82.327
A1	1275	12	153	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675	
A1	1276	12	154	H	X	C terminal	0.25	-0.49	-2.33	5.70630	7.1763(	80.668	-3.3763	-78.338
A1	1277	12	155	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564
A1	1278	12	156	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557
A1	1279	12	157	W	X	C terminal	-0.56	-0.79	2.09	6.50540	9.4442(	81.36	-8.5954	-83.45
A1	1280	12	158	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011
A1	1281	12	159	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662	
A1	1282	12	160	G	X	C terminal	0.43	0.25	-1.15					
A1	1283	12	161	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112
A1	1284	12	162	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888
A1	1285	12	163	W	X	C terminal	-0.56	-0.79	2.09	6.50540	9.4442(	81.36	-8.5954	-83.45
A1	1286	12	164	I	X	C terminal	-0.51	-0.07	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222
A1	1287	12	165	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557
A1	1288	12	166	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564
A1	1289	12	167	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627
A1	1290	12	168	Q	X	C terminal	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018
A1	1291	12	169	G	X	C terminal	0.43	0.25	-1.15					
A1	1292	12	170	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675	
A1	1293	12	171	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627
A1	1294	12	172	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936
A1	1295	12	173	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557
A1	1296	12	174	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974
A1	1297	12	175	I	X	C terminal	-0.51	-0.07	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222
A1	1298	12	176	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011
A1	1299	12	177	V	X	C terminal	-1	-0.46	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53
A1	1300	12	178	F	X	C terminal	-0.85	-0.81	1.71	5.96373	8.4757(	83.928	-7.6737	-85.638
A1	1301	12	179	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675	
A1	1302	12	180	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627
A1	1303	12	181	M	X	C terminal	-1	-1	0.67	5.73531	7.9317(	81.657	-6.4053	-82.327
A1	1304	12	182	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112
A1	1305	12	183	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208
A1	1306	12	184	G	X	C terminal	0.43	0.25	-1.15					
A1	1307	12	185	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112
A1	1308	12	186	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208
A1	1309	12	187	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675	
A1	1310	12	188	V	X	C terminal	-1	-0.46	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53
A1	1311	12	189	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888
A1	1312	12	190	P	X	C terminal	-0.1	-0.07	-0.14	5.57682	5.6186(	80.704	-5.4368	-80.564
A1	1313	12	191	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208

A1	1314	12	192	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1315	12	193	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1316	12	194	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1317	12	195	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1318	12	196	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1319	12	197	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1320	12	198	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1321	12	199	N	X	C terminal	0.44	-1	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A1	1322	12	200	W	X	C terminal	-0.56	-0.79	2.09	6.5054	9.4442	81.36	-8.5954	-83.45	
A1	1323	12	201	Y	X	C terminal	0.01	0.53	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A1	1324	12	202	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1325	12	203	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1326	12	204	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1327	12	205	N	X	C terminal	0.44	-1	-0.85	5.1002	5.3919	59.477	-4.2502	-58.627	
A1	1328	12	206	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1329	12	207	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1330	12	208	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1331	12	209	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1332	12	210	R	X	C terminal	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	1333	12	211	H	X	C terminal	0.25	-0.49	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A1	1334	12	212	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1335	12	213	W	X	C terminal	-0.56	-0.79	2.09	6.5054	9.4442	81.36	-8.5954	-83.45	
A1	1336	12	214	Y	X	C terminal	0.01	0.53	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A1	1337	12	215	G	X	C terminal	0.43	0.25	-1.15						
A1	1338	12	216	E	X	C terminal	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	1339	12	217	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1340	12	218	M	X	C terminal	-1	-1	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	
A1	1341	12	219	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1342	12	220	G	X	C terminal	0.43	0.25	-1.15						
A1	1343	12	221	G	X	C terminal	0.43	0.25	-1.15						
A1	1344	12	222	F	X	C terminal	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	1345	12	223	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1346	12	224	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1347	12	225	R	X	C terminal	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	1348	12	226	E	X	C terminal	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	1349	12	227	G	X	C terminal	0.43	0.25	-1.15						
A1	1350	12	228	G	X	C terminal	0.43	0.25	-1.15						
A1	1351	12	229	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1352	12	230	G	X	C terminal	0.43	0.25	-1.15						
A1	1353	12	231	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1354	12	232	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1355	12	233	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1356	12	234	A	X	C terminal	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	1357	12	235	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1358	12	236	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1359	12	237	A	X	C terminal	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	1360	12	238	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1361	12	239	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1362	12	240	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	1363	12	241	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1364	12	242	F	X	C terminal	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	1365	12	243	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	1366	12	244	R	X	C terminal	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	1367	12	245	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	1368	12	246	M	X	C terminal	-1	-1	0.67	5.7353	7.9317	81.657	-6.4053	-82.327	

A1	1369	12	247	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1370	12	248	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A1	1371	12	249	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A1	1372	12	250	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A1	1373	12	251	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1374	12	252	Q	X	C terminal	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A1	1375	12	253	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A1	1376	12	254	I	X	C terminal	-0.51	-0.07	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A1	1377	12	255	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A1	1378	12	256	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A1	1379	12	257	H	X	C terminal	0.25	-0.49	-2.33	5.70630	7.1763(	80.668	-3.3763	-78.338	
A1	1380	12	258	C	X	C terminal	1.31	-0.95	0.02	4.54756	6.0951(		-4.5675		
A1	1381	12	259	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A1	1382	12	260	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A1	1383	12	261	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1384	12	262	V	X	C terminal	-1	-0.46	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	
A1	1385	12	263	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A1	1386	12	264	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A1	1387	12	265	M	X	C terminal	-1	-1	0.67	5.73531	7.9317(	81.657	-6.4053	-82.327	
A1	1388	12	266	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	
A1	1389	12	267	Q	X	C terminal	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A1	1390	12	268	Q	X	C terminal	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A1	1391	12	269	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A1	1392	12	270	G	X	C terminal	0.43	0.25	-1.15						
A1	1393	12	271	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A1	1394	12	272	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936	
A1	1395	12	273	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A1	1396	12	274	K	X	C terminal	-0.67	0.4	-2.8	5.58972	8.1697(	72.811	-2.7897	-70.011	
A1	1397	12	275	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A1	1398	12	276	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936	
A1	1399	12	277	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936	
A1	1400	12	278	L	X	C terminal	-0.73	-0.78	1.25	5.70759	5.7170(	72.686	-6.9575	-73.936	
A1	1401	12	279	Q	X	C terminal	0.44	-0.19	-0.77	5.51489	7.3776(	74.788	-4.7448	-74.018	
A1	1402	12	280	G	X	C terminal	0.43	0.25	-1.15						
A1	1403	12	281	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1404	12	282	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A1	1405	12	283	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A1	1406	12	284	I	X	C terminal	-0.51	-0.07	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A1	1407	12	285	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A1	1408	12	286	I	X	C terminal	-0.51	-0.07	1.12	5.80934	6.2971(	75.102	-6.9293	-76.222	
A1	1409	12	287	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1410	12	288	A	X	C terminal	0.05	-0.43	-0.5	3.92448	3.9245(	31.474	-3.4244	-30.974	
A1	1411	12	289	E	X	C terminal	-0.53	-0.14	-3.63	5.31362	6.5667(	69.742	-1.6836	-66.112	
A1	1412	12	290	G	X	C terminal	0.43	0.25	-1.15						
A1	1413	12	291	N	X	C terminal	0.44	-1	-0.85	5.10027	5.3919(	59.477	-4.2502	-58.627	
A1	1414	12	292	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1415	12	293	R	X	C terminal	0.5	-0.3	-1.81	6.40037	9.7848(	57.551	-4.5903	-55.741	
A1	1416	12	294	F	X	C terminal	-0.85	-0.81	1.71	5.96373	8.4757(	83.928	-7.6737	-85.638	
A1	1417	12	295	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A1	1418	12	296	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A1	1419	12	297	S	X	C terminal	0.43	2.62	-0.46	4.22627	4.8442(		-3.7662		
A1	1420	12	298	V	X	C terminal	-1	-0.46	0.46	5.29870	6.9578(	60.07	-5.7587	-60.53	
A1	1421	12	299	T	X	C terminal	0.31	0.09	-0.25	4.90291	6.2805(	51.458	-4.6529	-51.208	
A1	1422	12	300	Y	X	C terminal	0.01	0.53	0.71	6.18798	8.0649(	91.178	-6.8979	-91.888	
A1	1423	12	301	D	X	C terminal	0.54	-0.13	-3.64	4.74709	5.6096(	55.197	-1.1070	-51.557	

A1	1424	12	302	G	X	C terminal	0.43	0.25	-1.15						
A1	1425	12	303	C	X	C terminal	1.31	-0.95	0.02	4.5475	6.0951		-4.5675		
A1	1426	12	304	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1427	12	305	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1428	12	306	H	X	C terminal	0.25	-0.49	-2.33	5.7063	7.1763	80.668	-3.3763	-78.338	
A1	1429	12	307	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1430	12	308	G	X	C terminal	0.43	0.25	-1.15						
A1	1431	12	309	A	X	C terminal	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	1432	12	310	W	X	C terminal	-0.56	-0.79	2.09	6.5054	9.4442	81.36	-8.5954	-83.45	
A1	1433	12	311	G	X	C terminal	0.43	0.25	-1.15						
A1	1434	12	312	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1435	12	313	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1436	12	314	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1437	12	315	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1438	12	316	E	X	C terminal	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	1439	12	317	Y	X	C terminal	0.01	0.53	0.71	6.1879	8.0649	91.178	-6.8979	-91.888	
A1	1440	12	318	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1441	12	319	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1442	12	320	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1443	12	321	K	X	C terminal	-0.67	0.4	-2.8	5.5897	8.1697	72.811	-2.7897	-70.011	
A1	1444	12	322	T	X	C terminal	0.31	0.09	-0.25	4.9029	6.2805	51.458	-4.6529	-51.208	
A1	1445	12	323	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1446	12	324	R	X	C terminal	0.5	-0.3	-1.81	6.4003	9.7848	57.551	-4.5903	-55.741	
A1	1447	12	325	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	1448	12	326	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1449	12	327	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1450	12	328	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1451	12	329	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1452	12	330	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1453	12	331	A	X	C terminal	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	1454	12	332	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1455	12	333	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	
A1	1456	12	334	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1457	12	335	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1458	12	336	G	X	C terminal	0.43	0.25	-1.15						
A1	1459	12	337	A	X	C terminal	0.05	-0.43	-0.5	3.9244	3.9245	31.474	-3.4244	-30.974	
A1	1460	12	338	P	X	C terminal	-0.1	-0.07	-0.14	5.5768	5.6186	80.704	-5.4368	-80.564	
A1	1461	12	339	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1462	12	340	Q	X	C terminal	0.44	-0.19	-0.77	5.5148	7.3776	74.788	-4.7448	-74.018	
A1	1463	12	341	E	X	C terminal	-0.53	-0.14	-3.63	5.3136	6.5667	69.742	-1.6836	-66.112	
A1	1464	12	342	F	X	C terminal	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	1465	12	343	G	X	C terminal	0.43	0.25	-1.15						
A1	1466	12	344	F	X	C terminal	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	1467	12	345	D	X	C terminal	0.54	-0.13	-3.64	4.7470	5.6096	55.197	-1.1070	-51.557	
A1	1468	12	346	I	X	C terminal	-0.51	-0.07	1.12	5.8093	6.2971	75.102	-6.9293	-76.222	
A1	1469	12	347	G	X	C terminal	0.43	0.25	-1.15						
A1	1470	12	348	S	X	C terminal	0.43	2.62	-0.46	4.2262	4.8442		-3.7662		
A1	1471	12	349	V	X	C terminal	-1	-0.46	0.46	5.2987	6.9578	60.07	-5.7587	-60.53	
A1	1472	12	350	C	X	C terminal	1.31	-0.95	0.02	4.5475	6.0951		-4.5675		
A1	1473	12	351	F	X	C terminal	-0.85	-0.81	1.71	5.9637	8.4757	83.928	-7.6737	-85.638	
A1	1474	12	352	L	X	C terminal	-0.73	-0.78	1.25	5.7075	5.7170	72.686	-6.9575	-73.936	

Appendix C, Table 1: Sequence information for COL1A1 and COL1A2

Appendix C, Table 2

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
85.5	-1	-1	K	V	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
95.5	-1	-1	P	H	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
127.5	-1	-1	H	S	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
131.5	-1	-1	Q	G	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
139.5	-1	-1	R	D	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
145.5	-1	-1	Q	P	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
149.5	-1	-1	P	G	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
152.5	-1	-1	P	G	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
159.5	-1	-1	G	P	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
168.5	-1	-1	A	P	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
169.5	-1	-1	P	Q	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
178.5	-1	-1	K	S	N-term	141784	429545	179549	315189	0	0	0	104690	538590	0	0	450990	313880	337810	276430	321970	269180	846460	361920	347950	212710	533910	798470	586330
179.5	-1	-1	S	T	N-term	300889	862952	115107	623411	0	0	0	196790	909060	0	0	923460	0	915110	142330	523480	713550	133980	192220	786930	399100	119690	125880	117400
180.5	-1	-1	T	G	N-term	0	110270	218287	671118	0	0	0	493660	0	0	0	237180	0	0	250270	629350	0	252440	368040	827710	0	188640	254840	120369
181.5	-1	-1	G	I	N-term	0	315425	735417	380296	0	0	0	140380	0	0	0	556640	0	0	105670	418060	0	126170	127420	519350	0	0	610770	578210
182.5	-1	-1	I	S	N-term	0	0	0	254677	0	0	0	233870	0	0	0	0	0	0	0	317150	0	0	0	0	0	0	0	701560
183.5	-1	-1	S	V	N-term	0	0	0	123257	0	0	0	462490	0	0	0	0	0	0	0	113480	0	0	0	162230	0	0	0	217320
184.5	-1	-1	V	P	N-term	0	0	261900	544332	0	0	0	103590	0	0	0	0	0	0	0	829630	0	0	104760	395900	0	0	0	951800
185.5	-1	-1	P	G	N-term	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
186.5	0	0.5	G	P	Overlap	0	116017	281225	440045	0	0	0	167540	0	0	0	0	0	0	365660	572750	0	0	403230	339970	0	464070	356010	847460
192.5	0	6.5	G	P	Overlap	0	0	0	299200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	119680
194.5	0	8.5	R	G	Overlap	171873	672672	618070	207695	0	0	0	981321	629496	0	0	430289	313880	442994	811266	200149	340535	135905	110377	253611	252620	888647	557245	372719
202.5	0	16.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
204.5	0	18.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
207.5	0	21.5	G	F	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
208.5	0	22.5	F	Q	Overlap	0	341825	843425	0	0	0	0	0	0	0	0	0	0	136730	0	0	0	0	337370	0	0	0	0	0
209.5	0	23.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
210.5	0	24.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
211.5	0	25.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
212.5	0	26.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
213.5	0	27.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
214.5	0	28.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
215.5	0	29.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
216.5	0	30.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
217.5	0	31.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
218.5	0	32.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
219.5	0	33.5	G	A	Overlap	0	0	161235	745915	0	0	0	376250	0	0	0	0	0	0	207120	783660	0	0	233520	0	0	0	204300	220000
227.5	0	41.5	R	G	Overlap	0	0	161235	745915	0	0	0	376250	0	0	0	0	0	0	207120	783660	0	0	233520	0	0	0	204300	220000

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
229.5	0	43.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
235.5	0	49.5	K	N	Overlap	0	0	0	340386	0	0	0	0	0	0	0	0	0	0	0	795880	0	0	0	110770	0	0	0	253050
236.5	0	50.5	N	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
237.5	0	51.5	G	D	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
238.5	0	52.5	D	D	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
239.5	0	53.5	D	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
247.5	0	61.5	R	P	Overlap	0	0	0	340386	0	0	0	0	0	0	0	0	0	0	0	795880	0	0	0	110770	0	0	0	253050
251.5	0	65.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
252.5	0	66.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
256.5	0	70.5	P	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
260.5	0	74.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
262.5	0	76.5	L	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
270.5	0	84.5	G	M	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
272.5	0	86.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
274.5	0	88.5	H	R	Gap	0	0	381220	723125	0	0	0	165010	0	0	0	0	0	0	0	0	0	0	573460	289250	0	0	951420	0
275.5	0	89.5	R	G	Gap	631200	110552	221225	842249	110860	515330	148044	123688	0	320310	163414	227871	0	224859	161774	704204	252480	134811	260386	136611	0	505080	299324	107080
276.5	0	90.5	G	F	Gap	0	0	0	157462	0	0	0	0	0	0	0	171980	0	0	0	457870	0	0	0	0	0	0	0	0
277.5	0	91.5	F	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
278.5	0	92.5	S	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
284.5	0	98.5	K	G	Gap	631200	110059	223406	840424	110860	515330	153231	133895	0	320310	162100	233707	0	223640	167465	706698	252480	134060	259433	136049	0	505080	304628	106080
285.5	0	99.5	G	D	Gap	0	0	284425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	113770	0
286.5	0	100.5	D	A	Gap	0	0	928925	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	371570	0
287.5	0	101.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
293.5	0	107.5	K	G	Gap	0	492752	661270	638155	0	0	639589	110385	0	0	131490	491843	0	121930	724063	208400	0	751710	152680	851880	0	0	163685	100050
297.5	0	111.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
298.5	0	112.5	S	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
300.5	0	114.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
302.5	0	116.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
303.5	0	117.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
311.5	0	125.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
313.5	0	127.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
316.5	0	130.5	E	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
317.5	0	131.5	R	G	Gap	0	965650	161337	402920	0	0	0	562110	0	0	722940	719700	0	0	146400	528270	0	386260	472050	544110	0	0	196720	467330
318.5	0	132.5	G	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
319.5	0	133.5	R	P	Gap	0	0	0	129307	0	0	0	0	0	0	0	0	0	0	0	517230	0	0	0	0	0	0	0	0
320.5	0	134.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
321.5	0	135.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
326.5	0	140.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
327.5	0	141.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
328.5	0	142.5	A	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
329.5	0	143.5	R	G	Gap	0	965650	161337	404213	0	0	0	562110	0	0	722940	719700	0	0	146400	533442	0	386260	472050	544110	0	0	196720	467330
331.5	0	145.5	N	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
332.5	0	146.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
336.5	0	150.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
337.5	0	151.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
338.5	0	152.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
339.5	0	153.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
340.5	0	154.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
341.5	0	155.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
342.5	0	156.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
343.5	0	157.5	P	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
344.5	0	158.5	T	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
345.5	0	159.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
346.5	0	160.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
347.5	0	161.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
348.5	0	162.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
349.5	0	163.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
351.5	0	165.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
352.5	0	166.5	F	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
354.5	0	168.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
357.5	0	171.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
359.5	0	173.5	K	G	Gap	110452	690535	395520	388557	555170	104080	445700	102560	300090	272240	858610	212010	0	0	312460	216680	135360	248990	556919	729870	276440	0	626840	586480
360.5	0	174.5	G	E	Gap	0	0	0	174980	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	450700	0	0	0	249220
368.5	0	182.5	R	G	Gap	237227	205466	101451	711412	120110	271520	208410	271160	668650	702360	363611	476600	116730	0	694620	398550	275060	629540	143845	143318	595310	122090	156139	966252
369.5	0	183.5	G	S	Gap	0	0	0	535650	0	0	0	0	0	0	0	128450	0	0	0	479150	0	0	0	153500	0	0	0	0
377.5	0	191.5	R	G	Gap	126774	136413	623682	321640	645930	167440	163840	168600	368560	430120	277750	265874	116730	0	382160	182349	139700	380550	900269	700345	318870	122090	934550	377280
378.5	0	192.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
384.5	0	198.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
387.5	0	201.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
388.5	0	202.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
389.5	0	203.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
390.5	0	204.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
391.5	0	205.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
392.5	0	206.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
393.5	0	207.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
394.5	0	208.5	N	P	Gap	0	0	159037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	636150	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
396.5	0	210.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
398.5	0	212.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
404.5	0	218.5	K	G	Gap	0	0	159037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	636150	0	0	0	0	0
405.5	0	219.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
406.5	0	220.5	A	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
407.5	0	221.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
408.5	0	222.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
409.5	0	223.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
410.5	0	224.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
411.5	0	225.5	G	I	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
412.5	0	226.5	I	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
413.5	0	227.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
414.5	0	228.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
418.5	0	232.5	F	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
420.5	3	0.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
422.5	3	2.5	R	G	Overlap	0	826182	276172	479722	0	0	116671	254316	0	283700	226261	279250	0	230940	191530	335340	0	711630	620600	844770	0	0	662970	459530
423.5	3	3.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
424.5	3	4.5	P	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
425.5	3	5.5	S	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
426.5	3	6.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
428.5	3	8.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
429.5	3	9.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
431.5	3	11.5	S	G	Overlap	0	0	0	0	0	0	641110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
433.5	3	13.5	P	P	Overlap	0	0	337125	630525	0	0	110260	212280	0	0	134850	114700	0	0	0	0	0	0	0	0	0	0	0	137510
435.5	3	15.5	G	P	Overlap	0	0	0	949750	0	0	0	0	0	0	0	379900	0	0	0	0	0	0	0	0	0	0	0	0
437.5	3	17.5	K	G	Overlap	0	826182	242459	407172	0	0	0	420360	0	283700	914110	126560	0	230940	191530	335340	0	711630	620600	844770	0	0	662970	322020
438.5	3	18.5	G	N	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
439.5	3	19.5	N	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
444.5	3	24.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
447.5	3	27.5	G	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
449.5	3	29.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
451.5	3	31.5	D	T	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
455.5	3	35.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
457.5	3	37.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
458.5	3	38.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
459.5	3	39.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
461.5	3	41.5	T	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
462.5	3	42.5	G	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
463.5	3	43.5	I	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
464.5	3	44.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
465.5	3	45.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
466.5	3	46.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
468.5	3	48.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
470.5	3	50.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
471.5	3	51.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
473.5	3	53.5	E	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
474.5	3	54.5	G	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
475.5	3	55.5	K	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
476.5	3	56.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
479.5	3	59.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
480.5	3	60.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
481.5	3	61.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
482.5	3	62.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
483.5	3	63.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485.5	3	65.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
492.5	3	72.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
494.5	3	74.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
500.5	3	80.5	R	G	Gap	0	0	783200	415985	0	0	0	0	0	0	0	146890	0	0	0	391750	0	0	313280	0	0	0	0	112530
501.5	3	81.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
502.5	3	82.5	F	P	Gap	0	0	0	0	0	0	193090	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
503.5	3	83.5	P	G	Gap	0	0	0	197887	206000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	791550	0	0	0	0
504.5	3	84.5	G	S	Gap	0	434550	716400	0	0	0	0	0	0	173820	0	0	0	0	0	0	0	0	0	0	0	0	286560	0
506.5	3	86.5	D	G	Gap	0	0	413350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	165340	0
508.5	3	88.5	V	A	Gap	0	0	503975	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	201590	0
509.5	3	89.5	A	G	Gap	0	0	119560	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	478240	0
512.5	3	92.5	K	G	Gap	0	434550	149960	613872	206000	0	0	0	0	173820	0	146890	0	0	0	391750	0	0	313280	791550	0	0	286560	112530
513.5	3	93.5	G	P	Gap	0	0	0	0	0	0	193090	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
517.5	3	97.5	E	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
518.5	3	98.5	R	G	Gap	0	0	509878	819812	0	0	0	0	0	0	0	0	0	0	605810	568650	0	0	0	148800	0	0	197893	122260
520.5	3	100.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
527.5	3	107.5	K	G	Gap	0	0	341745	819812	0	0	0	0	0	0	0	0	0	0	605810	568650	0	0	0	148800	0	0	761170	122260
528.5	3	108.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
529.5	3	109.5	S	P	Gap	0	0	0	0	0	0	0	118890	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
530.5	3	110.5	P	G	Gap	0	0	0	0	0	0	0	182640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
531.5	3	111.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
532.5	3	112.5	E	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
533.5	3	113.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
534.5	3	114.5	G	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
535.5	3	115.5	R	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
538.5	3	118.5	E	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
539.5	3	119.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
540.5	3	120.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
541.5	3	121.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
543.5	3	123.5	G	A	Gap	0	0	0	0	0	0	0	182640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
544.5	3	124.5	A	K	Gap	0	0	0	0	0	0	0	118890	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
545.5	3	125.5	K	G	Gap	0	0	167175	163310	0	0	0	0	0	0	0	362400	0	0	0	269640	0	0	668700	0	0	0	0	347360
546.5	3	126.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
547.5	3	127.5	L	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
549.5	3	129.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
550.5	3	130.5	S	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
557.5	3	137.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
558.5	3	138.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
559.5	3	139.5	K	T	Gap	0	602925	223327	163310	0	0	0	0	0	0	0	362400	0	241170	0	269640	0	0	893310	0	0	0	0	347360
560.5	3	140.5	T	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
561.5	3	141.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
562.5	3	142.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
563.5	3	143.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
564.5	3	144.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
566.5	3	146.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
567.5	3	147.5	G	Q	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
569.5	3	149.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
570.5	3	150.5	G	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
571.5	3	151.5	R	P	Gap	0	602925	561525	0	0	0	0	0	0	0	0	0	0	241170	0	0	0	0	224610	0	0	0	0	0
572.5	3	152.5	P	G	Gap	0	0	0	208475	0	0	0	0	0	0	0	833900	0	0	0	0	0	0	0	0	0	0	0	0
573.5	3	153.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
576.5	3	156.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
579.5	3	159.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
580.5	3	160.5	A	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
581.5	3	161.5	R	G	Gap	0	320505	718972	112997	0	0	0	0	0	0	0	833900	0	375380	0	0	0	644590	231820	0	0	262050	557690	443650
582.5	3	162.5	G	Q	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
583.5	3	163.5	Q	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
593.5	3	173.5	K	G	Gap	0	320505	718972	110912	0	0	0	0	0	0	0	0	0	375380	0	0	0	644590	231820	0	0	262050	557690	443650
595.5	3	175.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
605.5	3	185.5	R	G	Gap	667925	100969	290665	552730	211220	0	243500	312540	267170	226290	161830	123200	0	243500	260310	363500	0	102610	511580	688420	0	351400	228940	103580

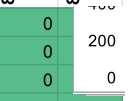
Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
606.5	3	186.5	G	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
607.5	3	187.5	V	P	Gap	0	204980	858040	190927	0	0	0	0	0	0	0	291480	0	819920	781540	138240	0	0	177780	236270	0	0	872820	360050
608.5	3	188.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
609.5	3	189.5	G	P	Gap	0	0	843700	365175	0	0	0	0	0	0	0	146070	0	0	0	0	0	0	337480	0	0	0	0	0
610.5	3	190.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
617.5	3	197.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
618.5	3	198.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
619.5	3	199.5	K	D	Gap	507262	298157	462962	119897	211220	611150	113179	573450	396390	115396	256619	240155	0	733359	310298	649104	683870	262420	742782	184104	948790	814530	542148	206557
620.5	3	200.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
621.5	3	201.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
624.5	3	204.5	G	A	Gap	0	0	0	362950	0	0	0	0	0	0	0	145180	0	0	0	0	0	0	0	0	0	0	0	0
626.5	3	206.5	Q	G	Gap	0	105090	0	278647	0	0	0	0	0	0	0	170570	0	420360	0	0	0	0	0	0	0	0	0	944020
627.5	3	207.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
628.5	3	208.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
630.5	3	210.5	G	P	Gap	0	0	279325	767270	0	174320	124250	240860	0	0	0	699300	0	0	0	610650	0	0	111730	0	0	0	0	238850
633.5	3	213.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
636.5	3	216.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
638.5	3	218.5	R	G	Gap	440470	196188	165666	706658	0	785470	101254	284996	129220	927670	947890	122730	0	485863	421730	332845	683870	159810	221223	112900	948790	463130	304480	124206
641.5	3	221.5	Q	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
645.5	3	225.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
647.5	3	227.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
652.5	3	232.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
653.5	3	233.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
654.5	6	0.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
658.5	6	4.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
660.5	6	6.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
661.5	6	7.5	E	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
664.5	6	10.5	K	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
665.5	6	11.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
666.5	6	12.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
667.5	6	13.5	E	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
668.5	6	14.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
669.5	6	15.5	G	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
670.5	6	16.5	V	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
671.5	6	17.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
672.5	6	18.5	G	D	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
673.5	6	19.5	D	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
674.5	6	20.5	L	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
675.5	6	21.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
683.5	6	29.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
686.5	6	32.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
692.5	6	38.5	R	G	Overlap	526584	201410	521697	100700	145500	691500	128260	537490	787690	620140	263210	267330	133280	507350	443830	909600	115890	171850	106460	182080	735390	644290	315150	103030
693.5	6	39.5	G	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
694.5	6	40.5	V	Q	Overlap	0	0	104385	744700	0	0	0	0	0	0	0	0	0	0	0	297880	0	0	417540	0	0	0	0	0
695.5	6	41.5	Q	G	Overlap	0	0	107692	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	430770	0	0	0	0	0
696.5	6	42.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
701.5	6	47.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
702.5	6	48.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
704.5	6	50.5	R	G	Overlap	526584	201410	523818	100775	145500	691500	128260	537490	787690	620140	263210	267330	133280	507350	443830	912578	115890	171850	107308	182080	735390	644290	315150	103030
707.5	6	53.5	N	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
716.5	6	62.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
717.5	6	63.5	G	D	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
718.5	6	64.5	D	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
719.5	6	65.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
720.5	6	66.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
721.5	6	67.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
726.5	6	72.5	G	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
729.5	6	75.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
730.5	6	76.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
731.5	6	77.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
732.5	6	78.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
740.5	6	86.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
741.5	6	87.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
742.5	6	88.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
743.5	6	89.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
745.5	6	91.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
749.5	6	95.5	K	G	Gap	0	0	247270	776682	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	310673	0	0	989080	0
751.5	6	97.5	D	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
752.5	6	98.5	R	G	Gap	0	230850	328275	200670	0	0	0	0	0	0	0	0	0	923400	0	0	0	0	131310	421740	0	0	0	380940
754.5	6	100.5	D	A	Gap	0	0	537615	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	215046	0
758.5	6	104.5	K	G	Gap	264710	484022	192136	168811	0	0	493350	932000	0	318570	208370	499880	0	116890	148130	271770	165740	665620	689900	461761	893100	697140	530590	181310
760.5	6	106.5	A	D	Gap	0	493600	153130	152153	0	0	0	0	0	0	0	251270	0	197440	0	155930	0	0	157490	342290	0	0	455030	248220
761.5	6	107.5	D	G	Gap	0	0	0	109639	0	0	0	0	0	0	0	0	0	0	0	268960	0	0	0	0	0	0	0	411660
766.5	6	112.5	K	D	Gap	0	230850	119092	414402	0	0	0	0	0	0	0	0	0	923400	0	0	0	0	131310	127667	0	0	345060	380940
770.5	6	116.5	R	G	Gap	264710	533382	214435	184560	0	0	493350	932000	0	318570	208370	525007	0	136634	148130	290052	165740	665620	847390	493739	893100	697140	604037	210248
771.5	6	117.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
772.5	6	118.5	L	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
774.5	6	120.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
776.5	6	122.5	I	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
777.5	6	123.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
778.5	6	124.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
779.5	6	125.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
780.5	6	126.5	G	P	Gap	0	0	0	568700	0	0	0	0	0	0	0	227480	0	0	0	0	0	0	0	0	0	0	0	0
781.5	6	127.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
782.5	6	128.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
783.5	6	129.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
784.5	6	130.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
785.5	6	131.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
786.5	6	132.5	G	D	Gap	0	0	0	0	110865	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
787.5	6	133.5	D	K	Gap	0	0	0	0	833574	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
788.5	6	134.5	K	G	Gap	0	0	0	0	250860	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
789.5	6	135.5	G	E	Gap	0	0	0	0	158627	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
790.5	6	136.5	E	T	Gap	0	0	0	0	284340	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
791.5	6	137.5	T	G	Gap	0	0	0	0	479930	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
792.5	6	138.5	G	P	Gap	0	0	0	0	404230	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
793.5	6	139.5	P	S	Gap	0	0	0	0	810000	445800	0	0	0	995020	128080	0	0	0	0	0	0	0	0	0	0	0	0	0
794.5	6	140.5	S	G	Gap	0	0	0	0	0	798572	204272	0	0	229210	0	0	0	0	0	0	0	0	0	0	0	0	0	0
795.5	6	141.5	G	P	Gap	0	0	0	0	870475	269215	268794	325940	0	0	213690	0	0	0	0	0	0	0	0	0	0	0	0	0
796.5	6	142.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
798.5	6	144.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
801.5	6	147.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
803.5	6	149.5	R	G	Gap	674655	515857	110117	422433	162515	335198	468876	408218	358390	115410	387413	537393	0	108862	627096	285526	847330	648538	171504	758997	149290	210856	167512	591469
806.5	6	152.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
808.5	6	154.5	D	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
809.5	6	155.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
812.5	6	158.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
813.5	6	159.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
816.5	6	162.5	G	P	Gap	0	0	822325	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	328930	0	0	0	0	0
818.5	6	164.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
819.5	6	165.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
820.5	6	166.5	F	A	Gap	0	480875	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192350	0	0	0	0	0
821.5	6	167.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
822.5	6	168.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
825.5	6	171.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
826.5	6	172.5	A	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
827.5	6	173.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
833.5	6	179.5	K	G	Gap	0	0	563107	100167	0	0	0	0	0	0	0	0	0	0	0	0	0	0	225243	0	0	0	0	400670
834.5	6	180.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
842.5	6	188.5	K	G	Gap	0	0	126760	100167	0	0	0	0	0	0	0	0	0	0	0	0	0	0	507040	0	0	0	0	400670
844.5	6	190.5	D	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
845.5	6	191.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
848.5	6	194.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
849.5	6	195.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
850.5	6	196.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
851.5	6	197.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
852.5	6	198.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
853.5	6	199.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
854.5	6	200.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
855.5	6	201.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
856.5	6	202.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
857.5	6	203.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
858.5	6	204.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
859.5	6	205.5	P	I	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
860.5	6	206.5	I	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
861.5	6	207.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
862.5	6	208.5	N	V	Gap	0	0	126760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	507040	0	0	0	0	0
863.5	6	209.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
864.5	6	210.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
865.5	6	211.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
867.5	6	213.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
869.5	6	215.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
872.5	6	218.5	R	G	Gap	0	0	681375	192027	0	0	0	0	0	0	0	0	0	0	0	768110	0	0	272550	0	0	0	0	0
873.5	6	219.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
874.5	6	220.5	S	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
875.5	6	221.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
876.5	6	222.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
877.5	6	223.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
878.5	6	224.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
879.5	6	225.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
880.5	6	226.5	A	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
881.5	6	227.5	T	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
887.5	6	233.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
888.5	9	0.5	G	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
889.5	9	1.5	R	V	Overlap	0	0	681375	192027	0	0	0	0	0	0	0	0	0	0	0	768110	0	0	272550	0	0	0	0	0
891.5	9	3.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
892.5	9	4.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
893.5	9	5.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
894.5	9	6.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
896.5	9	8.5	S	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
897.5	9	9.5	G	N	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
898.5	9	10.5	N	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
899.5	9	11.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
900.5	9	12.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
901.5	9	13.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
903.5	9	15.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
906.5	9	18.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
908.5	9	20.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
909.5	9	21.5	G	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
910.5	9	22.5	K	E	Overlap	0	0	681050	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	272420	0
914.5	9	26.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
916.5	9	28.5	P	R	Overlap	0	743625	0	0	0	0	0	0	0	0	0	0	0	297450	0	0	0	0	0	0	0	0	0	0
917.5	9	29.5	R	G	Overlap	646900	0	0	540523	0	0	0	0	258760	0	0	0	0	0	0	264200	0	0	0	478850	0	0	0	168060
925.5	9	37.5	R	A	Overlap	646900	743625	875772	540523	0	0	0	0	258760	0	0	0	0	297450	0	264200	0	0	778890	478850	0	0	272420	168060
926.5	9	38.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
927.5	9	39.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
928.5	9	40.5	E	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
930.5	9	42.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
931.5	9	43.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
933.5	9	45.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
934.5	9	46.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
936.5	9	48.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
938.5	9	50.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
939.5	9	51.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
940.5	9	52.5	E	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
941.5	9	53.5	K	G	Overlap	576545	304155	556649	0	0	0	416370	0	0	106170	637800	0	530680	642670	144950	0	0	395140	163688	0	177550	726420	380980	0
945.5	9	57.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
946.5	9	58.5	A	D	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
947.5	9	59.5	D	G	Overlap	0	0	855225	0	0	0	0	0	0	0	0	0	0	0	273380	0	0	0	844320	0	0	0	230320	0
948.5	9	60.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
949.5	9	61.5	P	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
950.5	9	62.5	A	G	Overlap	0	0	0	0	0	0	152810	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
951.5	9	63.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
952.5	9	64.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
953.5	9	65.5	P	G	Overlap	0	0	786000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	314400	0	0	0	0	0
954.5	9	66.5	G	T	Overlap	0	368970	669257	0	578690	0	468140	0	0	0	0	0	0	874600	811160	0	0	601280	109720	0	0	0	768670	0
955.5	9	67.5	T	P	Overlap	0	256425	388290	0	0	0	199410	0	0	0	0	0	0	102570	0	0	0	0	976330	0	0	0	576830	0
956.5	9	68.5	P	G	Overlap	0	0	317000	0	0	0	163320	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	126800	0
957.5	9	69.5	G	P	Overlap	0	121247	131425	0	0	718640	0	0	0	0	0	0	0	0	0	0	0	484990	248410	0	0	0	277290	0
959.5	9	71.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
960.5	9	72.5	G	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
962.5	9	74.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
963.5	9	75.5	G	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
965.5	9	77.5	R	G	Overlap	576545	517402	211448	217777	578690	718640	109443	0	0	106170	268680	296500	530680	832700	701884	223620	0	105809	633743	0	177550	726420	114994	617840
966.5	9	78.5	G	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
967.5	9	79.5	V	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
968.5	9	80.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
974.5	9	86.5	R	G	Gap	0	138582	121504	217777	0	0	0	0	0	0	204900	296500	0	0	448480	223620	0	554330	420680	0	0	0	0	617840
977.5	9	89.5	R	G	Gap	185552	149227	912710	259824	0	0	0	805970	0	0	358820	708780	0	596910	397400	285820	0	0	229210	253440	742210	0	602520	429160
978.5	9	90.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
979.5	9	91.5	F	P	Gap	0	0	742550	128932	0	0	0	0	0	0	0	0	0	0	0	245110	0	0	297020	0	0	0	0	270620
980.5	9	92.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
981.5	9	93.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
982.5	9	94.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
984.5	9	96.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
986.5	9	98.5	S	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
990.5	9	102.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
991.5	9	103.5	K	Q	Gap	185552	407455	191606	766463	0	0	364020	234699	0	111090	125098	101162	0	596910	619470	294307	0	810180	232180	158254	742210	111640	347202	117888
993.5	9	105.5	G	P	Gap	0	0	0	150012	0	0	0	0	0	0	0	0	0	0	0	892980	0	0	0	344820	0	0	0	165930
1001.5	9	113.5	R	G	Gap	0	258227	103171	742444	0	0	364020	226640	0	111090	892160	302840	0	0	241645	268743	0	810180	123560	156064	0	111640	286950	113735
1002.5	9	114.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1003.5	9	115.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1004.5	9	116.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1005.5	9	117.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1006.5	9	118.5	P	M	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1007.5	9	119.5	M	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1008.5	9	120.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1009.5	9	121.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1010.5	9	122.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
1011.5	9	123.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1012.5	9	124.5	L	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1013.5	9	125.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1014.5	9	126.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1017.5	9	129.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1018.5	9	130.5	E	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1019.5	9	131.5	S	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1020.5	9	132.5	G	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1021.5	9	133.5	R	E	Gap	782005	270639	482251	108526	405370	0	161890	435520	0	0	0	106790	0	796990	164365	767366	216150	248950	144776	182560	966520	366180	316880	164130
1023.5	9	135.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1024.5	9	136.5	A	P	Gap	0	0	0	867867	0	0	0	0	0	0	0	0	0	0	0	622770	0	0	0	163680	0	0	0	121190
1033.5	9	145.5	R	D	Gap	782005	270639	446467	108801	405370	0	161890	435520	0	0	0	106790	0	796990	144790	749887	216150	248950	132420	184196	966520	366180	316880	165341
1040.5	9	152.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1045.5	9	157.5	E	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1047.5	9	159.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1049.5	9	161.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1050.5	9	162.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1052.5	9	164.5	P	G	Gap	0	0	0	206387	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	825550
1053.5	9	165.5	G	A	Gap	0	0	161215	602425	147000	0	0	0	0	0	0	0	0	0	0	0	0	0	644860	0	0	0	0	240970
1054.5	9	166.5	A	P	Gap	0	111847	509245	254081	180960	0	0	0	0	0	0	348450	0	447390	517980	140430	0	0	151900	277290	0	0	0	563760
1055.5	9	167.5	P	G	Gap	0	100510	232277	413500	0	0	0	0	0	0	0	0	0	402040	0	0	0	0	929110	165400	0	0	0	0
1056.5	9	168.5	G	A	Gap	0	833075	107403	193445	399330	0	0	137280	0	0	0	415150	0	333230	349760	810280	0	0	949210	372310	0	0	299718	278930
1057.5	9	169.5	A	P	Gap	0	0	0	129340	0	0	421580	0	0	0	0	517360	0	0	0	0	0	0	0	0	0	0	0	0
1058.5	9	170.5	P	G	Gap	0	0	320325	806162	0	0	0	0	0	0	0	0	0	0	0	549750	0	0	128130	0	0	0	0	267490
1059.5	9	171.5	G	A	Gap	0	142212	361755	106507	399400	0	0	131750	0	0	0	0	0	568850	617060	107540	0	0	829960	0	0	0	0	318490
1060.5	9	172.5	A	P	Gap	0	357800	161602	499425	156670	0	0	764460	0	0	237190	199770	0	0	0	0	0	143120	409220	0	0	0	0	0
1062.5	9	174.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1066.5	9	178.5	P	A	Gap	0	0	641850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	256740	0
1068.5	9	180.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1069.5	**9**	**181.5**	**K**	**S**	Gap	403329	223414	386945	962950	118587	146373	371687	230105	300449	926971	375325	294138	103840	422193	295730	103875	110056	284180	546577	123953	108467	945860	667941	127936
1070.5	**9**	**182.5**	**S**	**G**	Gap	459275	0	0	0	0	0	691450	415120	0	0	0	0	0	0	0	0	0	0	0	0	183710	0	0	0
1071.5	**9**	**183.5**	**G**	**D**	Gap	0	444882	185004	114599	891515	0	206550	221290	0	635830	210670	0	0	0	946660	479880	0	114370	155300	210310	0	0	279380	200100
1072.5	**9**	**184.5**	**D**	**R**	Gap	0	414232	143879	392215	0	0	0	232680	0	662530	0	0	0	0	386700	217640	0	994400	376590	350320	0	0	499190	100090
1073.5	**9**	**185.5**	**R**	**G**	Gap	238127	118754	247549	566585	584658	956620	233017	170980	125810	612720	225274	100982	735290	216525	151319	538649	242190	145770	346528	882087	510980	514490	267076	744622
1074.5	**9**	**186.5**	**G**	**E**	Gap	0	716617	192572	155735	103840	0	117780	265880	0	459990	193070	0	0	160230	134340	379230	0	804180	200680	297970	0	0	242200	287050
1075.5	**9**	**187.5**	**E**	**T**	Gap	0	100840	308682	103444	263395	650310	246810	218760	0	252530	174420	144850	0	244070	283830	350340	0	990610	413730	200580	0	349760	362750	163680
1076.5	**9**	**188.5**	**T**	**G**	Gap	0	274510	102506	517424	395010	0	658820	933290	0	0	703620	425390	0	653590	803740	205160	0	444450	162910	806900	0	0	963810	101510
1077.5	**9**	**189.5**	**G**	**P**	Gap	0	233652	613630	197031	257100	183960	904720	526830	0	107050	564970	372050	0	435710	443990	713920	0	283390	773020	355470	0	108460	672540	324060

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
1078.5	**9**	**190.5**	**P**	**A**	Gap	598350	124211	185002	821111	195190	0	598460	432610	0	0	714400	0	0	338460	126200	852470	239340	106930	391080	200640	0	514560	151290	119280
1079.5	**9**	**191.5**	**A**	**G**	Gap	144098	159459	296062	792342	147670	936990	240120	158940	954240	444820	270800	158920	0	297700	251680	362750	480970	212300	408020	131070	0	833540	253750	133700
1080.5	**9**	**192.5**	**G**	**P**	Gap	0	127193	943422	309968	405630	240270	104887	462390	0	774690	852580	417880	0	434040	885510	211565	0	387900	113040	494300	0	0	905200	492220
1081.5	**9**	**193.5**	**P**	**A**	Gap	0	0	439625	0	0	0	0	437880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	175850	0
1082.5	**9**	**194.5**	**A**	**G**	Gap	0	175282	320017	165621	337502	0	295955	265301	0	0	826390	344230	0	377754	683340	600250	0	262349	562307	299077	0	610270	566790	268962
1083.5	**9**	**195.5**	**G**	**P**	Gap	0	215587	982275	442080	138580	160580	260452	653870	0	298610	839100	486068	0	703350	709940	237075	0	129140	114204	783360	0	0	123802	699280
1084.5	**9**	**196.5**	**P**	**I**	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1085.5	**9**	**197.5**	**I**	**G**	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1086.5	**9**	**198.5**	**G**	**P**	Gap	0	935572	263646	659220	120495	978020	360892	265979	0	455570	314580	535956	0	328672	219020	608775	0	0	194970	629860	0	0	326016	862290
1087.5	**9**	**199.5**	**P**	**V**	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1088.5	**9**	**200.5**	**V**	**G**	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1089.5	**9**	**201.5**	**G**	**A**	Gap	0	302260	140859	356177	123009	0	173771	195599	0	0	122738	182874	0	120904	101059	256832	0	0	160470	494380	0	0	179172	490624
1090.5	**9**	**202.5**	**A**	**R**	Gap	168140	418900	358335	149604	187652	0	190970	161600	818920	0	0	0	0	167560	499230	640350	415600	0	934110	321582	175070	0	0	212800
1091.5	**9**	**203.5**	**R**	**G**	Gap	112667	567854	118005	257842	277433	346931	767495	503469	619292	188752	522370	464534	252378	116875	939531	240176	137322	685897	158708	383639	226181	228011	167122	361102
1092.5	**9**	**204.5**	**G**	**P**	Gap	383317	129982	320440	492857	0	0	559400	583070	348370	0	149640	556810	0	336250	217290	619750	0	148160	503970	129600	118490	355210	410860	0
1093.5	9	205.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1100.5	9	212.5	R	G	Gap	480880	218335	505015	948067	170210	101650	144519	820490	195163	312370	246606	741078	750090	518922	469492	800187	0	244241	662485	152287	165334	789392	641478	139510
1101.5	9	213.5	G	D	Gap	0	0	106135	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	424540	0
1102.5	9	214.5	D	K	Gap	0	0	208727	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	834910	0
1103.5	9	215.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1105.5	9	217.5	E	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1112.5	9	224.5	R	G	Gap	0	0	0	987850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	255140	0	0	0	140000
1147.5	12	25.5	Q	G	Overlap	0	545126	147588	499385	0	0	598960	408980	0	304840	138140	300330	0	127090	476610	494130	0	630810	399530	469980	0	248310	129350	100340
1148.5	12	26.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1149.5	12	27.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1151.5	12	29.5	G	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1157.5	12	35.5	G	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1158.5	12	36.5	K	D	Overlap	0	570216	159007	544972	0	0	598960	408980	0	304840	180415	300330	0	137126	556488	547553	0	630810	438789	535580	0	248310	123550	106672
1159.5	12	37.5	D	G	Overlap	0	0	145012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	580050	0
1160.5	12	38.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1161.5	12	39.5	L	N	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1162.5	12	40.5	N	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1163.5	12	41.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1164.5	12	42.5	L	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1165.5	12	43.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1166.5	12	44.5	G	P	Overlap	0	0	0	184310	0	0	0	0	0	0	0	328190	0	0	0	409050	0	0	0	0	0	0	0	0
1169.5	12	47.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1172.5	12	50.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
1174.5	12	52.5	R	G	Overlap	0	250900	128685	474301	0	0	0	0	0	0	422750	328190	0	100360	798780	575135	0	0	392590	656000	0	0	0	633250
1218.5	12	96.5	K	A	C terminal	0	899100	963611	0	0	0	470270	0	0	0	962850	0	0	0	662800	0	0	242750	298360	0	0	116890	279700	0
1227.5	12	105.5	R	A	C terminal	0	899100	963611	0	0	0	470270	0	0	0	962850	0	0	0	662800	0	0	242750	298360	0	0	116890	279700	0

Appendix C, Table 2: Cleavage site information for COL1A1

Appendix C, Table 3

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
22.5	-1	-1	C	Q	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23.5	-1	-1	Q	S	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31.5	-1	-1	R	K	Telopeptide	0	0	118830	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	475320	0
37.5	-1	-1	D	R	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41.5	-1	-1	R	G	Telopeptide	0	0	118830	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	475320	0
44.5	-1	-1	R	G	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52.5	-1	-1	R	D	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
62.5	-1	-1	P	G	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
79.5	-1	-1	A	Q	Telopeptide	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81.5	-1	-1	F	D	N-terminal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
84.5	-1	-1	K	G	N-terminal	451400	730305	273665	232437	0	0	0	394920	0	0	0	0	0	778420	234150	153000	0	101120	521280	369750	180560	113160	339230	407000
85.5	-1	-1	G	G	N-terminal	0	656170	651522	222126	0	0	0	327550	0	0	0	614960	0	880810	562090	272690	0	789240	114780	0	0	954630	896200	615200
86.5	-1	-1	G	G	N-terminal	0	272050	102832	457420	0	0	0	621690	0	0	0	139280	0	0	135000	509230	0	0	156540	539380	0	108820	119790	779680
87.5	2	0.5	G	P	Overlap	0	0	977275	895651	0	0	0	788750	0	0	0	168400	0	0	140510	668220	0	0	125570	102060	0	0	124830	189210
88.5	2	1.5	P	G	Overlap	0	113772	351962	457831	0	0	0	607170	0	0	0	121400	0	0	214130	384360	0	0	708540	626890	0	455090	485180	818860
89.5	2	2.5	G	P	Overlap	0	0	167387	366797	0	0	0	239450	0	0	0	0	0	0	0	294800	0	0	0	342730	0	0	669550	829660
97.5	2	10.5	R	G	Overlap	451400	177229	243601	128686	0	0	0	169872	0	0	0	339016	0	165923	250088	127788	0	180044	413289	133957	180560	362952	311029	252661
110.5	2	23.5	G	F	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
111.5	2	24.5	F	Q	Overlap	0	0	160755	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	643020	0	0	0	0	0
112.5	2	25.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
113.5	2	26.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
116.5	2	29.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
117.5	2	30.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
118.5	2	31.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
119.5	2	32.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
120.5	2	33.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
121.5	2	34.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
130.5	2	43.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
133.5	2	46.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
134.5	2	47.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
138.5	2	51.5	K	A	Overlap	0	165495	132762	660751	0	942980	0	0	0	194560	0	254030	0	601090	153830	933040	0	0	252220	128750	0	414370	125000	123680
139.5	2	52.5	A	G	Overlap	0	0	885575	466075	0	0	0	0	0	0	0	0	0	0	0	186430	0	0	354230	0	0	0	0	0
140.5	2	53.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
141.5	2	54.5	E	D	Overlap	0	0	0	691400	0	0	0	0	0	0	0	0	0	0	0	276560	0	0	0	0	0	0	0	0
142.5	2	55.5	D	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
144.5	2	57.5	H	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
149.5	2	62.5	G	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
150.5	2	63.5	R	P	Overlap	0	165495	141618	672326	0	942980	0	0	0	194560	0	254030	0	601090	153830	139603	0	0	287643	128750	0	414370	125000	123680

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
152.5	2	65.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
154.5	2	67.5	R	G	Overlap	533902	768242	126369	812605	0	598460	501540	620740	505310	164200	614070	123370	0	170590	107050	369850	0	863770	173700	155400	163030	339100	163320	120320
157.5	2	70.5	V	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
163.5	2	76.5	R	G	Overlap	533902	768242	126369	812605	0	598460	501540	620740	505310	164200	614070	123370	0	170590	107050	369850	0	863770	173700	155400	163030	339100	163320	120320
165.5	2	78.5	F	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
167.5	2	80.5	G	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
174.5	2	87.5	F	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
175.5	2	88.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
178.5	2	91.5	R	G	Gap	0	151432	527045	783155	0	0	0	576740	0	0	141510	0	0	292540	895930	172720	0	184600	434880	122510	0	128590	144220	173480
187.5	2	100.5	K	G	Gap	0	151432	527045	783155	0	0	0	576740	0	0	141510	0	0	292540	895930	172720	0	184600	434880	122510	0	128590	144220	173480
192.5	2	105.5	A	P	Gap	0	0	668400	932500	0	0	0	0	0	0	267360	0	0	0	0	0	0	0	0	0	0	0	0	373000
196.5	2	109.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
197.5	2	110.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
198.5	2	111.5	E	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
200.5	2	113.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
201.5	2	114.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
203.5	2	116.5	G	E	Gap	0	0	668400	932500	0	0	0	0	0	0	267360	0	0	0	0	0	0	0	0	0	0	0	0	373000
204.5	2	117.5	E	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
205.5	2	118.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
206.5	2	119.5	G	T	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
214.5	2	127.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
220.5	2	133.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
222.5	2	135.5	R	V	Gap	0	597827	141255	255690	0	114600	244760	0	0	155570	0	0	0	159390	118510	477920	0	641840	359450	0	0	0	870610	544840
226.5	2	139.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
227.5	2	140.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
232.5	2	145.5	R	G	Gap	0	637074	161936	338995	0	114600	549962	401102	0	155570	189665	480621	0	175088	129550	564215	0	641840	398246	912190	0	0	100983	652484
242.5	2	155.5	G	P	Gap	0	0	177787	0	0	0	0	0	0	0	711150	0	0	0	0	0	0	0	0	0	0	0	0	0
244.5	2	157.5	A	G	Gap	0	0	0	373270	0	0	0	204260	0	0	0	864190	0	0	0	628890	0	0	0	0	0	0	0	0
245.5	2	158.5	G	P	Gap	0	144137	188675	393045	0	0	103690	804910	0	0	0	343400	0	576550	0	250550	0	0	754700	424760	0	0	0	553470
248.5	2	161.5	G	S	Gap	0	0	104016	101567	0	0	117490	193830	0	0	118550	0	0	0	110400	406270	0	0	103070	0	0	0	840440	0
249.5	2	162.5	S	A	Gap	0	844950	581962	0	0	0	584680	0	0	0	0	0	0	337980	0	0	0	0	177600	0	0	0	551850	0
250.5	2	163.5	A	G	Gap	0	754500	0	160967	0	0	0	506500	0	0	0	0	0	301800	0	0	0	0	0	294150	0	0	0	349720
251.5	2	164.5	G	P	Gap	0	883825	795600	140145	0	0	255540	443990	0	0	0	508020	0	353530	0	143250	0	0	318240	193280	0	0	0	173250
252.5	2	165.5	P	P	Gap	0	0	0	0	0	0	0	113060	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
254.5	2	167.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
255.5	2	168.5	F	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
257.5	2	170.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
262.5	2	175.5	K	G	Gap	0	103197	103157	127053	0	0	346550	895490	0	0	596540	0	0	163690	832150	799230	0	193790	269760	186380	0	553090	0	241910

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
263.5	2	176.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
264.5	2	177.5	E	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
265.5	2	178.5	L	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
266.5	2	179.5	G	P	Gap	0	0	651750	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	260700	0	0	0	0	0
267.5	2	180.5	P	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
268.5	2	181.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
269.5	2	182.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
270.5	2	183.5	N	P	Gap	0	280240	259747	135712	0	0	221160	371420	0	0	407500	521200	0	913940	292640	187710	0	0	575610	142760	0	207020	129990	160260
271.5	2	184.5	P	G	Gap	464775	773992	864920	204467	663900	0	538440	280070	185910	170470	629600	0	0	176180	538680	796085	0	116370	229140	115870	0	0	0	101990
272.5	2	185.5	G	P	Gap	0	160272	0	364147	0	0	177090	0	0	0	0	342990	0	394310	0	0	0	246780	0	576290	0	0	0	846000
280.5	2	193.5	R	G	Gap	0	368893	455914	572268	663900	0	287705	741039	0	170470	163364	864190	0	825574	429723	216657	0	556940	110058	853558	0	760110	129990	113244
281.5	2	194.5	G	E	Gap	0	0	0	198497	0	0	0	0	0	0	0	0	0	0	0	793990	0	0	0	0	0	0	0	0
283.5	2	196.5	V	G	Gap	464775	0	0	0	0	0	0	0	185910	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
287.5	2	200.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
288.5	2	201.5	L	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
290.5	2	203.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
293.5	2	206.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
294.5	2	207.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
296.5	2	209.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
297.5	2	210.5	N	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
299.5	2	212.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
300.5	2	213.5	A	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
301.5	2	214.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
302.5	2	215.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
303.5	2	216.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
307.5	2	220.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
308.5	2	221.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
309.5	2	222.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
310.5	2	223.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
311.5	2	224.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
312.5	2	225.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
313.5	2	226.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
314.5	2	227.5	G	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
323.5	5	2.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
325.5	5	4.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
326.5	5	5.5	G	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
327.5	5	6.5	I	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
328.5	5	7.5	P	G	Overlap	0	134647	278065	127294	129030	0	478110	494860	0	0	930520	949080	0	538590	240820	468990	0	0	508290	245270	0	0	270100	207520

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
329.5	5	8.5	G	P	Overlap	0	386388	150262	614350	262860	268050	796510	934840	0	746450	972280	142160	0	965190	187290	562690	0	505720	192850	911790	0	0	123680	840760
330.5	5	9.5	P	V	Overlap	0	598675	108793	140424	0	0	0	169390	0	0	193120	213570	0	0	0	121640	0	0	279450	345340	0	239470	136410	202060
331.5	5	10.5	V	G	Overlap	0	0	550800	0	0	0	0	0	0	0	220320	0	0	0	0	0	0	0	0	0	0	0	0	0
338.5	5	17.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
340.5	5	19.5	R	G	Overlap	0	527023	189498	882069	275763	268050	844321	100126	0	746450	110667	153786	0	150378	211372	621753	0	505720	271624	150240	0	239470	164331	125034
341.5	5	20.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
342.5	5	21.5	L	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
352.5	5	31.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
358.5	5	37.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
362.5	5	41.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
364.5	5	43.5	V	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
365.5	5	44.5	G	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
366.5	5	45.5	Q	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
368.5	5	47.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
370.5	5	49.5	P	G	Overlap	0	0	0	202417	0	0	0	0	0	0	0	0	0	0	0	796800	0	0	0	536200	0	0	0	193790
371.5	5	50.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
378.5	5	57.5	K	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
379.5	5	58.5	R	G	Overlap	0	930975	647850	487850	0	0	0	0	0	0	0	190150	0	372390	0	489550	0	0	259140	127170	0	0	0	0
380.5	5	59.5	G	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
381.5	5	60.5	S	T	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
382.5	5	61.5	T	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
383.5	5	62.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
384.5	5	63.5	E	I	Overlap	0	0	0	202417	0	0	0	0	0	0	0	0	0	0	0	796800	0	0	0	536200	0	0	0	193790
385.5	5	64.5	I	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
386.5	5	65.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
387.5	5	66.5	P	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
388.5	5	67.5	A	G	Overlap	0	0	0	317925	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	127170	0	0	0	0
389.5	5	68.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
392.5	5	71.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
395.5	5	74.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
396.5	5	75.5	L	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
397.5	5	76.5	R	G	Overlap	0	930975	647850	169925	0	0	0	0	0	0	0	190150	0	372390	0	489550	0	0	259140	0	0	0	0	0
398.5	5	77.5	G	N	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
399.5	5	78.5	N	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
400.5	5	79.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
403.5	5	82.5	R	G	Gap	0	0	403950	162200	0	0	0	0	0	0	0	648800	0	0	0	0	0	0	161580	0	0	0	0	0
409.5	5	88.5	D	G	Gap	0	0	246890	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	987560	0
411.5	5	90.5	R	A	Gap	235355	327085	205434	752979	0	0	0	372085	0	0	0	364710	0	447300	139990	717890	0	861040	340678	107570	941420	0	341070	118185

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
412.5	5	91.5	A	G	Gap	380975	0	147842	273127	0	0	0	0	152390	0	0	0	0	0	0	135360	0	0	591370	364210	0	0	0	592940
413.5	5	92.5	G	V	Gap	0	0	0	340122	0	0	0	132250	0	0	0	0	0	0	0	370380	0	0	0	215260	0	0	0	774850
421.5	5	100.5	R	G	Gap	470042	124793	421158	160568	409470	465220	646780	111838	764100	213160	110160	135473	174650	273190	357190	134220	0	171870	595761	250769	941420	327990	621522	243737
422.5	5	101.5	G	A	Gap	0	164065	216465	276063	0	0	0	193270	0	0	0	227230	0	656260	0	367470	0	0	0	362510	0	0	865860	351550
430.5	5	109.5	R	G	Gap	196590	922493	219081	858037	409470	465220	646780	750558	611710	213160	110160	105717	174650	229116	217200	622934	0	857660	270650	142982	0	327990	278315	127367
431.5	5	110.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
432.5	5	111.5	P	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
433.5	5	112.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
434.5	5	113.5	G	D	Gap	0	0	503675	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	201470	0
435.5	5	114.5	D	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
438.5	5	117.5	R	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
441.5	5	120.5	E	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
443.5	5	122.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
444.5	5	123.5	L	M	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
448.5	5	127.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
449.5	5	128.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
450.5	5	129.5	F	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
452.5	5	131.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
453.5	5	132.5	S	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
461.5	5	140.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
462.5	5	141.5	K	E	Gap	0	0	120985	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	483940	0	0	0	0	0
463.5	5	142.5	E	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
464.5	5	143.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
466.5	5	145.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
467.5	5	146.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
468.5	5	147.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
469.5	5	148.5	P	G	Gap	0	0	0	279287	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	608150	0	0	0	509000
470.5	5	149.5	G	I	Gap	954150	275875	151992	220408	0	0	148170	322200	0	0	291570	703780	0	110350	137410	493750	381660	0	0	303370	0	0	178990	458510
471.5	5	150.5	I	D	Gap	254875	0	0	480775	0	0	0	0	0	0	0	0	0	0	0	0	101950	0	0	192310	0	0	0	0
472.5	5	151.5	D	G	Gap	0	0	841142	911412	0	0	0	162610	0	0	0	115060	0	0	353810	378490	0	0	175360	174170	0	0	283540	141040
473.5	5	152.5	G	R	Gap	0	0	0	104838	0	0	0	0	0	0	0	219930	0	0	0	0	0	0	0	211090	0	0	0	186270
474.5	5	153.5	R	P	Gap	344472	178075	415288	110973	920910	126990	228040	186200	129250	709370	270490	182850	0	366540	384830	107440	393770	190400	591513	167440	854870	844230	414320	150730
475.5	5	154.5	P	G	Gap	0	116442	373497	482620	219840	0	220150	108370	0	0	0	586020	0	465770	440990	266760	0	0	661960	142590	0	0	391040	179220
476.5	5	155.5	G	P	Gap	0	0	246180	261592	0	0	0	460810	0	0	0	633170	0	0	0	889440	0	0	517410	461400	0	0	467310	432710
477.5	5	156.5	P	I	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
479.5	5	158.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
481.5	5	160.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
482.5	5	161.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
484.5	5	163.5	R	G	Gap	379501	178467	427934	117619	942894	126990	229741	192592	129250	709370	273405	198176	0	368109	390183	109468	533886	190400	598619	179266	854870	844230	449528	161927
486.5	5	165.5	E	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
493.5	5	172.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
496.5	5	175.5	K	G	Gap	0	0	0	520707	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	782230	0	0	0	130060
501.5	5	180.5	D	P	Gap	0	0	918225	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	367290	0
504.5	5	183.5	K	A	Gap	0	781872	724290	125756	0	0	501310	266650	0	0	267260	399240	0	262840	247860	386190	0	499090	277340	191373	0	0	210470	269040
508.5	5	187.5	K	G	Gap	272071	412410	478122	941920	273150	573940	172310	701760	695870	400040	187920	117980	771100	0	254140	0	214110	0	607130	207320	727480	124960	180570	157650
509.5	5	188.5	G	H	Gap	182502	593415	179164	305388	0	0	881850	133540	460280	0	740380	138250	0	159030	111470	103240	209760	607220	419880	660990	474220	176140	111270	443500
511.5	5	190.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
517.5	5	196.5	R	G	Gap	454573	426162	577650	979827	273150	573940	173693	717780	115615	400040	191332	119761	771100	421870	613470	141859	423870	110631	757933	215061	120170	125136	260131	163474
520.5	5	199.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
523.5	5	202.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
526.5	5	205.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
527.5	5	206.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
529.5	5	208.5	Q	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
531.5	5	210.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
532.5	5	211.5	P	G	Gap	0	0	931950	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	372780	0	0	0	0	0
541.5	5	220.5	K	G	Gap	0	0	931950	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	372780	0	0	0	0	0
544.5	5	223.5	Q	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
547.5	5	226.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
551.5	5	230.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
552.5	5	231.5	F	Q	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
553.5	5	232.5	Q	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
554.5	5	233.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
555.5	5	234.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
556.5	5	235.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
557.5	8	2.5	G	P	Overlap	0	0	0	572550	0	0	0	0	0	0	0	0	0	0	0	229020	0	0	0	0	0	0	0	0
558.5	8	3.5	P	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
559.5	8	4.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
567.5	8	12.5	K	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
571.5	8	16.5	R	G	Overlap	0	0	0	572550	0	0	0	0	0	0	0	0	0	0	0	229020	0	0	0	0	0	0	0	0
572.5	8	17.5	G	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
573.5	8	18.5	I	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
574.5	8	19.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
575.5	8	20.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
576.5	8	21.5	E	F	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
577.5	8	22.5	F	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
578.5	8	23.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
583.5	8	28.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
584.5	8	29.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
585.5	8	30.5	A	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
586.5	8	31.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
589.5	8	34.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
590.5	8	35.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
591.5	8	36.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
592.5	8	37.5	P	G	Overlap	0	321725	512870	601325	0	0	0	0	0	0	0	0	0	128690	0	0	0	0	118390	0	0	0	867580	240530
593.5	8	38.5	G	E	Overlap	0	703215	229935	556710	881800	0	191210	606970	0	180270	267900	244170	0	178350	160610	232960	0	849090	217430	983040	0	0	273800	766670
594.5	8	39.5	E	S	Overlap	0	121390	118020	196540	635440	0	675410	446040	0	0	926830	904780	0	0	107630	877840	0	485560	156150	607900	0	0	115620	0
595.5	8	40.5	S	G	Overlap	0	0	401237	123435	303640	0	0	271320	0	0	575850	404030	0	0	0	0	0	0	102910	253020	0	0	0	200320
596.5	8	41.5	G	A	Overlap	0	466740	251612	555820	140590	0	204780	885200	0	207590	169320	166830	0	0	163770	318750	0	128640	373920	0	0	372970	299440	173770
597.5	8	42.5	A	A	Overlap	0	123652	105697	398225	0	0	685990	308010	0	0	667150	480530	0	0	775530	0	0	494610	153000	824650	0	0	125520	720200
598.5	8	43.5	A	G	Overlap	0	556450	105796	505680	104760	0	635800	468390	0	0	524650	0	0	0	140640	0	0	0	230080	119510	0	222580	0	827620
599.5	8	44.5	G	P	Overlap	0	486275	550000	976132	0	0	0	127620	0	0	360680	242790	0	0	437670	426840	0	194510	720320	180310	0	0	681330	143180
601.5	8	46.5	T	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
602.5	8	47.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
603.5	8	48.5	P	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
605.5	8	50.5	G	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
606.5	8	51.5	S	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
607.5	8	52.5	R	G	Overlap	347900	514759	158862	337267	440387	129820	595710	426213	0	960660	106734	832724	139160	106100	108787	106633	0	749216	261682	566681	0	152754	158246	592484
608.5	8	53.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
612.5	8	57.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
614.5	8	59.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
618.5	8	63.5	N	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
619.5	8	64.5	K	G	Overlap	347900	248614	136155	0	129490	129820	0	0	0	572800	0	0	139160	600840	0	0	0	243140	544620	0	0	931990	0	0
620.5	8	65.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
622.5	8	67.5	P	G	Overlap	0	382800	0	0	0	0	0	0	0	0	0	0	0	153120	0	0	0	0	0	0	0	0	0	0
623.5	8	68.5	G	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
624.5	8	69.5	V	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
625.5	8	70.5	V	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
626.5	8	71.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
627.5	8	72.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
629.5	8	74.5	G	T	Overlap	0	0	0	354350	317340	0	0	685590	0	0	0	141740	0	0	0	0	0	0	0	0	0	0	0	0
630.5	8	75.5	T	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
631.5	8	76.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
632.5	8	77.5	G	P	Overlap	0	0	928000	0	144570	0	0	0	0	0	0	0	0	0	0	0	0	0	371200	0	0	0	0	0
634.5	8	79.5	S	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
635.5	8	80.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
643.5	8	88.5	R	G	Gap	0	0	928000	354350	461910	0	0	685590	0	0	0	141740	0	0	0	0	0	0	371200	0	0	0	0	0
645.5	8	90.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
652.5	8	97.5	K	G	Gap	0	0	0	126746	0	0	0	120070	0	0	0	0	0	0	0	264170	0	0	0	301690	0	0	0	178880
655.5	8	100.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
661.5	8	106.5	R	G	Gap	0	0	0	126746	0	0	0	120070	0	0	0	0	0	0	0	264170	0	0	0	301690	0	0	0	178880
662.5	8	107.5	G	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
663.5	8	108.5	D	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
664.5	8	109.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
669.5	8	114.5	R	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
670.5	8	115.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
673.5	8	118.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
680.5	8	125.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
681.5	8	126.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
682.5	8	127.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
683.5	8	128.5	G	P	Gap	0	0	312590	442175	0	0	634570	0	0	0	456880	176870	0	0	0	0	0	0	0	0	0	0	793480	0
684.5	8	129.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
686.5	8	131.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
687.5	8	132.5	A	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
688.5	8	133.5	N	G	Gap	0	0	187397	675620	0	0	0	135740	0	0	0	0	0	0	165430	275450	0	0	316240	152670	0	0	267920	900330
689.5	8	134.5	G	D	Gap	0	0	0	114288	0	0	0	289250	0	0	0	711520	0	0	0	0	0	0	0	386000	0	0	0	0
690.5	8	135.5	D	R	Gap	0	231275	728455	194756	0	0	0	0	0	0	0	253690	0	925100	0	936980	0	0	201220	393200	0	0	271260	266760
691.5	8	136.5	R	G	Gap	107503	517335	822577	208521	361633	429760	579680	145650	372750	178840	452110	200530	0	992790	356300	903220	162280	641210	137620	356280	230460	256500	110570	367430
692.5	8	137.5	G	E	Gap	0	135347	910720	345022	0	0	942380	328070	0	0	673980	0	0	0	471800	856090	0	541390	125120	846210	0	0	124590	448270
693.5	8	138.5	E	A	Gap	0	246867	372750	242113	362540	0	487640	227090	0	0	0	248970	0	987470	0	802380	0	0	0	434240	0	0	149100	429080
694.5	8	139.5	A	G	Gap	0	404247	108416	408642	680370	0	734800	256080	0	0	596050	750410	0	113900	729400	103770	0	477990	161190	874380	0	0	139930	581380
695.5	8	140.5	G	P	Gap	0	144471	401842	123166	168940	0	244110	115920	0	0	234250	207280	0	369160	398040	417570	0	158340	421210	226020	0	503870	553870	204160
696.5	8	141.5	P	A	Gap	0	377750	514072	127744	0	0	154160	526490	0	0	0	0	0	151100	371090	257290	0	0	978290	0	0	0	706910	485250
697.5	8	142.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
698.5	8	143.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
701.5	8	146.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
703.5	8	148.5	A	G	Gap	0	0	0	0	667440	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
704.5	8	149.5	G	P	Gap	0	0	0	0	150090	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
706.5	8	151.5	R	G	Gap	107503	545737	986289	241963	613181	429760	633626	182347	372750	178840	492804	242672	0	107533	428369	101142	162280	667237	150848	423489	230460	261538	151550	418956
708.5	8	153.5	S	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
712.5	8	157.5	R	G	Gap	0	0	438907	734575	0	0	907280	130220	0	0	0	293830	0	0	0	0	0	0	175563	0	0	0	0	0
717.5	8	162.5	P	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
721.5	8	166.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
722.5	8	167.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
723.5	8	168.5	F	A	Gap	0	0	276400	734575	0	0	713990	130220	0	0	0	293830	0	0	0	0	0	0	110560	0	0	0	0	0
724.5	8	169.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
725.5	8	170.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
728.5	8	173.5	G	A	Gap	0	0	0	0	0	0	193290	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
731.5	8	176.5	G	Q	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
736.5	8	181.5	K	G	Gap	0	0	162507	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	650030	0	0	0	0	0
739.5	8	184.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
742.5	8	187.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
745.5	8	190.5	K	G	Gap	0	0	494850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	197940	0	0	0	0	0
748.5	8	193.5	N	G	Gap	0	0	0	0	151350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
749.5	8	194.5	G	P	Gap	0	0	0	0	101210	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
752.5	8	197.5	G	P	Gap	0	0	0	0	154120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
754.5	8	199.5	T	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
755.5	8	200.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
756.5	8	201.5	P	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
757.5	8	202.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
758.5	8	203.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
759.5	8	204.5	A	A	Gap	0	0	494850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	197940	0	0	0	0	0
760.5	8	205.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
761.5	8	206.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
763.5	8	208.5	S	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
764.5	8	209.5	G	P	Gap	0	0	0	0	406680	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
766.5	8	211.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
775.5	8	220.5	R	G	Gap	0	0	251425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100570	0	0	0	0	0
776.5	8	221.5	G	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
777.5	8	222.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
786.5	8	231.5	F	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
791.5	11	2.5	G	R	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
792.5	11	3.5	R	T	Overlap	0	0	251425	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100570	0	0	0	0	0
793.5	11	4.5	T	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
794.5	11	5.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
795.5	11	6.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
796.5	11	7.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
797.5	11	8.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
798.5	11	9.5	P	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
799.5	11	10.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
800.5	11	11.5	G	I	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
801.5	11	12.5	I	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
806.5	11	17.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
809.5	11	20.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
812.5	11	23.5	G	K	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
813.5	11	24.5	K	E	Overlap	0	0	252925	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	101170	0
817.5	11	28.5	R	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
820.5	11	31.5	R	G	Overlap	0	0	220135	185634	0	858250	0	910840	0	0	106500	143070	0	0	189580	910090	0	0	0	637220	0	0	584460	0
828.5	11	39.5	R	T	Overlap	0	0	222664	185634	0	858250	0	910840	0	0	106500	143070	0	0	189580	910090	0	0	0	637220	0	0	594577	0
830.5	11	41.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
831.5	11	42.5	E	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
833.5	11	44.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
835.5	11	46.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
836.5	11	47.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
840.5	11	51.5	F	V	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
842.5	11	53.5	G	E	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
844.5	11	55.5	K	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
845.5	11	56.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
851.5	11	62.5	G	T	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
852.5	11	63.5	T	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
853.5	11	64.5	A	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
854.5	11	65.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
855.5	11	66.5	P	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
856.5	11	67.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
857.5	11	68.5	G	T	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
858.5	11	69.5	T	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
859.5	11	70.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
860.5	11	71.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
861.5	11	72.5	P	Q	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
862.5	11	73.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
863.5	11	74.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
864.5	11	75.5	L	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
865.5	11	76.5	L	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
866.5	11	77.5	G	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
867.5	11	78.5	A	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
868.5	11	79.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
869.5	11	80.5	G	F	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
876.5	11	87.5	S	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
877.5	11	88.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
880.5	11	91.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
881.5	11	92.5	G	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
882.5	11	93.5	L	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
883.5	11	94.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
884.5	11	95.5	G	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
885.5	11	96.5	V	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
886.5	11	97.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
887.5	11	98.5	G	S	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
888.5	11	99.5	S	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
889.5	11	100.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
890.5	11	101.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
891.5	11	102.5	E	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
892.5	11	103.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
893.5	11	104.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
894.5	11	105.5	P	L	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
895.5	11	106.5	L	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
896.5	11	107.5	G	I	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
897.5	11	108.5	I	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
898.5	11	109.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
899.5	11	110.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
900.5	11	111.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
902.5	11	113.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
903.5	11	114.5	A	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
904.5	11	115.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
906.5	11	117.5	P	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
908.5	11	119.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
909.5	11	120.5	N	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
911.5	11	122.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
912.5	11	123.5	N	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
913.5	11	124.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
914.5	11	125.5	G	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
916.5	11	127.5	N	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
924.5	11	135.5	R	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
925.5	11	136.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
926.5	11	137.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
927.5	11	138.5	N	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
931.5	11	142.5	D	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
936.5	11	147.5	R	D	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
943.5	11	154.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
946.5	11	157.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
948.5	11	159.5	Y	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
949.5	11	160.5	P	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
950.5	11	161.5	G	N	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
951.5	11	162.5	N	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
952.5	11	163.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
953.5	11	164.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
954.5	11	165.5	P	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
955.5	11	166.5	V	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
956.5	11	167.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
957.5	11	168.5	A	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
958.5	11	169.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
959.5	11	170.5	G	A	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
960.5	11	171.5	A	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
961.5	11	172.5	P	G	Gap	0	232955	445972	199210	0	0	0	0	0	0	0	0	0	931820	396790	0	0	0	138710	391410	0	0	0	405430
962.5	11	173.5	G	P	Gap	0	124020	430925	125679	220770	150980	0	103660	0	400220	248810	189260	0	280920	418880	616030	0	175140	602780	236120	0	0	453230	186070
964.5	11	175.5	Q	G	Gap	0	0	0	0	0	0	0	237290	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
965.5	11	176.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
968.5	11	179.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
971.5	11	182.5	G	K	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
972.5	11	183.5	K	H	Gap	0	147316	475522	145600	220770	150980	0	127389	0	400220	248810	189260	0	374102	458559	616030	0	175140	741490	275261	0	0	453230	226613
975.5	11	186.5	S	R	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
976.5	11	187.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
977.5	11	188.5	G	E	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
978.5	11	189.5	E	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
979.5	11	190.5	P	G	Gap	0	359100	740652	347814	879330	0	581510	241340	0	0	465990	129530	0	143640	512290	665750	0	0	133210	759470	0	0	652230	552260
980.5	11	191.5	G	P	Gap	0	0	811325	633375	225048	136490	0	0	0	0	324530	0	0	0	0	0	0	0	0	253350	0	0	0	0
981.5	11	192.5	P	V	Gap	568400	417265	563257	422392	585750	0	248980	172630	0	0	296160	824890	227360	115680	665770	0	0	512260	129110	120310	0	0	0	478220
982.5	11	193.5	V	G	Gap	0	444375	735850	250974	0	0	0	140240	0	0	0	0	0	177750	0	357860	0	0	294340	616110	0	0	0	352000
983.5	11	194.5	G	A	Gap	0	766805	466100	955247	218160	881240	247940	165310	0	263560	0	302590	0	153360	0	325800	0	101460	0	156430	0	255460	186440	162830
984.5	11	195.5	A	V	Gap	151882	128427	135770	382659	888610	0	621880	595130	0	0	101800	734060	0	0	0	271390	607530	513710	307070	725910	0	0	134210	459930
985.5	11	196.5	V	G	Gap	0	108596	141582	459893	0	0	0	126480	0	0	0	0	0	291590	0	944750	0	870770	407930	174510	0	557170	158400	0
986.5	11	197.5	G	P	Gap	0	143880	442777	108297	0	0	179820	0	0	0	204470	155990	0	358760	490460	102770	0	216760	597690	143370	0	0	478490	171450
988.5	11	199.5	A	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
989.5	11	200.5	G	A	Gap	0	0	0	0	163980	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
990.5	11	201.5	A	V	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
992.5	11	203.5	G	P	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
994.5	11	205.5	R	G	Gap	161564	136917	259527	305697	596641	128696	135589	576369	759130	368135	239650	861688	227360	291527	221542	320677	209503	167218	320153	433531	358570	521116	256765	382412
995.5	11	206.5	G	P	Gap	0	390765	724910	109975	0	0	413000	0	0	0	0	304700	0	952370	882270	119460	0	610690	123220	156260	0	0	785170	133710
997.5	11	208.5	S	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1003.5	11	214.5	R	G	Gap	145808	132668	247867	261444	663200	126450	130273	547080	759130	365500	232580	809417	0	280072	216342	303654	148750	162750	302692	329282	358570	512990	239855	331897
1004.5	11	215.5	G	D	Gap	0	0	710500	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	284200	0
1005.5	11	216.5	D	K	Gap	0	0	115637	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	462550	0
1006.5	11	217.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1012.5	11	223.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1015.5	11	226.5	R	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1021.5	11	232.5	K	G	Gap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1027.5	14	4.5	Q	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1028.5	14	5.5	G	L	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1029.5	14	6.5	L	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1030.5	14	7.5	P	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1031.5	14	8.5	G	L	Overlap	0	0	0	207792	0	0	0	0	0	0	0	0	0	0	0	101970	0	0	0	365260	0	0	0	363940
1032.5	14	9.5	L	A	Overlap	0	0	400900	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	160360	0	0	0	0	0
1033.5	14	10.5	A	G	Overlap	0	0	110355	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	441420	0	0	0	0	0
1034.5	14	11.5	G	H	Overlap	0	0	0	154110	0	0	0	0	0	0	0	265100	0	0	0	129090	0	0	0	0	0	0	0	460840
1035.5	14	12.5	H	H	Overlap	0	0	115702	805140	0	0	0	0	0	0	0	218560	0	0	0	152940	0	0	462810	147260	0	0	0	0
1036.5	14	13.5	H	G	Overlap	0	0	191752	912972	0	0	0	0	0	0	0	734690	0	0	0	132940	0	0	767010	0	0	0	0	158780
1037.5	14	14.5	G	D	Overlap	0	0	0	200894	0	0	0	101560	0	0	0	997860	0	0	0	734800	0	0	0	313800	0	0	0	316510
1038.5	14	15.5	D	Q	Overlap	0	0	422427	128452	0	0	0	0	0	0	0	445300	0	0	0	153240	0	0	932960	175650	0	0	756750	140390
1039.5	14	16.5	Q	G	Overlap	0	0	0	235967	0	0	0	0	0	0	0	943870	0	0	0	0	0	0	0	0	0	0	0	0
1040.5	14	17.5	G	A	Overlap	0	0	0	312425	0	0	0	0	0	0	0	157170	0	0	0	224040	0	0	0	405860	0	0	0	462630
1041.5	14	18.5	A	P	Overlap	0	0	241860	407150	162830	0	0	261290	0	0	0	157430	0	0	392640	211260	0	0	574800	676010	0	0	0	583900
1042.5	14	19.5	P	G	Overlap	0	0	324150	355250	0	0	0	120950	0	0	0	0	0	0	0	0	0	0	129660	0	0	0	0	142100
1043.5	14	20.5	G	A	Overlap	0	228677	452325	224283	0	0	763560	167380	0	0	0	779220	0	0	0	153430	0	914710	180930	294960	0	0	0	370820
1045.5	14	22.5	V	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1051.5	14	28.5	R	G	Overlap	269340	148459	328462	105843	834463	853390	155744	972232	101370	308334	137601	154393	0	333754	341787	777993	301530	172484	458145	174986	674460	567650	376316	155146
1052.5	14	29.5	G	P	Overlap	616967	787250	127181	239259	0	0	527480	548090	0	0	648170	915680	0	177340	0	286490	178460	951040	201350	0	683270	424560	242560	578980
1053.5	14	30.5	P	A	Overlap	163427	486225	188752	177062	217260	245090	267910	128770	0	194490	169000	138930	653710	0	237620	163770	0	0	317800	487760	0	0	305910	428270
1055.5	14	32.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1060.5	14	37.5	A	G	Overlap	0	0	423000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	169200	0	0	0	0	0
1062.5	14	39.5	K	D	Overlap	277143	147569	319012	102491	103544	877899	156356	965217	101370	308234	139378	139174	653710	331243	338606	766141	319376	171021	445061	172987	681292	571895	353005	146447
1063.5	14	40.5	D	G	Overlap	0	0	331165	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	132466	0
1065.5	14	42.5	R	T	Overlap	125012	779923	175645	209442	0	510850	106410	244656	402100	151814	632534	393897	299110	149814	138014	131991	186900	113719	315642	344713	243030	332540	185671	321675
1066.5	14	43.5	T	G	Overlap	236247	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	944990	0	0	0	0	0	0	0
1067.5	14	44.5	G	Q	Overlap	0	0	540225	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	216090	0	0	0	0	0

Position	Row (QStagger)	Position (QStagger)	AA Before	AA After	Region	NonStillborn Avg Time 0	NonStillborn Avg Time 1	NonStillborn Avg Time 2	NonStillborn Avg Time 3	11_0_intensity	11_1_intensity	11_2_intensity	11_3_intensity	12_0_intensity	12_1_intensity	12_2_intensity	12_3_intensity	15_0_intensity	15_1_intensity	15_2_intensity	15_3_intensity	6_0_intensity	6_1_intensity	6_2_intensity	6_3_intensity	8_0_intensity	8_1_intensity	8_2_intensity	8_3_intensity
1068.5	14	45.5	Q	P	Overlap	0	235636	152819	814436	0	0	0	296090	0	0	0	190740	0	531400	0	127370	0	334420	568210	173130	0	767260	430670	138000
1069.5	14	46.5	P	G	Overlap	0	255050	411345	209107	558640	0	0	650170	0	0	0	0	0	102020	348140	314310	0	0	818070	528720	0	0	479170	276280
1070.5	14	47.5	G	A	Overlap	0	0	0	491950	0	0	0	0	0	0	0	0	0	0	0	196780	0	0	0	0	0	0	0	0
1078.5	14	55.5	R	G	Overlap	148637	769859	171684	208599	558640	510850	104580	244471	402100	149770	626920	366007	299110	147081	133198	131585	281399	112554	316306	348880	243030	333307	174542	317332
1085.5	14	62.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1086.5	14	63.5	P	A	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1088.5	14	65.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1097.5	14	74.5	G	P	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1101.5	14	78.5	P	S	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1102.5	14	79.5	S	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1103.5	14	80.5	G	G	Overlap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1104.5	14	81.5	G	G	C-terminal	0	0	172295	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	689180	0
1107.5	14	84.5	D	F	C-terminal	0	0	361225	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	144490	0
1116.5	14	93.5	R	A	C-terminal	0	0	533520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	213408	0

Appendix C, Table 3: Cleavage site information for COL1A2

References

Abelson, P. H. (1954). Amino acids in fossils. *Science*, *119*(3096), 576–576.

Ahmed, H. E., & Darwish, S. S. (2012). Effect of museum conditions on historical dyed silk fabric with madder dye. *Journal of Polymers and the Environment*, *20*(2), 596–606.

Al-Ammar, A., Drummond, J. L., & Bedran-Russo, A. K. (2009). The use of collagen cross-linking agents to enhance dentin bond strength. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials*, *91*(1), 419–424.

Alexander, J. J. G. (1986). *The York Gospels: a facsimile with introductory essays*. presentation to the members of the Roxburghe Club.

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, P. F., Samaniego, J. A., Gilbert, M. T. P., Willerslev, E., Zhang, G., Scofield, R. P., Holdaway, R. N., & Bunce, M. (2012). The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings. Biological Sciences / The Royal Society*, *279*(1748), 4724–4733.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Arrhenius, S. (1889). Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte. *Zeitschrift Für Physikalische Chemie*, *4*(1), 96–116.

Asara, J. M., Schweitzer, M. H., Freimark, L. M., Phillips, M., & Cantley, L. C. (2007). Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry. *Science*, *316*(5822), 280–285.

Axelsson, K. M., Larsen, R., Sommer, D. V. P., & Melin, R. (2016). Degradation of collagen in parchment under the influence of heat-induced oxidation: Preliminary study of changes at macroscopic, microscopic, and molecular levels. *Studies in Conservation*, *61*(1), 46–57.

Bada, J. L. (1972). The dating of fossil bones using the racemization of isoleucine. *Earth and Planetary Science Letters*, *15*(3), 223–231.

Bada, J. L. (1985). AMINO ACID RACEMIZATION DATING OF FOSSIL BONES. *Annual Review of Earth and Planetary Sciences*, *13*(1), 241–268.

Bansal, A. K., Shetty, D. C., Bindal, R., & Pathak, A. (2012). Amelogenin: A novel protein with diverse applications in genetic and molecular profiling. *Journal of Oral and Maxillofacial Pathology: JOMFP*, *16*(3), 395–399.

Barbieri, R., Mekni, R., Levasseur, A., Chabrière, E., Signoli, M., Tzortzis, S., Aboudharam, G., & Drancourt, M. (2017). Paleoproteomics of the Dental Pulp: The plague paradigm. *PloS One*, *12*(7), e0180552.

Barton, H. (2007). Starch residues on museum artefacts: implications for determining tool use. *Journal of Archaeological Science*, *34*(10), 1752–1762.

Beckerman, S. M. (2015). *Corded Ware Coastal Communities: Using ceramic analysis to reconstruct third millennium BC societies in the Netherlands*. Sidestone Press.

Bell, A. W., Deutsch, E. W., Au, C. E., Kearney, R. E., Beavis, R., Sechi, S., Nilsson, T., Bergeron, J. J. M., & HUPO Test Sample Working Group. (2009). A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nature Methods*, *6*(6), 423–430.

Bello, S. M., Parfitt, S. A., & Stringer, C. B. (2011). Earliest directly-dated human skull-cups. *PloS One*, *6*(2), e17026.

Berlett, B. S., & Stadtman, E. R. (1997). Protein oxidation in aging, disease, and oxidative stress. *The Journal of Biological Chemistry*, *272*(33), 20313–20316.

Berman, H. M. (2000). The Protein Data Bank. In *Nucleic Acids Research* (Vol. 28, Issue 1, pp. 235–242). https://doi.org/10.1093/nar/28.1.235

Berrill, A., Biddlecombe, J., & Bracewell, D. (2011). Chapter 13 - Product Quality During Manufacture and Supply. In C. Van Der Walle (Ed.), *Peptide and Protein Delivery* (pp. 313–339). Academic Press.

Bienkiewicz, K. J. (1983). *Physical chemistry of leather making*.

Binder, M., Roberts, C., Spencer, N., Antoine, D., & Cartwright, C. (2014). On the Antiquity of Cancer: Evidence for Metastatic Carcinoma in a Young Man from Ancient Nubia (c. 1200BC). *PloS One*, *9*(3), e90924.

Blakely, R. L., & Beck, L. (1984). TOOTH-TOOL USE VERSUS DENTAL MUTILATION: A CASE STUDY FROM THE PREHISTORIC SOUTHEAST. *Midcontinental Journal of Archaeology, MCJA*, *9*(2), 269–284.

Bleasdale, M., Richter, K. K., Janzen, A., Brown, S., Scott, A., Zech, J., Wilkin, S., Wang, K., Schiffels, S., Desideri, J., Besse, M., Reinold, J., Saad, M., Babiker, H., Power, R. C., Ndiema, E., Ogola, C., Manthi, F. K., Zahir, M., … Boivin, N. (2021). Ancient proteins provide evidence of dairy consumption in eastern Africa. *Nature Communications*, *12*(1), 632.

Bleicher, N., Kelstrup, C., Olsen, J. V., & Cappellini, E. (2015). Molecular evidence of use of hide glue in 4th millennium BC Europe. *Journal of Archaeological Science*, *63*, 65–71.
Bradfield, J., Forssman, T., Spindler, L., & Antonites, A. R. (2019). Identifying the animal species used to manufacture bone arrowheads in South Africa. *Archaeological and Anthropological Sciences*, *11*(6), 2419–2434.

Brandt, L. Ø., Schmidt, A. L., Mannering, U., Sarret, M., Kelstrup, C. D., Olsen, J. V., & Cappellini, E. (2014). Species identification of archaeological skin objects from Danish bogs: comparison between mass spectrometry-based peptide sequencing and microscopy-based methods. *PloS One*, *9*(9), e106875.

Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M., & Paabo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. In *Proceedings of the National Academy of Sciences* (Vol. 104, Issue 37, pp. 14616–14621). https://doi.org/10.1073/pnas.0704665104

Bringans, S. D., Dyer, J. M., Plowman, J. E., & Bryson, W. G. (2006). Kynurenine Located within Keratin Proteins Isolated from Photoyellowed Wool Fabric. *Textile Research Journal*, *76*(4), 288–294.

Buckley, M., & 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), e1–e1.

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

Buckley, M., Fraser, S., Herman, J., Melton, N. D., Mulville, J., & Pálsdóttir, A. H. (2014). Species identification of archaeological marine mammals using collagen fingerprinting. *Journal of Archaeological Science*, *41*, 631–641.

Buckley, M., Lawless, C., & Rybczynski, N. (2019). Collagen sequence analysis of fossil camels, Camelops and c.f. Paracamelus, from the Arctic and sub-Arctic of Plio-Pleistocene North America. *Journal of Proteomics*, *194*, 218–225.

Buckley, M., & Melton, N. D. (2013). Proteomics analysis of ancient food vessel stitching reveals> 4000-year-old milk protein. *Rapid Communications in Mass Spectrometry: RCM*. https://onlinelibrary.wiley.com/doi/abs/10.1002/rcm.6481

Buckley, M., Walker, A., Ho, S. Y. W., Yang, Y., Smith, C., Ashton, P., Oates, J. T., Cappellini, E., Koon, H., Penkman, K., Elsworth, B., Ashford, D., Solazzo, C., Andrews, P., Strahler, J., Shapiro, B., Ostrom, P., Gandhi, H., Miller, W., … Collins, M. J. (2008). Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry" [Review of *Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry"*]. *Science*, *319*(5859), 33; author reply 33.

Buckley, M., Whitcher Kansa, S., Howard, S., Campbell, S., Thomas-Oates, J., & Collins, M. (2010). Distinguishing between archaeological sheep and goat bones using a single collagen peptide. *Journal of Archaeological Science*, *37*(1), 13–20.

Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., & Zahedi, R. P. (2012). Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *Journal of Proteomics*, *75*(4), 1454–1462.

Capasso, S., Mazzarella, L., Sorrentino, G., Balboni, G., & Kirby, A. J. (1996). Kinetics and mechanism of the cleavage of the peptide bond next to asparagine. *Peptides*, *17*(6), 1075–1077.

Cappellini, E., Collins, M. J., & Gilbert, M. T. P. (2014). Biochemistry. Unlocking ancient protein palimpsests. *Science*, *343*(6177), 1320–1322.

Cappellini, E., Gentry, A., Palkopoulou, E., Ishida, Y., Cram, D., Roos, A.-M., Watson, M., Johansson, U. S., Fernholm, B., Agnelli, P., Barbagli, F., Littlewood, D. T. J., Kelstrup, C. D., Olsen, J. V., Lister, A. M., Roca, A. L., Dalén, L., & Gilbert, M. T. P. (2014). Resolution of the type material of the Asian elephant, Elephas maximus Linnaeus, 1758 (Proboscidea, Elephantidae). *Zoological Journal of the Linnean Society*, *170*(1), 222–232.

Cappellini, E., Gilbert, M. T. P., Geuna, F., Fiorentino, G., Hall, A., Thomas-Oates, J., Ashton, P. D., Ashford, D. A., Arthur, P., Campos, P. F., Kool, J., Willerslev, E., & Collins, M. J. (2010). A multidisciplinary study of archaeological grape seeds. *Die Naturwissenschaften*, *97*(2), 205–217.

Cappellini, E., Jensen, L. J., Szklarczyk, D., Ginolhac, A., da Fonseca, R. A. R., Stafford, T. W., Holen, S. R., Collins, M. J., Orlando, L., Willerslev, E., Gilbert, M. T. P., & Olsen, J. V. (2012). Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *Journal of Proteome Research*, *11*(2), 917–926.

Cappellini, E., Prohaska, A., Racimo, F., Welker, F., Pedersen, M. W., Allentoft, M. E., de Barros Damgaard, P., Gutenbrunner, P., Dunne, J., Hammann, S., Roffet-Salque, M., Ilardo, M., Moreno-Mayar, J. V., Wang, Y., Sikora, M., Vinner, L., Cox, J., Evershed, R. P., & Willerslev, E. (2018). Ancient Biomolecules and Evolutionary Inference. *Annual Review of Biochemistry*, *87*, 1029–1060.

Cappellini, E., Welker, F., Pandolfi, L., Ramos-Madrigal, J., Samodova, D., Rüther, P. L., Fotakis, A. K., Lyon, D., Moreno-Mayar, J. V., Bukhsianidze, M., Rakownikow Jersie-Christensen, R., Mackie, M., Ginolhac, A., Ferring, R., Tappen, M., Palkopoulou, E., Dickinson, M. R., Stafford, T. W., Jr, Chan, Y. L., … Willerslev, E. (2019). Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny. *Nature*, *574*(7776), 103–107.

Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., Nesvizhskii, A., & Working Group on Publication Guidelines for Peptide and Protein Identification Data. (2004). The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Molecular & Cellular Proteomics: MCP*, *3*(6), 531–533.

Catak, S., Monard, G., Aviyente, V., & Ruiz-López, M. F. (2009). Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms. *The Journal of Physical Chemistry. A*, *113*(6), 1111–1120.

Cecil Curwen, E. (1934). Excavations in Whitehawk Neolithic Camp, Brighton, 1932–31. *The Antiquaries Journal*, *14*(2), 99–133.

Cersoy, S., Zirah, S., Marie, A., & Zazzo, A. (2019). Toward a versatile protocol for radiocarbon and proteomics analysis of ancient collagen. *Journal of Archaeological Science*, *101*, 1–10.

Charlton, S., Ramsøe, A., Collins, M., Craig, O. E., Fischer, R., Alexander, M., & Speller, C. F. (2019). New insights into Neolithic milk consumption through proteomic analysis of dental calculus. *Archaeological and Anthropological Sciences*. https://doi.org/10.1007/s12520-019-00911-7

Chen, T., Dewhirst, F. E., Paster, B. J., Tanner, A., & Wade, W. (n.d.). *HOMD: Human Oral Microbiome Database [Internet]. All Human Oral Microbial Taxa. 2017 [cited 2017 Nov 25]*.
Chow, W. Y., Forman, C. J., Bihan, D., Puszkarska, A. M., Rajan, R., Reid, D. G., Slatter, D. A., Colwell, L. J., Wales, D. J., Farndale, R. W., & Duer, M. J. (2018). Proline provides site-specific flexibility for in vivo collagen. *Scientific Reports*, *8*(1), 13809.

Chung, L., Dinakarpandian, D., Yoshida, N., Lauer-Fields, J. L., Fields, G. B., Visse, R., & Nagase, H. (2004). Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *The EMBO Journal*, *23*(15), 3020–3030.

Clark, G. (1954). *Excavations at Star Carr: An Early Mesolithic Site at Seamer Near Scarborough, Yorkshire*. CUP Archive.

Clarkson, C. (1992). REDISCOVERING PARCHMENT: THE NATURE OF THE BEAST. *Paper Conservator*, *16*(1), 5–26.

Cleland, T. P. (2018a). Solid Digestion of Demineralized Bone as a Method To Access Potentially Insoluble Proteins and Post-Translational Modifications. *Journal of Proteome Research*, *17*(1), 536–542.

Cleland, T. P. (2018b). Human Bone Paleoproteomics Utilizing the Single-Pot, Solid-Phase-Enhanced Sample Preparation Method to Maximize Detected Proteins and Reduce Humics. *Journal of Proteome Research*, *17*(11), 3976–3983.

Cleland, T. P., & Vashishth, D. (2015). Bone protein extraction without demineralization using principles from hydroxyapatite chromatography. *Analytical Biochemistry*, *472*, 62–66.

Colleary, C., Lamadrid, H. M., O'Reilly, S. S., Dolocan, A., & Nesbitt, S. J. (2021). Molecular preservation in mammoth bone and variation based on burial environment. *Scientific Reports*, *11*(1), 2662.

Collins, M. J., Gernaey, A. M., Nielsen-Marsh, C. M., Vermeer, C., & Westbroek, P. (2000). Slow rates of degradation of osteocalcin: Green light for fossil bone protein? In *Geology* (Vol. 28, Issue 12, pp. 1139–1142). https://doi.org/10.1130/0091-7613(2000)028<1139:srodoo>2.3.co;2

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

Collins, M. J., & Riley, M. S. (2000). Amino acid racemization in biominerals: the impact of protein degradation and loss. *Oxford University Press*.
Collins, M. J., Riley, M. S., Child, A. M., & Turner-Walker, G. (1995). A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. *Journal of Archaeological Science*, *22*(2), 175–183.

Colonese, A. C., Hendy, J., Lucquin, A., Speller, C. F., Collins, M. J., Carrer, F., Gubler, R., Kühn, M., Fischer, R., & Craig, O. E. (2017). New criteria for the molecular identification of cereal grains associated with archaeological artefacts. *Scientific Reports*, *7*(1), 6633.

Colwell, C. (2018, April 6). Rights of the dead and the living clash when scientists extract DNA from human remains. *The Conversation*. http://theconversation.com/rights-of-the-dead-and-the-living-clash-when-scientists-extract-dna-from-human-remains-94284

Cooper, A., & Poinar, H. N. (2000). Ancient DNA: do it right or not at all [Review of *Ancient DNA: do it right or not at all*]. *Science*, *289*(5482), 1139.

Corthals, A., Koller, A., Martin, D. W., Rieger, R., Chen, E. I., Bernaski, M., Recagno, G., & Dávalos, L. M. (2012). Detecting the immune system response of a 500 year-old Inca mummy. *PloS One*, *7*(7), e41244.

Covington, A. D. (2009). *Tanning Chemistry: The Science of Leather*. Royal Society of Chemistry.

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372.

Craig, O. E., Chapman, J., Heron, C., Willis, L. H., Bartosiewicz, L., Taylor, G., Whittle, A., & Collins, M. (2005). Did the first farmers of central and eastern Europe produce dairy foods? *Antiquity*, *79*(306), 882–894.

Craig, O., Mulville, J., Pearson, M. P., Sokol, R., Gelsthorpe, K., Stacey, R., & Collins, M. (2000). Detecting milk proteins in ancient pots. *Nature*, *408*(6810), 312.

Cramp, L. J. E., Evershed, R. P., Lavento, M., Halinen, P., Mannermaa, K., Oinonen, M., Kettunen, J., Perola, M., Onkamo, P., & Heyd, V. (2014). Neolithic dairy farming at the extreme of agriculture in northern Europe. *Proceedings. Biological Sciences / The Royal Society*, *281*(1791), 20140819.

cRAP protein sequences. (n.d.). Retrieved June 27, 2020, from https://www.thegpm.org/crap/

Crisp, M. K. (2013). *Amino acid racemization dating: Method development using African ostrich (Struthio camelus) eggshell* [Phd, University of York]. http://etheses.whiterose.ac.uk/4770/

Dallongeville, S., Koperska, M., Garnier, N., Reille-Taillefert, G., Rolando, C., & Tokarski, C. (2011). Identification of animal glue species in artworks using proteomics: application to a 18th century gilt sample. *Analytical Chemistry*, *83*(24), 9431–9437.

Davis, S. J. M., Albarella, U., Detry, C., Ginja, C., Götherström, A., Pires, A. E., Sendim, A., & Svensson, E. M. (2018). An osteometrical method for sexing cattle bones: the metacarpals from 17th century Carnide, Lisbon, Portugal. *Annalen Des Naturhistorischen Museums in Wien. Serie A. Mineralogie Und Petrographie, Geologie Und Palaeontologie, Anthropologie Und Praehistorie*, *120*, 367–388.

Debono Spiteri, C., Gillis, R. E., Roffet-Salque, M., Castells Navarro, L., Guilaine, J., Manen, C., Muntoni, I. M., Saña Segui, M., Urem-Kotsou, D., Whelton, H. L., Craig, O. E., Vigne, J.-D., & Evershed, R. P. (2016). Regional asynchronicity in dairy production and processing in early farming communities of the northern Mediterranean. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(48), 13594–13599.

de Jong, E. W., Westbroek, P., Westbroek, J. W., & Bruning, J. W. (1974). Preservation of antigenic properties of macromolecules over 70 Myr. *Nature*, *252*(5478), 63–64.

Delsuc, F., Kuch, M., Gibb, G. C., Karpinski, E., Hackenberger, D., Szpak, P., Martínez, J. G., Mead, J. I., McDonald, H. G., MacPhee, R. D. E., Billet, G., Hautier, L., & Poinar, H. N. (2019). Ancient Mitogenomes Reveal the Evolutionary History and Biogeography of Sloths. *Current Biology: CB*, *29*(12), 2031–2042.e6.

Demarchi, B., Boano, R., Ceron, A., Bello, F. D., Favero-Longo, S. E., Fiddyment, S., Marochetti, E. F., Mangiapane, G., Mattonai, M., Pennacini, C., Ribechini, E., Woolley, J., Zilberstein, G., & Righetti, P. G. (2020). Never boring: Non-invasive palaeoproteomics of mummified human skin. *Journal of Archaeological Science*, *119*, 105145.

Demarchi, B., & Collins, M. (2014). Amino Acid Racemization Dating. In *Encyclopedia of Scientific Dating Methods* (pp. 1–22). https://doi.org/10.1007/978-94-007-6326-5_73-1

Demarchi, B., Collins, M. J., Tomiak, P. J., Davies, B. J., & Penkman, K. E. H. (2013). Intra-crystalline protein diagenesis (IcPD) in Patella vulgata. Part II: Breakdown and temperature sensitivity. In *Quaternary Geochronology* (Vol. 16, pp. 158–172). https://doi.org/10.1016/j.quageo.2012.08.001

Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C. L., Woolley, J., Crisp, M. K., Wilson, J., Fotakis, A., Fischer, R., Kessler, B. M., Rakownikow Jersie-Christensen, R., Olsen, J. V., Haile, J., Thomas, J., Marean, C. W., Parkington, J., Presslee, S., Lee-Thorp, J., Ditchfield, P., … Collins, M. J. (2016). Protein sequences bound to mineral surfaces persist into deep time. *eLife*, *5*. https://doi.org/10.7554/eLife.17092

Dickinson, M. R., Lister, A. M., & Penkman, K. E. H. (2019). A new method for enamel amino acid racemization dating: a closed system approach. *Quaternary Geochronology*. https://www.sciencedirect.com/science/article/pii/S1871101418300773

Dixit, S. N., Seyer, J. M., Kang, A. H., & Gross, J. (1978). Covalent structure of collagen: amino acid sequence of chick skin collagen alpha1(1)-CB6B. *Biochemistry*, *17*(26), 5719–5722.

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

Doherty, S., Alexander, M. M., Vnouček, J., Newton, J., & Collins, M. J. (2021). Measuring the impact of parchment production on skin collagen stable isotope (δ13C and δ15N) values. *STAR: Science & Technology of Archaeological Research*, *7*(1), 1–12.

Dooley, K. A., Lomax, S., Zeibel, J. G., Miliani, C., Ricciardi, P., Hoenigswald, A., Loew, M., & Delaney, J. K. (2013). Mapping of egg yolk and animal skin glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy. *The Analyst*, *138*(17), 4838–4848.

Dunne, J., Evershed, R. P., Salque, M., Cramp, L., Bruni, S., Ryan, K., Biagetti, S., & di Lernia, S. (2012). First dairying in green Saharan Africa in the fifth millennium BC. *Nature*, *486*(7403), 390–394.

Dunne, J., Rebay-Salisbury, K., Salisbury, R. B., Frisch, A., Walton-Doyle, C., & Evershed, R. P. (2019). Milk of ruminants in ceramic baby bottles from prehistoric child graves. *Nature*, *574*(7777), 246–248.

Entezami, P., Fox, D. A., Clapham, P. J., & Chung, K. C. (2011). Historical perspective on the etiology of rheumatoid arthritis. *Hand Clinics*, *27*(1), 1–10.

Evershed, R. P., Payne, S., Sherratt, A. G., Copley, M. S., Coolidge, J., Urem-Kotsu, D., Kotsakis, K., Ozdoğan, M., Ozdoğan, A. E., Nieuwenhuyse, O., Akkermans, P. M. M. G., Bailey, D., Andeescu, R.-R., Campbell, S., Farid, S., Hodder, I., Yalman, N., Ozbaşaran, M., Biçakci, E., … Burton, M. M. (2008). Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature*, *455*(7212), 528–531.

Fagernäs, Z., García-Collado, M. I., Hendy, J., Hofman, C. A., Speller, C., Velsko, I., & Warinner, C. (2020). A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus. *Journal of Archaeological Science*, *118*, 105135.

Fiddyment, S., Holsinger, B., Ruzzier, C., Devine, A., Binois, A., Albarella, U., Fischer, R., Nichols, E., Curtis, A., Cheese, E., Teasdale, M. D., Checkley-Scott, C., Milner, S. J., Rudy, K. M., Johnson, E. J., Vnouček, J., Garrison, M., McGrory, S., Bradley, D. G., & Collins, M. J. (2015). Animal origin of 13th-century uterine vellum revealed using noninvasive peptide fingerprinting. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), 15066–15071.

Fiddyment, S., Teasdale, M. D., Vnouček, J., Lévêque, É., Binois, A., & Collins, M. J. (2019). So you want to do biocodicology? A field guide to the biological analysis of parchment. *Heritage Science*, *7*(1), 35.

Fischer, R., & Kessler, B. M. (2015a). Gel-aided sample preparation (GASP)—A simplified method for gel-assisted proteomic sample generation from protein extracts and intact cells. *Proteomics*, *15*(7), 1224–1229.

Fischer, R., & Kessler, B. M. (2015b). Gel-aided sample preparation (GASP)-A simplified method for gel-assisted proteomic sample generation from protein extracts and intact cells. *Proteomics*, *15*(7), 1224–1229.

Frank, A., Wochnik, A. S., Bein, T., & Scheu, C. (2017). A biomolecule-assisted, cost-efficient route for growing tunable CuInS 2 films for green energy application. *RSC Advances*, *7*(33), 20219–20230.

Froment, C., Hourset, M., Sáenz-Oyhéréguy, N., Mouton-Barbosa, E., Willmann, C., Zanolli, C., Esclassan, R., Donat, R., Thèves, C., Burlet-Schiltz, O., & Mollereau, C. (2020). Analysis of 5000 year-old human teeth using optimized large-scale and targeted proteomics approaches for detection of sex-specific peptides. *Journal of Proteomics*, *211*, 103548.

Geber, J., Tromp, M., Scott, A., Bouwman, A., Nanni, P., Grossmann, J., Hendy, J., & Warinner, C. (2019). Relief food subsistence revealed by microparticle and proteomic analyses of dental calculus from victims of the Great Irish Famine. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1908839116
Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M., & Barnes, I. (2005). Assessing ancient DNA studies. *Trends in Ecology & Evolution*, *20*(10), 541–544.

Ginolhac, A., Rasmussen, M., Gilbert, M. T. P., Willerslev, E., & Orlando, L. (2011). mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* , *27*(15), 2153–2155.

Gold, D. A., Robinson, J., Farrell, A. B., Harris, J. M., Thalmann, O., & Jacobs, D. K. (2014). Attempted DNA extraction from a Rancho La Brea Columbian mammoth (Mammuthus columbi): prospects for ancient DNA from asphalt deposits. *Ecology and Evolution*, *4*(4), 329–336.

Gotherstrom, A., Collins, M. J., Angerbjorn, A., & Liden, K. (2002). Bone preservation and DNA amplification. *Archaeometry*, *44*(3), 395–404.

Greco, E., El-Aguizy, O., Ali, M. F., Foti, S., Cunsolo, V., Saletti, R., & Ciliberto, E. (2018). Proteomic Analyses on an Ancient Egyptian Cheese and Biomolecular Evidence of Brucellosis. *Analytical Chemistry*, *90*(16), 9673–9676.

Green, E. J., & Speller, C. F. (2017). Novel Substrates as Sources of Ancient DNA: Prospects and Hurdles. *Genes*, *8*(7). https://doi.org/10.3390/genes8070180

Hao, P., Ren, Y., Alpert, A. J., & Sze, S. K. (2011). Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Molecular & Cellular Proteomics: MCP*, *10*(10), O111.009381.

Hare, P.E. (1968). Racemization of amino acids in fossil shells. *Carnegie Inst. Wash. Yearb.*, *66*, 526–528.

Hare, P. E. (1976). Relative reaction rates and activation energies for some amino acid reactions. *Carnegie Institute of Washington Yearbook*, *70*, 801–806.

Helgason, A., Pálsson, S., Lalueza-Fox, C., Ghosh, S., Sigurðardóttir, S., Baker, A., Hrafnkelsson, B., Árnadóttir, L., Þorsteinsdóttir, U., & Stefánsson, K. (2007). A Statistical Approach to Identify Ancient Template DNA. In *Journal of Molecular Evolution* (Vol. 65, Issue 1, pp. 92–102). https://doi.org/10.1007/s00239-006-0259-8

Hendy, J., Collins, M., Teoh, K. Y., Ashford, D. A., Thomas-Oates, J., Donoghue, H. D., Pap, I., Minnikin, D. E., Spigelman, M., & Buckley, M. (2016). The challenge of identifying tuberculosis proteins in archaeological tissues. *Journal of Archaeological Science*, *66*, 146–153.

Hendy, J., Colonese, A. C., Franz, I., Fernandes, R., Fischer, R., Orton, D., Lucquin, A., Spindler, L., Anvari, J., Stroud, E., Biehl, P. F., Speller, C., Boivin, N., Mackie, M., Jersie-Christensen, R. R., Olsen, J. V., Collins, M. J., Craig, O. E., & Rosenstock, E. (2018b). Ancient proteins from ceramic vessels at Çatalhöyük West reveal the hidden cuisine of early farmers. *Nature Communications*, *9*(1), 4064.

Hendy, J., Warinner, C., Bouwman, A., Collins, M. J., Fiddyment, S., Fischer, R., Hagan, R., Hofman, C. A., Holst, M., Chaves, E., Klaus, L., Larson, G., Mackie, M., McGrath, K., Mundorff, A. Z., Radini, A., Rao, H., Trachsel, C., Velsko, I. M., & Speller, C. F. (2018a). Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings. Biological Sciences / The Royal Society*, *285*(1883). https://doi.org/10.1098/rspb.2018.0977

Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C., & Collins, M. J. (2018c). A guide to ancient protein studies. *Nature Ecology & Evolution*, *2*(5), 791–799.

High, K., Milner, N., Panter, I., Demarchi, B., & Penkman, K. E. H. (2016). Lessons from Star Carr on the vulnerability of organic archaeological remains to environmental change. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), 12957–12962.

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

Hill, R. C., Wither, M. J., Nemkov, T., Barrett, A., D'Alessandro, A., Dzieciatkowska, M., & Hansen, K. C. (2015). Preserved Proteins from Extinct Bison latifrons Identified by Tandem Mass Spectrometry; Hydroxylysine Glycosides are a Common Feature of Ancient Collagen. *Molecular & Cellular Proteomics: MCP*, *14*(7), 1946–1958.

Hill, R. L. (1965). Hydrolysis of proteins. *Advances in Protein Chemistry*, *20*, 37–107.
Hirosawa, M., Hoshida, M., Ishikawa, M., & Toya, T. (1993). MASCOT: multiple alignment

system for protein sequences based on three-way dynamic programming. *Computer Applications in the Biosciences: CABIOS*, *9*(2), 161–167.

Hofreiter, M., Jaenicke, V., Serre, D., Haeseler, A. von, & Pääbo, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research*, *29*(23), 4793–4799.

Hong, C., Jiang, H., Lü, E., Wu, Y., Guo, L., Xie, Y., Wang, C., & Yang, Y. (2012). Identification of milk component in ancient food residue by proteomics. *PloS One*, *7*(5), e37053.

Höss, M., Jaruga, P., Zastawny, T. H., Dizdaroglu, M., & Pääbo, S. (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, *24*(7), 1304–1307.
Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95.

Huq, N. L., Tseng, A., & Chapman, G. E. (1990). Partial amino acid sequence of osteocalcin from an extinct species of ratite bird. *Biochemistry International*, *21*(3), 491–496.

Jeong, C., Wilkin, S., Amgalantugs, T., Bouwman, A. S., Taylor, W. T. T., Hagan, R. W., Bromage, S., Tsolmon, S., Trachsel, C., Grossmann, J., Littleton, J., Makarewicz, C. A., Krigbaum, J., Burri, M., Scott, A., Davaasambuu, G., Wright, J., Irmer, F., Myagmar, E., … Warinner, C. (2018). Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), E11248–E11255.

Jersie-Christensen, R. R., Lanigan, L. T., Lyon, D., Mackie, M., Belstrøm, D., Kelstrup, C. D., Fotakis, A. K., Willerslev, E., Lynnerup, N., Jensen, L. J., Cappellini, E., & Olsen, J. V. (2018). Quantitative metaproteomics of medieval dental calculus reveals individual oral health status. *Nature Communications*, *9*(1), 4744.

Jia, L., & Sun, Y. (2017). Protein asparagine deamidation prediction based on structures with machine learning methods. *PloS One*, *12*(7), e0181347.

Jiang, X., Ye, M., Jiang, X., Liu, G., Feng, S., Cui, L., & Zou, H. (2007). Method Development of Efficient Protein Extraction in Bone Tissue for Proteome Analysis. In *Journal of Proteome Research* (Vol. 6, Issue 6, pp. 2287–2294). https://doi.org/10.1021/pr070056t

Johnson, R. R. (1970). Ancient and Medieval Accounts of the“ Invention” of Parchment. *California Studies in Classical Antiquity*, *3*, 115–122.

Jones, E., Oliphant, T., Peterson, P., & Others. (2001). *SciPy: Open source scientific tools for Python*.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* , *29*(13), 1682–1684.

Kato, K., Nakayoshi, T., Kurimoto, E., & Oda, A. (2019). Computational Studies on the Nonenzymatic Deamidation Mechanisms of Glutamine Residues. *ACS Omega*, *4*(2), 3508–3513.

Kemp, C. (2015). Museums: The endangered dead. *Nature*, *518*(7539), 292–294.
Kennedy, C. J., & Wess, T. J. (2003). The structure of collagen within parchment--a review. *Restaurator. International Journal for the Preservation of Library and Archival Material*,

24(2), 61–80.

Key, F. M., Posth, C., Krause, J., Herbig, A., & Bos, K. I. (2017). Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends in Genetics: TIG*, *33*(8), 508–520.

Kistler, L., Ware, R., Smith, O., Collins, M., & Allaby, R. G. (2017). A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research*, *45*(11), 6310–6320.

Knecht, H. (1997). Projectile Points of Bone, Antler, and Stone. In H. Knecht (Ed.), *Projectile Technology* (pp. 191–212). Springer US.

Kofford, M. W., Schwartz, L. B., Schechter, N. M., Yager, D. R., Diegelmann, R. F., & Graham, M. F. (1997). Cleavage of type I procollagen by human mast cell chymase initiates collagen fibril formation and generates a unique carboxyl-terminal propeptide. *The Journal of Biological Chemistry*, *272*(11), 7127–7131.

Koppenhoefer, R. M. (n.d.). The lipids of sheep skins: I. Lipids of fresh sheep skin. *Journal of the American Leather Chemists Association*, *33*, 203–215.

Kossiakoff, A. A. (1988). Tertiary structure is a principal determinant to protein deamidation. *Science*, *240*(4849), 191–194.

Krause, J., Briggs, A. W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., & Pääbo, S. (2010). A Complete mtDNA Genome of an Early Modern Human from Kostenki, Russia. In *Current Biology* (Vol. 20, Issue 3, pp. 231–236). https://doi.org/10.1016/j.cub.2009.11.068

Kriausakul, N., & Mitterer, R. M. (1980). Some factors affecting the epimerization of isoleucine in peptides and proteins. *Biogeochemistry of Amino Acids*, 283–296.

Kuckova, S., Crhova, M., Vankova, L., Hnizda, A., Hynek, R., & Kodicek, M. (2009). Towards proteomic analysis of milk proteins in historical building materials. *International Journal of Mass Spectrometry*, *284*(1), 42–46.

Lanigan, L. T., Mackie, M., Feine, S., Hublin, J.-J., Schmitz, R. W., Wilcke, A., Collins, M. J., Cappellini, E., Olsen, J. V., Taurozzi, A. J., & Welker, F. (2020). Multi-protease analysis of Pleistocene bone proteomes. *Journal of Proteomics*, 103889.

Lara-Villoslada, F., Olivares, M., & Xaus, J. (2005). The balance between caseins and whey proteins in cow's milk determines its allergenicity. *Journal of Dairy Science*, *88*(5), 1654–1660.

Lee, J., Kim, M.-H., Lee, K.-B., van Elslande, E., Walter, P., & Lee, Y. (2014). Analysis of natural dyes in archeological textiles using TOF-SIMS and other analytical techniques. *Surface and Interface Analysis: SIA*, *46*(S1), 312–316.

Lee, L. (1992). THE CONSERVATION OF PLEATED ILLUMINATED VELLUM LEAVES IN THE ASHMOLE BESTIARY. *Paper Conservator*, *16*(1), 46–49.

Le Meillour, L., Zazzo, A., Lesur, J., Cersoy, S., & Zirah, S. (2018). Identification of degraded bone and tooth splinters from arid environments using palaeoproteomics. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *511*, 472–482.

Leo, G., Bonaduce, I., Andreotti, A., Marino, G., Pucci, P., Colombini, 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*, *83*(6), 2056–2064.

Liardon, R., & Ledermann, S. (1986). Racemization kinetics of free and protein-bound amino acids under moderate alkaline treatment. *Journal of Agricultural and Food Chemistry*, *34*(3), 557–565.

Li, C., Ning, C., Hagelberg, E., Li, H., Zhao, Y., Li, W., Abuduresule, I., Zhu, H., & Zhou, H. (2015). Analysis of ancient human mitochondrial DNA from the Xiaohe cemetery: insights into prehistoric population movements in the Tarim Basin, China. *BMC Genetics*, *16*, 78.
Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*(6422), 709–715.

Li, X., Lin, C., & O'Connor, P. B. (2010). Glutamine deamidation: differentiation of glutamic acid and gamma-glutamic acid in peptides by electron capture dissociation. *Analytical Chemistry*, *82*(9), 3606–3615.

Lluveras-Tenorio, A., Vinciguerra, R., Galano, E., Blaensdorf, C., Emmerling, E., Perla Colombini, M., Birolo, L., & Bonaduce, I. (2017). GC/MS and proteomics to unravel the painting history of the lost Giant Buddhas of Bāmiyān (Afghanistan). *PloS One*, *12*(4), e0172990.

Lodish, H., Berk, A., Lawrence Zipursky, S., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). *Collagen: The Fibrous Proteins of the Matrix*. W. H. Freeman.

Lopez-Polin, L. (2012). Possible interferences of some conservation treatments with subsequent studies on fossil bones: A conservator's overview. *Quaternary International: The Journal of the International Union for Quaternary Research*, *275*, 120–127.

Lugli, F., Di Rocco, G., Vazzana, A., Genovese, F., Pinetti, D., Cilli, E., Carile, M. C., Silvestrini, S., Gabanini, G., Arrighi, S., Buti, L., Bortolini, E., Cipriani, A., Figus, C., Marciani, G., Oxilia, G., Romandini, M., Sorrentino, R., Sola, M., & Benazzi, S. (2019). Enamel peptides reveal the sex of the Late Antique "Lovers of Modena." *Scientific Reports*, *9*(1), 13130.

Ma, B.-G., Chen, L., Ji, H.-F., Chen, Z.-H., Yang, F.-R., Wang, L., Qu, G., Jiang, Y.-Y., Ji, C., & Zhang, H.-Y. (2008). Characters of very ancient proteins. *Biochemical and Biophysical Research Communications*, *366*(3), 607–611.
Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., & Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Communications in Mass Spectrometry: RCM*, *17*(20), 2337–2342.

Mackie, M., Hendy, J., Lowe, A. D., Sperduti, A., Holst, M., Collins, M. J., & Speller, C. F. (2017). Preservation of the metaproteome: variability of protein preservation in ancient dental calculus. *Science and Technology of Archaeological Research*, *3*(1), 74–86.

Mackie, M., Radini, A., & Speller, C. F. (2017). *The Sustainability of Dental Calculus for Archaeological Research* (J. Favreau & R. Patalano (Eds.); p. 8).

Mackie, M., Rüther, P., Samodova, D., Di Gianvincenzo, F., Granzotto, C., Lyon, D., Peggie, D. A., Howard, H., Harrison, L., Jensen, L. J., Olsen, J. V., & Cappellini, E. (2018). Palaeoproteomic Profiling of Conservation Layers on a 14th Century Italian Wall Painting. *Angewandte Chemie* , *57*(25), 7369–7374.

Mahajan, S. (2019). Role of Human Tooth Wear Analysis in Archaeology: A Review. *Ancient Asia*, *10*. https://www.ancient-asia-journal.com/articles/10.5334/aa.181/print/

Maixner, F., Overath, T., Linke, D., Janko, M., Guerriero, G., van den Berg, B. H. J., Stade, B., Leidinger, P., Backes, C., Jaremek, M., Kneissl, B., Meder, B., Franke, A., Egarter-Vigl, E., Meese, E., Schwarz, A., Tholey, A., Zink, A., & Keller, A. (2013). Paleoproteomic study of the Iceman's brain tissue. In *Cellular and Molecular Life Sciences* (Vol. 70, Issue 19, pp. 3709–3722). https://doi.org/10.1007/s00018-013-1360-y

Makarewicz, C., Marom, N., & Bar-Oz, G. (2017). Palaeobiology: Ensure equal access to ancient DNA [Review of *Palaeobiology: Ensure equal access to ancient DNA*]. *Nature*, *548*(7666), 158.

Maki, J., Tanabe, T., & Sekiya, H. (2014). The visual recognition of parasitic helminths in Japan before the introduction of parasitology from Germany-A preliminary note on the confirmation from Jomon Period onward. **松山大学論集**, *26*(5), 231–248.

Malaspinas, A.-S., Tange, O., Moreno-Mayar, J. V., Rasmussen, M., DeGiorgio, M., Wang, Y., Valdiosera, C. E., Politis, G., Willerslev, E., & Nielsen, R. (2014). bammds: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). *Bioinformatics* , *30*(20), 2962–2964.

Maltby, M., Allen, M., Best, J., Fothergill, B. T., & Demarchi, B. (2018). Counting Roman chickens: Multidisciplinary approaches to human-chicken interactions in Roman Britain. *Journal of Archaeological Science: Reports*, *19*, 1003–1015.

Manguy, J., Jehl, P., Dillon, E. T., Davey, N. E., Shields, D. C., & Holton, T. A. (2017). Peptigram: A Web-Based Application for Peptidomics Data Visualization. *Journal of Proteome Research*, *16*(2), 712–719.

Masson, M., Molnár, E., Donoghue, H. D., Besra, G. S., Minnikin, D. E., Wu, H. H. T., Lee, O. Y.-C., Bull, I. D., & Pálfi, G. (2013). Osteological and biomolecular evidence of a 7000-year-old case of hypertrophic pulmonary osteopathy secondary to tuberculosis from neolithic hungary. *PloS One*, *8*(10), e78252.

Mays, S., Roberts, D., Marshall, P., Pike, A. W. G., van Heekeren, V., Bronk Ramsey, C., Dunbar, E., Reimer, P., Linscott, B., Radini, A., Lowe, A., Dowle, A., Speller, C., Vallender, J., & Bedford, J. (2018). Lives before and after Stonehenge: An osteobiographical study of four prehistoric burials recently excavated from the Stonehenge World Heritage Site. *Journal of Archaeological Science: Reports*, *20*, 692–710.

McCoy, V. E., Gabbott, S. E., Penkman, K., Collins, M. J., Presslee, S., Holt, J., Grossman, H., Wang, B., Solórzano Kraemer, M. M., Delclòs, X., & Peñalver, E. (2019). Ancient amino acids from fossil feathers in amber. *Scientific Reports*, *9*(1), 6420.

McGrath, K., Rowsell, K., Gates St-Pierre, C., Tedder, A., Foody, G., Roberts, C., Speller, C., & Collins, M. (2019). Identifying Archaeological Bone via Non-Destructive ZooMS and the Materiality of Symbolic Expression: Examples from Iroquoian Bone Points. *Scientific Reports*, *9*(1), 11027.

Mellars P, Dark P (1998) Star Carr in Context (*McDonald Institute for Archaeological Research, Cambridge, UK*).

Menderes, O., Covington, A. D., Waite, E. R., & Collins, M. J. (1999). The mechanism and effects of collagen amide group hydrolysis during liming. *Journal of the Society of Leather Technologists and Chemists*, *83*(2), 107–110.

Meyer, M., Arsuaga, J.-L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B., Martínez, I., Gracia, A., de Castro, J. M. B., Carbonell, E., Viola, B., Kelso, J., Prüfer, K., & Pääbo, S. (2016). Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. In *Nature* (Vol. 531, Issue 7595, pp. 504–507). https://doi.org/10.1038/nature17405

Miller, G. H., Magee, J. W., & Jull, A. (1997). Low-latitude glacial cooling in the Southern Hemisphere from amino-acid racemization in emu eggshells. *Nature*, *385*(6613), 241.

Miyahara, M., Hayashi, K., Berger, J., Tanzawa, K., Njieha, F. K., Trelstad, R. L., & Prockop, D. J. (1984). Formation of collagen fibrils by enzymic cleavage of precursors of type I collagen in vitro. *The Journal of Biological Chemistry*, *259*(15), 9891–9898.

Moonesinghe, R., Khoury, M. J., & Janssens, A. C. J. W. (2007). Most published research findings are false-but a little replication goes a long way. *PLoS Medicine*, *4*(2), e28.

Mühlemann, B., Vinner, L., Margaryan, A., Wilhelmson, H., de la Fuente Castro, C., Allentoft, M. E., de Barros Damgaard, P., Hansen, A. J., Holtsmark Nielsen, S., Strand, L. M., Bill, J., Buzhilova, A., Pushkina, T., Falys, C., Khartanovich, V., Moiseyev, V., Jørkov, M. L. S., Østergaard Sørensen, P., Magnusson, Y., … Sikora, M. (2020). Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science*, *369*(6502). https://doi.org/10.1126/science.aaw8977

Müller, R., Roberts, C. A., & Brown, T. A. (2016). Complications in the study of ancient tuberculosis: Presence of environmental bacteria in human archaeological remains. *Journal of Archaeological Science*, *68*, 5–11.

Murray, K., Rodwell, V., Bender, D., Botham, K. M., Weil, P. A., & Kennelly, P. J. (2009). *Harper's illustrated biochemistry. 28*. Citeseer.

Murray, M. S. (2008). Zooarchaeology and Arctic marine mammal biogeography, conservation, and management. *Ecological Applications: A Publication of the Ecological Society of America*, *18*(2 Suppl), S41–S55.
Naihui, W., Samantha, B., Peter, D., Sandra, H., Maxim, K., Sindy, L., Oshan, W., Stefano, G., Michael, C., Liora, H. K., Matthew, S., Glenn, S., Michael, S., Kristine, R. K., & Katerina, D. (2021). Testing the efficacy and comparability of ZooMS protocols on archaeological bone. *Journal of Proteomics*, *233*, 104078.

Némethy, G., & Scheraga, H. A. (1986). Stabilization of collagen fibrils by hydroxyproline. *Biochemistry*, *25*(11), 3184–3188.

Nicholson, G. J., Tomiuk, J., Czarnetzki, A., Bachmann, L., & Pusch, C. M. (2002). Detection of bone glue treatment as a major source of contamination in ancient DNA analyses. *American Journal of Physical Anthropology*, *118*(2), 117–120.

Nistelberger, H. M., Pálsdóttir, A. H., Star, B., Leifsson, R., Gondek, A. T., Orlando, L., Barrett, J. H., Hallsson, J. H., & Boessenkool, S. (2019). Sexing Viking Age horses from burial and non-burial sites in Iceland using ancient DNA. *Journal of Archaeological Science*, *101*, 115–122.

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P. L. F., Fumagalli, M., Vilstrup, J. T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C. D., … Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, *499*(7456), 74–78.

Oró, J., & Skewes, H. B. (1965). Free amino-acids on human fingers: the question of contamination in microanalysis. *Nature*, *207*(5001), 1042–1045.

Orton, D., Anvari, J., Gibson, C., Last, J., Bogaard, A., Rosenstock, E., & Biehl, P. F. (2018). A tale of two tells: dating the Çatalhöyük West Mound. *Antiquity*, *92*(363), 620–639.

Ostrom, P. H., Schall, M., Gandhi, H., Shen, T.-L., Hauschka, P. V., Strahler, J. R., & Gage, D. A. (2000). New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. *Geochimica et Cosmochimica Acta*, *64*(6), 1043–1050.

Oxlund, H., Barckman, M., Ortoft, G., & Andreassen, T. T. (1995). Reduced concentrations of collagen cross-links are associated with reduced strength of bone. *Bone*, *17*(4 Suppl), 365S – 371S.

Papadopoulou, D., Sakalis, A., Merousis, N., & Tsirliganis, N. C. (2007). Study of decorated archeological ceramics by micro X-ray fluorescence spectroscopy. *Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment*, *580*(1), 743–746.

Pearson, A. F. P., Jeffs, B., Witkin, A., & MacQuarrie, H. (2011). *Infernal traffic: Excavation of a liberated African graveyard in Rupert's Valley, St. Helena*. Council for British Archeology.

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

Penkman, K. E. H., Preece, R. C., Bridgland, D. R., Keen, D. H., Meijer, T., Parfitt, S. A., White, T. S., & Collins, M. J. (2011). A chronological framework for the British Quaternary based on Bithynia opercula. *Nature*, *476*(7361), 446–449.
Perri, A. R., Mitchell, K. J., Mouton, A., Álvarez-Carretero, S., Hulme-Beaman, A., Haile, J., Jamieson, A., Meachen, J., Lin, A. T., Schubert, B. W., Ameen, C., Antipina, E. E., Bover, P., Brace, S., Carmagnini, A., Carøe, C., Samaniego Castruita, J. A., Chatters, J. C., Dobney, K., … Frantz, L. A. F. (2021). Dire wolves were the last of an ancient New World canid lineage. *Nature*, *591*(7848), 87–91.

Perry-Gal, L., Erlich, A., Gilboa, A., & Bar-Oz, G. (2015). Earliest economic exploitation of chicken outside East Asia: Evidence from the Hellenistic Southern Levant. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(32), 9849–9854.

Persikov, A. V., Ramshaw, J. A. M., & Brodsky, B. (2005). Prediction of collagen stability from amino acid sequence. *The Journal of Biological Chemistry*, *280*(19), 19343–19349.

Peyrégne, S., & Prüfer, K. (2020). Present-Day DNA Contamination in Ancient DNA Datasets. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *42*(9), 2000081.

Piez, K. A., & Trus, B. L. (1977). Microfibrillar structure and packing of collagen:

hydrophobic interactions. *Journal of Molecular Biology*, *110*(4), 701–704.

Presslee, S., Slater, G. J., Pujos, F., Forasiepi, A. M., Fischer, R., Molloy, K., Mackie, M., Olsen, J. V., Kramarz, A., Taglioretti, M., Scaglia, F., Lezcano, M., Lanata, J. L., Southon, J., Feranec, R., Bloch, J., Hajduk, A., Martin, F. M., Salas Gismondi, R., … MacPhee, R. D. E. (2019). Palaeoproteomics resolves sloth relationships. *Nature Ecology & Evolution*, *3*(7), 1121–1130.

Presslee, S., Penkman, K., Fischer, R., Richards-Slidel, E., Southon, J., Hospitaleche, C. A., Collins, M., & MacPhee, R. (2021). Assessment of different screening methods for selecting palaeontological bone samples for peptide sequencing. *Journal of Proteomics*, *230*, 103986.

Procopio, N., Chamberlain, A. T., & Buckley, M. (2017). Intra- and Interskeletal Proteome Variations in Fresh and Buried Bones. *Journal of Proteome Research*, *16*(5), 2016–2029.

Qian, Y., Engel, M. H., Macko, S. A., Carpenter, S., & Deming, J. W. (1993). Kinetics of peptide hydrolysis and amino acid decomposition at high temperature. *Geochimica et Cosmochimica Acta*, *57*(14), 3281–3293.

Ramsøe, A., van Heekeren, V., Ponce, P., Fischer, R., Barnes, I., Speller, C., & Collins, M. J. (2020). DeamiDATE 1.0: Site-specific deamidation as a tool to assess authenticity of members of ancient proteomes. *Journal of Archaeological Science*, *115*, 105080.

Ramsøe, A., Crispin, M., Mackie, M., McGrath, K., Fischer, R., Demarchi, B., Collins, M. J., Hendy, J., & Speller, C. (2021). Assessing the degradation of ancient milk proteins through site-specific deamidation patterns. *Scientific Reports*, *11*(1), 7795.

Rao, H., Li, B., Yang, Y., Ma, Q., & Wang, C. (2015). Proteomic identification of organic additives in the mortars of ancient Chinese wooden buildings. *Analytical Methods*, *7*(1), 143–149.

Reed, R. (1972). *Ancient skins, parchments and leathers*.

Reed, R. (1975). *The nature and making of parchment*.
https://www.bcin.ca/bcin/detail.app?id=54399
Reed, R., & Reed, R. (1972). *Ancient Skins, Parchments and Leathers*. Seminar Press.

Renaud, G., Slon, V., Duggan, A. T., & Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biology*, *16*, 224.

Robinson, N. E., & Robinson, A. (2004). Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins. *Althouse Press*.

Robinson, N. E., & Robinson, A. B. (2001). Prediction of protein deamidation rates from primary and three-dimensional structure. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(8), 4367–4372.

Robinson, N. E., Robinson, Z. W., Robinson, B. R., Robinson, A. L., Robinson, J. A., Robinson, M. L., & Robinson, A. B. (2004). Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. *The Journal of Peptide Research: Official Journal of the American Peptide Society*, *63*(5), 426–436.

Rohland, N., Siedel, H., & Hofreiter, M. (2004). Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *BioTechniques*, *36*(5), 814–816, 818–821.

Rowntree, B. S. (1901). Poverty: A study of town life. *Macmillan*.

Rybczynski, N., Gosse, J. C., Harington, C. R., Wogelius, R. A., Hidy, A. J., & Buckley, M. (2013). Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nature Communications*, *4*, 1550.

Ryder, M. L. (1964). Parchment—its history, manufacture and composition. *Journal of the Society of Archivists. Society of Archivists* , *2*(9), 391–399.

San Antonio, J. D., Schweitzer, M. H., Jensen, S. T., Kalluri, R., Buckley, M., & Orgel, J. P. R. O. (2011). Dinosaur peptides suggest mechanisms of protein survival. *PloS One*, *6*(6), e20381.

Sawafuji, R., Cappellini, E., Nagaoka, T., Fotakis, A. K., Jersie-Christensen, R. R., Olsen, J. V., Hirata, K., & Ueda, S. (2017). Proteomic profiling of archaeological human bone. *Royal Society Open Science*, *4*(6), 161004.

Saxl, H. (1954). *An Investigation of the Qualities, the Methods of Manufacture and the Preservation of Historic Parchment and Vellum with a View to Identifying the Animal Species Used: Being a Thesis Presented to the University of Leeds, Department of Leather Industries*. University of Leeds.

Sayle, R. A., & Milner-White, E. J. (1995). RASMOL: biomolecular graphics for all. *Trends in Biochemical Sciences*, *20*(9), 374.

Schroeter, E. R., Blackburn, K., Goshe, M. B., & Schweitzer, M. H. (2019). Proteomic method to extract, concentrate, digest and enrich peptides from fossils with coloured (humic) substances for mass spectrometry analyses. *Royal Society Open Science*, *6*(8), 181433.

Schroeter, E. R., & Cleland, T. P. (2016). Glutamine deamidation: an indicator of antiquity, or preservational quality? *Rapid Communications in Mass Spectrometry: RCM*, *30*(2), 251–255.

Schroeter, E. R., DeHart, C. J., Cleland, T. P., Zheng, W., Thomas, P. M., Kelleher, N. L., Bern, M., & Schweitzer, M. H. (2017). Expansion for the Brachylophosaurus canadensis Collagen I Sequence and Additional Evidence of the Preservation of Cretaceous Protein. *Journal of Proteome Research*, *16*(2), 920–932.

Schroeter, E. R., DeHart, C. J., Schweitzer, M. H., Thomas, P. M., & Kelleher, N. L. (2016). Bone protein “extractomics”: comparing the efficiency of bone protein extractions of Gallus gallus in tandem mass spectrometry, with an eye towards paleoproteomics. *PeerJ*, *4*, e2603.

Shapiro, B., Hofreiter, M., La, J., & Er, A. (2012). Ancient DNA: methods and protocols. *Humana Press New York*.

Shevchenko, A., Schuhmann, A., Thomas, H., & Wetzel, G. (2018). Fine Endmesolithic fish caviar meal discovered by proteomics in foodcrusts from archaeological site Friesack 4 (Brandenburg, Germany). *PloS One*, *13*(11), e0206483.

Shevchenko, A., Yang, Y., Knaust, A., Thomas, H., Jiang, H., Lu, E., Wang, C., & Shevchenko, A. (2014). Proteomics identifies the composition and manufacturing recipe of the 2500-year old sourdough bread from Subeixi cemetery in China. *Journal of Proteomics*, *105*, 363–371.

Simpson, J. P. (2015). *Investigating the relationship between glutamine deamidation and collagen breakdown in bone* [Phd, University of York]. http://etheses.whiterose.ac.uk/id/eprint/13664

Simpson, J. P., Fascione, M., Bergström, E., Wilson, J., Collins, M. J., Penkman, K. E. H., & Thomas-Oates, J. (2019). Ionisation bias undermines the use of matrix-assisted laser desorption/ionisation for estimating peptide deamidation: Synthetic peptide studies demonstrate electrospray ionisation gives more reliable response ratios. *Rapid Communications in Mass Spectrometry: RCM*, *33*(12), 1049–1057.

Simpson, J. P., Penkman, K. E. H., Demarchi, B., Koon, H., Collins, M. J., Thomas-Oates, J., Shapiro, B., Stark, M., & Wilson, J. (2016). The effects of demineralisation and sampling point variability on the measurement of glutamine deamidation in type I collagen extracted from bone. *Journal of Archaeological Science*, *69*, 29–38.

Sirin, N., Matzenauer, C., Reckert, A., & Ritz-Timme, S. (2018). Age estimation based on aspartic acid racemization in dentine: what about caries-affected teeth? *International Journal of Legal Medicine*, *132*(2), 623–628.

Skoglund, P., Northoff, B. H., Shunkov, M. V., Derevianko, A. P., Pääbo, S., Krause, J., & Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2229–2234.

Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C., & 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.

Smith, C. I., Nielsen-Marsh, C. M., Jans, M. M. E., & Collins, M. J. (2007). Bone diagenesis in the European Holocene I: patterns and mechanisms. *Journal of Archaeological Science*, *34*(9), 1485–1493.
Smith, J. S. (2019). Archaeological oral health: a comparison of post-medieval and modern-day dental caries exposure of adults in East London. *British Dental Journal*, *227*(8), 721–725.

Smith, O., Dunshea, G., Sinding, M.-H. S., Fedorov, S., Germonpre, M., Bocherens, H., & Gilbert, M. T. P. (2019). Ancient RNA from Late Pleistocene permafrost and historical canids shows tissue-specific transcriptome survival. *PLoS Biology*, *17*(7), e3000166.

Smith, O., Momber, G., Bates, R., Garwood, P., Fitch, S., Pallen, M., Gaffney, V., & Allaby, R. G. (2015). Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. *Science*, *347*(6225), 998–1001.

Solazzo, C., Courel, B., Connan, J., van Dongen, B. E., Barden, H., Penkman, K., Taylor, S., Demarchi, B., Adam, P., Schaeffer, P., Nissenbaum, A., Bar-Yosef, O., & Buckley, M. (2016). Identification of the earliest collagen- and plant-based coatings from Neolithic artefacts (Nahal Hemar cave, Israel). *Scientific Reports*, *6*, 31053.

Solazzo, C., Dyer, J. M., Clerens, S., Plowman, J., Peacock, E. E., & Collins, M. J. (2013). Proteomic evaluation of the biodegradation of wool fabrics in experimental burials. *International Biodeterioration & Biodegradation*, *80*, 48–59.

Solazzo, C., Fitzhugh, W. W., Rolando, C., & Tokarski, C. (2008). Identification of protein remains in archaeological potsherds by proteomics. *Analytical Chemistry*, *80*(12), 4590–4597.

Solazzo, C., Wadsley, M., Dyer, J. M., Clerens, S., Collins, M. J., & Plowman, J. (2013). Characterisation of novel α-keratin peptide markers for species identification in keratinous tissues using mass spectrometry. *Rapid Communications in Mass Spectrometry: RCM*, *27*(23), 2685–2698.

Solazzo, C., Wilson, J., Dyer, J. M., Clerens, S., Plowman, J. E., von Holstein, I., Walton Rogers, P., Peacock, E. E., & Collins, M. J. (2014). Modeling deamidation in sheep α-keratin peptides and application to archeological wool textiles. *Analytical Chemistry*, *86*(1), 567–575.

Squires, K., Errickson, D., & Márquez-Grant, N. (2019). *Ethical Approaches to Human Remains: A Global Challenge in Bioarchaeology and Forensic Anthropology*. Springer Nature.
Stankiewicz, B. A., & van Bergen, P. F. (1998). Nitrogen and N-Containing Macromolecules in the Bio- and Geosphere: An Introduction. In *Nitrogen-Containing Macromolecules in the Bio- and Geosphere* (Vol. 707, pp. 1–12). American Chemical Society.

Steen, H., & Mann, M. (2004). The ABC's (and XYZ's) of peptide sequencing. *Nature Reviews. Molecular Cell Biology*, *5*(9), 699–711.

Stewart, N. A., Gerlach, R. F., Gowland, R. L., Gron, K. J., & Montgomery, J. (2017). Sex determination of human remains from peptides in tooth enamel. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(52), 13649–13654.

Stock, C. (1930). *Rancho La Brea: a record of Pleistocene life in California*. *1*, 83.

Stratton, L. P., Kelly, R. M., Rowe, J., Shively, J. E., Smith, D. D., Carpenter, J. F., & Manning, M. C. (2001). Controlling deamidation rates in a model peptide: Effects of temperature, peptide concentration, and additives. *Journal of Pharmaceutical Sciences*, *90*(12), 2141–2148.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2018). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613.

Taubenberger, A. V., Woodruff, M. A., Bai, H., Muller, D. J., & Hutmacher, D. W. (2010). The effect of unlocking RGD-motifs in collagen I on pre-osteoblast adhesion and differentiation. *Biomaterials*, *31*(10), 2827–2835.

Teasdale, M. D., Fiddyment, S., Vnouček, J., Mattiangeli, V., Speller, C., Binois, A., Carver, M., Dand, C., Newfield, T. P., Webb, C. C., Bradley, D. G., & Collins, M. J. (2017). The York Gospels: a 1000-year biological palimpsest. *Royal Society Open Science*, *4*(10), 170988.

Thomas P. Gilbert, M., Rudbeck, L., Willerslev, E., Hansen, A. J., Smith, C., Penkman, K. E. H., Prangenberg, K., Nielsen-Marsh, C. M., Jans, M. E., Arthur, P., Lynnerup, N., Turner-Walker, G., Biddle, M., Kjølbye-Biddle, B., & Collins, M. J. (2005). Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. *Journal of Archaeological Science*, *32*(5), 785–793.

Thompson, D. V. (1956). *The Materials and Techniques of Medieval Painting*. Courier Corporation.

Thomsen, P. F., Elias, S., Gilbert, M. T. P., Haile, J., Munch, K., Kuzmina, S., Froese, D. G., Sher, A., Holdaway, R. N., & Willerslev, E. (2009). Non-destructive sampling of ancient insect

DNA. *PloS One*, *4*(4), e5048.

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortensen, H. M., Hirbo, J. B., Osman, M., Ibrahim, M., Omar, S. A., Lema, G., Nyambo, T. B., Ghori, J., Bumpstead, S., Pritchard, J. K., Wray, G. A., & Deloukas, P. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, *39*(1), 31–40.

Tomiak, P. J., Penkman, K. E. H., Hendy, E. J., Demarchi, B., Murrells, S., Davis, S. A., McCullagh, P., & Collins, M. J. (2013). Testing the limitations of artificial protein degradation kinetics using known-age massive Porites coral skeletons. *Quaternary Geochronology*, *16*(C), 87–109.

Tsutaya, T., Mackie, M., Koenig, C., Sato, T., Weber, A. W., Kato, H., Olsen, J. V., & Cappellini, E. (2019). Palaeoproteomic identification of breast milk protein residues from the archaeological skeletal remains of a neonatal dog. *Scientific Reports*, *9*(1), 12841.

UniProt Consortium. (2015). UniProt: a hub for protein information. *Nucleic Acids Research*, *43*(Database issue), D204–D212.

Vanden Berghe, I. (2012). Towards an early warning system for oxidative degradation of protein fibres in historical tapestries by means of calibrated amino acid analysis. *Journal of Archaeological Science*, *39*(5), 1349–1359.

van der Valk, T., Pečnerová, P., Díez-Del-Molino, D., Bergström, A., Oppenheimer, J., Hartmann, S., Xenikoudakis, G., Thomas, J. A., Dehasque, M., Sağlıcan, E., Fidan, F. R., Barnes, I., Liu, S., Somel, M., Heintzman, P. D., Nikolskiy, P., Shapiro, B., Skoglund, P., Hofreiter, M., … Dalén, L. (2021). Million-year-old DNA sheds light on the genomic history of mammoths. *Nature*, *591*(7849), 265–269.

van Doorn, N. L., Wilson, J., Hollund, H., Soressi, M., & Collins, M. J. (2012). Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Communications in Mass Spectrometry: RCM*, *26*(19), 2319–2327.

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

Van Rossum, G., & Drake, F. L., Jr. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Villa, P., Pollarolo, L., Degano, I., Birolo, L., Pasero, M., Biagioni, C., Douka, K., Vinciguerra, R., Lucejko, J. J., & Wadley, L. (2015). A Milk and Ochre Paint Mixture Used 49,000 Years Ago at Sibudu, South Africa. *PloS One*, *10*(6), e0131273.

Vinciguerra, R., Illiano, A., De Chiaro, A., Carpentieri, A., Lluveras-Tenorio, A., Bonaduce, I., Marino, G., Pucci, P., Amoresano, A., & Birolo, L. (2019). Identification of proteinaceous binders in paintings: A targeted proteomic approach for cultural heritage. *Microchemical Journal, Devoted to the Application of Microtechniques in All Branches of Science*, *144*, 319–328.

Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P.-A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H.-J., Albar, J. P., … Hermjakob, H. (2014).

ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology*, *32*(3), 223–226.

von Holstein, I. C. C., Walton Rogers, P., Craig, O. E., Penkman, K. E. H., Newton, J., & Collins, M. J. (2016). Provenancing Archaeological Wool Textiles from Medieval Northern Europe by Light Stable Isotope Analysis (δ13C, δ15N, δ2H). *PloS One*, *11*(10), e0162330.

Wadsworth, C., & Buckley, M. (2014). Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Communications in Mass Spectrometry: RCM*, *28*(6), 605–615.

Wadsworth, C., Procopio, N., Anderung, C., Carretero, J.-M., Iriarte, E., Valdiosera, C., Elburg, R., Penkman, K., & Buckley, M. (2017). Comparing ancient DNA survival and proteome content in 69 archaeological cattle tooth and bone samples from multiple European sites. *Journal of Proteomics*, *158*, 1–8.

Wakankar, A. A., & Borchardt, R. T. (2006). Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. *Journal of Pharmaceutical Sciences*, *95*(11), 2321–2336.

Wall, J. D., & Kim, S. K. (2007). Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genetics*, *3*(10), 1862–1866.

Walt, S. van der, van der Walt, S., Chris Colbert, S., & Varoquaux, G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation. In *Computing in Science & Engineering* (Vol. 13, Issue 2, pp. 22–30). https://doi.org/10.1109/mcse.2011.37

Warinner, C., Hendy, J., Speller, C., Cappellini, E., Fischer, R., Trachsel, C., Arneborg, J., Lynnerup, N., Craig, O. E., Swallow, D. M., Fotakis, A., Christensen, R. J., Olsen, J. V., Liebert, A., Montalva, N., Fiddyment, S., Charlton, S., Mackie, M., Canci, A., … Collins, M. (2014). Direct evidence of milk consumption from ancient human dental calculus. *Scientific Reports*, *4*, 7104.

Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R. Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E., Seiler, R., Hansen, L. H., Castruita, J. A. S., … Cappellini, E. (2014). Pathogens and host immunity in the ancient human oral cavity. *Nature Genetics*, *46*(4), 336–344.

Warinner, C., Speller, C., & Collins, M. J. (2015). A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1660), 20130376.

Welker, F., Collins, M. J., Thomas, J. A., Wadsley, M., & Brace, S. (2015). Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature*. https://www.nature.com/articles/nature14249

Welker, F., Collins, M. J., Thomas, J. A., Wadsley, M., Brace, S., Cappellini, E., Turvey, S. T., Reguero, M., Gelfo, J. N., Kramarz, A., Burger, J., Thomas-Oates, J., Ashford, D. A., Ashton, P. D., Rowsell, K., Porter, D. M., Kessler, B., Fischer, R., Baessmann, C., … MacPhee, R. D. E. (2015). Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature*, *522*(7554), 81–84.

Welker, F., Hajdinjak, M., Talamo, S., Jaouen, K., Dannemann, M., David, F., Julien, M., Meyer, M., Kelso, J., Barnes, I., Brace, S., Kamminga, P., Fischer, R., Kessler, B. M., Stewart, J. R., Pääbo, S., Collins, M. J., & Hublin, J.-J. (2016). Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(40), 11162–11167.

Welker, F., Ramos-Madrigal, J., Gutenbrunner, P., Mackie, M., Tiwary, S., Rakownikow Jersie-Christensen, R., Chiva, C., Dickinson, M. R., Kuhlwilm, M., de Manuel, M., Gelabert, P., Martinón-Torres, M., Margvelashvili, A., Arsuaga, J. L., Carbonell, E., Marques-Bonet, T., Penkman, K., Sabidó, E., Cox, J., … Cappellini, E. (2020). The dental proteome of Homo antecessor. *Nature*, *580*(7802), 235–238.

Welker, F., Ramos-Madrigal, J., Kuhlwilm, M., Liao, W., Gutenbrunner, P., de Manuel, M., Samodova, D., Mackie, M., Allentoft, M. E., Bacon, A.-M., Collins, M. J., Cox, J., Lalueza-Fox, C., Olsen, J. V., Demeter, F., Wang, W., Marques-Bonet, T., & Cappellini, E. (2019). Enamel proteome shows that Gigantopithecus was an early diverging pongine. *Nature*. https://doi.org/10.1038/s41586-019-1728-8

Welker, F., Smith, G. M., Hutson, J. M., Kindler, L., Garcia-Moreno, A., Villaluenga, A., Turner, E., & Gaudzinski-Windheuser, S. (2017). Middle Pleistocene protein sequences from the rhinoceros genus Stephanorhinus and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. *PeerJ*, *5*, e3033.

Welker, F., Soressi, M. A., Roussel, M., van Riemsdijk, I., Hublin, J.-J., & Collins, M. J. (2017). Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen. *STAR: Science & Technology of Archaeological Research*, *3*(1), 15–27.

Welker, F., Soressi, M., Rendu, W., Hublin, J.-J., & Collins, M. (2015). Using ZooMS to identify fragmentary bone from the Late Middle/Early Upper Palaeolithic sequence of Les Cottés, France. *Journal of Archaeological Science*, *54*, 279–286.

Wenger, C. D., & Coon, J. J. (2013). A proteomics search algorithm specifically designed for high-resolution tandem mass spectra. *Journal of Proteome Research*, *12*(3), 1377–1386.

Wiktorowicz, C. J., Arnold, B., Wiktorowicz, J. E., Murray, M. L., & Kurosky, A. (2017). Hemorrhagic fever virus, human blood, and tissues in Iron Age mortuary vessels. *Journal of Archaeological Science*, *78*, 29–39.

Wilkins, M. R., Appel, R. D., Van Eyk, J. E., Chung, M. C. M., Görg, A., Hecker, M., Huber, L. A., Langen, H., Link, A. J., Paik, Y.-K., Patterson, S. D., Pennington, S. R., Rabilloud, T., Simpson, R. J., Weiss, W., & Dunn, M. J. (2006). Guidelines for the next 10 years of proteomics. *Proteomics*, *6*(1), 4–8.

Wilkin, S., Ventresca Miller, A., Taylor, W. T. T., Miller, B. K., Hagan, R. W., Bleasdale, M., Scott, A., Gankhuyg, S., Ramsøe, A., Uliziibayar, S., Trachsel, C., Nanni, P., Grossmann, J., Orlando, L., Horton, M., Stockhammer, P. W., Myagmar, E., Boivin, N., Warinner, C., & Hendy, J. (2020). Dairy pastoralism sustained eastern Eurasian steppe populations for 5,000 years. *Nature Ecology & Evolution*, *4*(3), 346–355.

Wilson, J., van Doorn, N. L., & Collins, M. J. (2012). Assessing the extent of bone degradation using glutamine deamidation in collagen. *Analytical Chemistry*, *84*(21), 9041–9048.

Wojtowicz, A., Yamauchi, M., Montella, A., Bandiera, P., Sotowski, R., & Ostrowski, K. (1999).

Persistence of Bone Collagen Cross-Links in Skeletons of the Nuraghi Population Living in Sardinia 1500–1200 B.C. *Calcified Tissue International*, *64*(5), 370–373.

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

Xie, M., Shevchenko, A., Wang, B., Shevchenko, A., Wang, C., & Yang, Y. (2016). Identification of a dairy product in the grass woven basket from Gumugou Cemetery (3800 BP, northwestern China). *Quaternary International: The Journal of the International Union for Quaternary Research*, *426*, 158–165.

Xiong, L., Boeren, S., Vervoort, J., & Hettinga, K. (2020). Dataset on proteomic changes of whey protein after different heat treatment. *Data in Brief*, *29*, 105227.

Yamamoto, M. (1989). Enamel Hypoplasia in the Deciduous Teeth of Edo Japanese. *Anthropological Science: Journal of the Anthropological Society of Nippon = Jinruigaku Zasshi*, *97*(4), 475–482.

Yang, Y., Shevchenko, A., Knaust, A., Abuduresule, I., Li, W., Hu, X., Wang, C., & Shevchenko, A. (2014). Proteomics evidence for kefir dairy in Early Bronze Age China. *Journal of Archaeological Science*, *45*, 178–186.