Making Eggshell Visible in the Archaeological Record

John Stewart

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

University of York

Archaeology

October 2013

Abstract

Despite its presence in many types of deposit, eggshell has long been a neglected archaeological resource. The difficulty of recovering the material, combined with analytical constraints on subsequent taxonomic identification, has led to systematic underuse. This thesis will begin to address this shortcoming by pursuing two main lines of research. First, a novel identification technique based on ZooMS (zooarchaeology by mass spectrometry) and peptide mass fingerprinting is developed for eggshell. The new technique is then applied to archaeological eggshell assemblages from Anglo-Scandinavian York and Norse-era sites in the northern and western isles of Scotland. This provides new insights into egg use in these locations during this period, and raises a range of new questions regarding the use of domestic and wild resources. Second, an investigation into patterns of diagenesis in the eggshell proteins which form the basis of the technique is conducted. The principal aims of this analysis are to explore the prevalent diagenetic processes affecting eggshell proteins and amino acids, and thus to test whether it is possible to produce an estimate of the expected temporal span of the technique based on high-temperature diagenesis, and to establish the potential usefulness of the material for IcPD (intra-crystalline protein diagenesis) dating. The main outcomes of this study are (i) the development and successful application of a new taxonomic identification technique for archaeological eggshell; (ii) enhanced understanding of egg use during the Norse era in Scotland and Anglo-Scandinavian York; (iii) the observation that high-temperature diagenesis cannot be used to accurately predict peptide survival at archaeological deposition temperatures; and (iv) that avian eggshell is not a viable substrate for absolute dating using IcPD (although it may still prove useful for relative dating).

Table of Contents

List of Figures

List of Tables

List of Additional Material

Acknowledgements

This project could not have been completed without the help of a lot of people. Constraints on space limit the expression of my gratitude towards all of the people who have helped at various stages, and this token of that gratitude seems wholly inadequate. My supervision team have provided support and advice from day one up to submission, so I give my foremost thanks to Matthew Collins, Kirsty Penkman, Bone Jones and Bea Demarchi. Other members of the academic staff at York, in particular Ol Craig and Terry O'Connor, have provided advice, enlightening discussion and/or access to material and resources. Alizon Holland has been an eternally (and inexplicably) patient source of administrative

support throughout, while Rick Allen and Adam Dowle, and more lately Matt von Tersch and Keri Rowsell, have provided support in the lab. Rick, along with Marc Wadsley, has also helped in developing software and solving various computer-related meltdowns. The only thing keeping me sane during the long hours of lab-work has been the consistently high standard of entertainment provided by other inmates of the S-Block labs. These are too many to list in full, but I am particularly grateful to my office-mates Marc, Harry, Jo, Jessie and Sophy, as well as Kirsty, Molly and Ashley, for consistently making me laugh even when I didn't feel like it (and in Ashley's case, occasionally letting me win at tennis).

In terms of access to material, I was fortunate to have the invaluable collaboration of Isla Gladstone at Yorkshire Museums Trust, and I remain grateful to her for making time in her insane schedule to provide me with what I needed. In the museum world, I am also grateful to Emma Bernard, Doug Russell, Jan Bolding Kristensen and Jane Sidell for their advice, practical support, and/or provision of material. This project would be of limited use if archaeological material had not been available from the outset, when there was no track record and little reason (according to one prominent member of the supervision team, who shall remain nameless) to presume that the technique would even work. Various people at YAT were of great service from the outset, including Rachel Cubitt and Toby Kendall, but most notably Bone Jones, whose relentless enthusiasm, interest and positivity (~~endured~~ valued by all who know him), made the whole thing seem not only possible but worthwhile.

The development of the project could not have occurred without the material and advice provided by a wide range of people in a wide range of places. My highest regards are given in particular to James Barrett in Cambridge, Niall Sharples and Julia Best in Cardiff, Gerry Bigelow at Bates College and David Griffiths in Oxford for providing the material which forms the basis of the investigation of Norse-era sites in Scotland. I am also indebted to all of the following people for providing material for analysis or advice: Colleen Batey, Beverley Ballin Smith, Jacqui Mulville, Dale Serjeantson, Erika Gal, Gitte Gotfredsen, Anja Prust, Christine Schuh, Beata Tugya, Veronique Laroulandie, Dimitra Mylona, Jane Thompson and Julie Bond.

The highest level of gratitude is reserved for those who kept my feet on the ground throughout, and ensured that I remained connected to the real world during the more intense moments of the past 3 years. My parents and my brother have helped to keep my mind off work at least some of the time, as have various friends and family members all over the country. The source of strength throughout this has been Emily, who has been my anchor for the past few years and whose contribution I cannot begin to put into words.

Declaration

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

John Stewart

Chapter 1: Introduction

Section 1: Brief History of Egg Use

1.1.1: Introduction

1.1.1.1: Overview

The first chapter provides the background information for this project. The first section summarises the current state of knowledge regarding egg use in the past, and outlines important developments in avian domestication, taphonomic factors acting on eggshell recovery, and existing and potential archaeological applications of eggshell. Subsequent sections will give detailed accounts of eggshell structure, deposition and proteins.

1.1.1.2: Why do people use bird eggs?

The eggs of birds are a nutritionally rich and versatile food source. Their main nutritional value is in their high protein content; in chicken eggs, for example, protein forms 12.6% of the total weight. This is comparable to the protein content of red meat (e.g. beef, 14% protein). Egg proteins contain all of the essential amino acids (i.e. those not synthesised within the body), a range of vitamins and minerals, and important trace elements. They are available at predictable times and locations, often in large densities (e.g. Sidell, 1993; Baldwin, 2009, 2010; Serjeantson, 2009). Although eggs of wild species and most domestic species are a seasonal resource, they could be stored for a number of months before the emergence of modern refrigeration technology (e.g. in waterglass, a combination of salt-water and peat ash; or in isinglass, derived

from the swim bladder of cod) (Serjeantson, 2009; Baldwin, 2010). Some domestic species can produce eggs throughout the year (Section 1.1.3), although it is not clear at what point this ability emerged. Today, eggs are a relatively cheap and mass-producible source of protein, and the eggs of the domestic chicken (Section 1.1.3.2) are an important food source in many parts of the world. For example, according to official United Nations estimates, the global average chicken stock is around 19 billion animals (many of which will represent laying hens); this is over 13 times more than the next most numerous domestic animal (UN Food and Agriculture Organisation, 2011).

Eggs have also held symbolic meaning in many cultures (most often as symbols of fertility and/or rebirth), and have been used in various burial rituals (Sidell, 1993a,b; Stadelman, 2000; Serjeantson, 2009). Applications of the eggshell have included use as liquid containers and as raw material for jewellery, while the whole egg has been used in clarification of wine, and as feed and/or tonic for livestock (Baldwin, 2009, 2010; Serjeantson, 2009; Texier et al., 2010). Some cultures have also observed taboos which prevent consumption of eggs (Serjeantson, 2009).

1.1.2: Collection and use of wild eggs in the past

While it has been established that ostrich eggs were being used in southern Africa by at least 100 millennia before present (Brooks et al., 1990; Texier et al., 2010), very little is known about egg use in most past societies. Eggshell fragments from archaeological sites can be very difficult to recover, and even

more difficult to taxonomically identify (Section 1.1.4); this, combined with the lack of documentary evidence for the vast majority of past cultures, limits the current state of understanding regarding most facets of past egg use. At present, the contribution of bird eggs to most past diets can only be speculated upon. Consumption of wild bird eggs was probably once commonplace, particularly among hunter-gatherers and coastal and island communities, but has now ceased in most parts of the world (Serjeantson, 2009). The near-complete cessation of wild egg collection is largely attributable to the prolific laying capabilities of the domestic chicken (Section 1.1.3.2). Collection of wild seabird eggs was widespread in many regions until around a century ago, and remains an important economic activity for some tribal peoples, including many aboriginal Australian and native North American groups (e.g. Hunn et al., 2003; Serjeantson, 2009). Collection of seabird eggs is either still practiced or has been recently practiced in some remote corners of northern Europe, including the Faroe Islands and many British and Irish islands (Baldwin 2009, 2010; Serjeantson, 2009). On the Isle of Man, for example, collection of gull eggs was revived during World War II food shortages, was commonly practiced until the late 1970s, and remains legal (subject to permit) to this day (Baldwin, 2009, 2010; Serjeantson, 2009).

The main factors determining the extent to which the eggs of a particular species were exploited are likely to include its breeding habits, and the size and number of eggs available; cultural factors (e.g. high status attached to certain foods), symbolic significance, and preferred tastes are also certain to have played roles (Serjeantson, 2009). In some cases, documentary evidence can

inform questions of the types of eggs used. For example, the Romans kept many types of bird, including several breeds of chicken, goose, peacock, pheasant and duck, and are believed to have consumed (to varying degrees) the eggs of all of these (Keepax, 1981). This level of documentary information is exceptional, and cultural factors affecting egg exploitation are difficult to reconstruct for most societies.

Setting cultural factors aside, colonial nesters are likely to have provided a more energetically efficient source of eggs than solitary nesters. This places emphasis on certain seabirds as egg providers; these often nest in high densities in predictable times and locations. Surviving documentary evidence shows that these provided the majority of wild eggs collected in many British and Irish coastal and island communities until fairly recently (Kightly, 1984; Baldwin, 2009, 2010; Serjeantson, 2009). Gull eggs were often collected in the highest volumes due to their high breeding densities and the relative accessibility of their nests (Baldwin, 2009, 2010). Many British-breeding seabirds, including puffins, shearwaters, gulls, guillemots, razorbills and gannets (but not fulmars), will re-lay up to 3-5 clutches per breeding season if their eggs are removed; this could facilitate repeated exploitation of the eggs (and young) of these species (Baldwin, 2009, 2010). Terrestrial species have also been used in the recent past in this part of the world; for example, the eggs of lapwing and rook were routinely eaten on the Isle of Man until the early 20th century (Baldwin, 2010).

While the accessibility of the nests is another potentially important factor, documentary records show that people have routinely collected eggs from steep sea cliffs in different locations around the U.K. and Ireland over a number of centuries (Kightly, 1984: Baldwin, 2009, 2010; Serjeantson, 2009). For some communities egg collection was a subsistence activity, while for others eggs were a cash crop to be sold at market (Baldwin, 2009). This was a dangerous occupation; the collector was lowered over the cliff with a rope tied around his waist, and a basket in which to collect eggs (secured either to a pole planted in the ground, or held by a team-mate). John Ray described techniques in use on the Isle of Man in the late 1670s which were still in use in Yorkshire during the 1940s (Kightly, 1984: Baldwin, 2010). It seems reasonable to propose that this practice might extend much further into the past than the relevant documentary records do.

These examples relate only to the relatively recent history of egg collection and use in the U.K. For people in other parts of the world, and in the deeper past, very little is known about egg use. At present, historic and ethnographic documents (where available) can shed light on egg use in particular times and places, while zooarchaeology and genetics can show when different domestic species became available in different areas. However, there is very little contribution at present from the archaeology of the egg itself; the core aim of this project is to derive an analytical framework which will allow this enormous gap in knowledge to be addressed.

1.1.3: Bird domestication and dispersal

1.1.3.1: Overview

As the sites considered in a later chapter (Chapter 4) are all located in Europe, and there has been generally a greater concentration of archaeological effort on that continent, the arrival of avian domesticates in Europe is considered here in more detail than in any other continent. It is likely that egg use changed dramatically following the domestication and dispersal of a few avian species. By far the most economically important of these has been the domestic chicken (*Gallus gallus domesticus*), which was also domesticated far earlier than any other species of bird. In terms of egg production, other important domestic species have included geese and (to a lesser extent) ducks; other commercially important species in various parts of the world have included turkeys, quail and pigeons (Brant, 1998; Beacham & Durand, 2007). Only chicken, goose and duck are summarised here due to a lack of detailed information regarding the domestication or keeping of other bird species; turkeys are not included as it is not currently clear whether it is possible to separate wild from domestic birds based on zooarchaeological evidence (Olsen, 2012).

1.1.3.2: Chicken domestication and dispersal

The vast majority of eggs consumed by modern people are provided by the domestic chicken (*Gallus gallus domesticus*); this is one of the most important (and by far the most numerous) domestic animal (UN Food and Agriculture Organisation, 2011). Their global distribution can be attributed to deliberate translocation; chickens (and their wild progenitors) are poor flyers, cannot swim, and have small home ranges (Storey et al., 2012). Although modern domestic

fowl are predominantly reared for food production, this may not have been the only (or even the main) motivation for their initial domestication (Blench & MacDonald, 2012). Chickens and their eggs have often held symbolic and ritual significance (for example, eggs are present as grave goods in widely dispersed cultures), and have also been widely used in cockfighting and for decorative purposes (Blench & MacDonald, 2012). In Japan, for example, there is little evidence of chickens being used for food production prior to the 19th century A.D. (Blench & MacDonald, 2012).

The progenitors of the domestic chicken are the wild junglefowl (*Gallus* spp.) of southern and southeastern Asia (e.g. Fumihito et al., 1996; Kanginakudru et al., 2008). Prior to the emergence of valid molecular techniques, a prolonged debate centred around whether there was a single origin (monophyly) of modern chickens, or whether these arose from a range of progenitors in a range of areas (Fumihito et al., 1996; Liu et al., 2006; Kanginakudru et al., 2008; Storey et al., 2012). Although it is possible that all variation in modern chicken mtDNA can be traced to wild species extant in modern Thailand (Fumihito et al., 1996), two centres of domestication in southern and south-eastern Asia are now generally recognised (West & Zhou, 1988; Liu et al., 2006; Tixier-Boichard et al., 2011; Storey et al., 2012). One of these was in Neolithic Southeast Asia at least 8000 years before present; the other was in the Indus Valley prior to 4500 years before present (West & Zhou, 1988). Consensus has not been reached on whether modern chicken breeds are derived from a single species of wild junglefowl (*G. gallus*) (e.g. Fumihito et al., 1996) or from hybridisation of several species (e.g. Kanginakudru et al., 2008; Storey et al., 2012); the most recent

genetic evidence suggests the latter. DNA evidence shows that birds from Southeast Asia were transported as part of the Austronesian expansion into the Pacific Islands (and ultimately to South America); they also travelled along trade routes to east Asia, the Middle East, Europe and probably Africa (West & Zhou, 1988; Storey et al., 2007, 2012); it seems that the Indus Valley was bypassed during this process, although birds domesticated there were spread along trade routes to east Africa, and their descendants are still found in the Indian subcontinent today (Coltherd, 1966; West & Zhou, 1988; Manaseryan & Balyan, 2002; Liu et al., 2006; Tixier-Boichard et al., 2011; Storey et al., 2012).

Chicken bones appear in southeast Europe during the late Neolithic/early Bronze Age (Kysely, 2010). In other parts of Europe, the chicken was a rather late addition to the core group of domestic animals; it does not appear in central, northern and western Europe until the Bronze/Iron Age transition (8^{th} – 9^{th} century B.C.) (West & Zhou, 1988; Kysely, 2010). The route(s) by which chickens came to central, northern and western Europe are not fully established, but there are several non-mutually-exclusive possibilities: across the Balkans from Bulgaria or Greece; across Ukraine or the Black Sea; or via Iberia during the period of Phoenician colonisation (9^{th} century B.C. onwards) (Kysely, 2010). Although precise dates of arrival are not available in many areas, the evidence suggests that they did not arrive in north-western Europe until the Middle Iron Age; in Sweden, for example, they do not appear until the 1^{st} century B.C. (Tyrberg, 2002). Bone assemblages from many Eurasian archaeological sites suggest that these rapidly became the dominant domestic bird in most regions following their introduction (Albarella & Thomas, 2002; Hamilton-Dyer, 2002;

Makowiecki & Gotfredsen, 2002; Serjeantson, 2002, 2009; Tyrberg, 2002; Baker, 2008; Benejaru et al., 2008; Thys & van Neer, 2008).

It is unclear what their initial purpose was; the presence of eggs as grave goods at many central European Iron Age sites suggests ritual significance, while the high ratio of males to females recovered at some Bronze Age and Iron Age sites suggests that cockfighting may have been a greater motivation than egg production (Kysely, 2010). It is likely that the early breeds present in Europe still had seasonal laying patterns governed by light regimes, and were not as productive in terms of egg yields as modern breeds (although the point at which laying was decoupled from natural light regimes is unknown). Modern breeds can produce upwards of 250 eggs per year, and documentary evidence suggests that they may have done so in England since at least the late medieval period ($12^{th} – 13^{th}$ century) (Stone, 2006).

1.1.3.3: Goose domestication and dispersal

Geese are less productive than chickens in terms of egg yields, but bone assemblages show that they were kept at many archaeological sites. It is unlikely that the primary focus of goose husbandry was ever egg production, as a goose will lay a maximum of 30-40 eggs in one year; the primary focus is likely to have been production of meat and/or down, while in Russia (and possibly in other regions) they were also bred for fighting (Serjeantson, 2002; MacDonald & Blench, 2012). The earliest domestication event is unknown, but at least six species of goose have been independently domesticated at some point

(MacDonald & Blench, 2012). The greylag goose (*Anser anser*) is the progenitor of most modern domestic geese; the only exception is in China, where the swan goose (*Anser cygnoides*) was the progenitor of modern domestics (Serjeantson, 2009; MacDonald & Blench, 2012).

The centre from which modern non-east-Asian domestic geese are derived was located in Egypt (MacDonald & Blench, 2012). The earliest tangible evidence of domestic geese takes the form of captive animals depicted on material dating from the Egyptian Old Kingdom (3rd millennium B.C.); the oldest evidence in the form of actual bones dates from the Middle Kingdom (early 2nd millennium B.C.). The point at which domestic geese arrived in Europe has not been conclusively determined, but Greek and Macedonian literary sources mention them during the first millennium B.C. (Harper, 1972; Serjeantson, 2002; MacDonald & Blench, 2012). The zooarchaeological evidence suggests that domestic geese only became widespread in Europe during the Roman period (MacDonald & Blench, 2012). They first appear in the British archaeological record during the 1st century AD (Harrison, 1980), and are present at some Anglo-Saxon and Viking/Norse sites in the U.K. (Dobney & Jaques, 2002; MacDonald & Blench, 2012). Domestic geese were widespread in Europe by the late medieval period, having been widely dispersed by the Normans, and are a distant second in predominance to chicken in most bone assemblages from this period (Albarella & Thomas, 2002; Hamilton-Dyer, 2002; Makowiecki & Gotfredsen, 2002; Serjeantson, 2002; Tyrberg, 2002; Baker, 2008; Bejenaru et al., 2008; Thys & van Neer, 2008).

Under traditional management techniques, geese have some advantages over chickens; they can be driven to market rather than carried; they can survive on poorer food than other domestic birds; and provide more meat per bird than chickens (MacDonald & Blench, 2012). However, the almost ubiquitous dominance of chicken in most cases (some medieval towns being an exception) highlights their relative disadvantages; far lower fecundity and egg yields (MacDonald & Blench, 2012). Domestic geese have declined in many areas in the last century or so, as chickens and turkeys are more amenable to modern factory and battery farming techniques (MacDonald & Blench, 2012).

1.1.3.4: Duck domestication and dispersal

Ducks were domesticated in different regions from a number of wild species; in Europe, the mallard (*Anas platyrhynchos*) was probably the main wild progenitor (Luff, 2012). Given the wide distribution and commonness of most *Anas* species (including mallards), it is likely that ducks were domesticated independently in different regions rather than deriving from a core area of domestication (Luff, 2012). Today, ducks account for less than 1% of the U.K. poultry industry; this is in contrast with East Asia, where over 75% of modern domestic ducks are found (Luff, 2012). The advantages of ducks over chickens are that they grow extremely fast, and are almost equally as fecund under favourable conditions; in hot, humid climates (such as those prevalent in much of East Asia) ducks can produce higher egg yields than chickens (Luff, 2012). Their disadvantages are that they require more food provision than either chickens or geese, are much less productive in northern European climates, and may have been more difficult to domesticate (particularly in cold conditions) (Luff, 2012).

It seems that ducks were domesticated in Europe much later than in China, where domestic ducks may have been present for at least 3000 years (Harper, 1972; Luff, 2012). In contrast, the earliest documentary evidence of duck husbandry in Europe comes from Rome at the end of the first millennium BC, and true domestication may not have occurred until the middle ages (Harper, 1972). There is no evidence to suggest that keeping of ducks was at all common in the wider Roman Empire at this time, and very limited evidence that it was even present outside this area (e.g. Tyrberg, 2002; Luff, 2012). Tame or domestic ducks were introduced to the U.K. by the Romans and may form a significant component of some Anglo-Saxon bone assemblages, although distinguishing domestic from wild types can be extremely difficult (Harrison, 1980; Dobney & Jaques, 2002; Luff, 2012). The first direct mention in the non-Greco-Roman historical literature comes from late 8th century AD French and German legal documents, which also suggest that these were much rarer than domestic chicken or geese (Harper, 1972; Luff, 2012). Historical documents suggest that ducks were bred at high-status sites in southern Poland from the 15th century, but were probably not truly domesticated there until the 19th century (Mackowiecki & Gotfredsen, 2002). Probably domestic ducks have been identified in medieval sites all over Europe (Albarella & Thomas, 2002; Hamilton-Dyer, 2002; Mackowiecki & Gotfredsen, 2002; Tyrberg, 2002; Baker, 2008; Bejenaru et al., 2008; Thys & van Neer, 2008).

1.1.4: Eggs in the archaeological record

1.1.4.1: Preservation and recovery of eggshell

Although the shell membrane can survive under exceptional preservation conditions, the hard, calcareous shell is usually the only part of the egg preserved over archaeological timescales. The archaeology of the egg must focus upon this material. Eggshell is predominantly composed of calcium carbonate (more detail on structure is given in section 1.2), and can survive for long periods at neutral to alkaline pH. Anaerobic conditions are also favourable to preservation, while acidic and/or wet conditions are unfavourable. Eggshell can be preserved for a very long time, and is often found in calcareous deposits which also contain mollusc shell (Keepax, 1981). Most existing finds were located in protected environments such as caves, wells, pits and ditches, although some exceptions have occurred (Keepax, 1981; Serjeantson, 2009). Where eggshell is preserved, it is often in tiny fragments which may not be instantly recognised. The fine sieving (at least < 5mm, and often < 2mm) which is often required to extract eggshell can place a limit on recovery of fragments; time and/or resource constraints have precluded such time-consuming and labour intensive sieving at many sites. Following deposition, it is likely that the thicker eggs of larger taxa such as geese will be preserved for longer than thinner eggs such as those of chicken or duck.

There are also a range of other taphonomic factors relating to the manner in which the people who originally used the egg handled and disposed of the material. These may differentially affect recovery of different types of eggshell in different areas. Although these are often neither testable nor provable, their

potential effects must be considered when interpreting the material recovered from archaeological deposits. For example, disposal patterns of eggs of different species may have differed for cultural reasons, or the eggs may have been consumed away from the main area of habitation (e.g. Baldwin, 2009). Eggs which were used for a specific purpose rather than for everyday consumption may also have been treated differently. All of these factors could affect the location in the ground of different types of eggshell, and could affect recovery and confuse interpretation.

1.1.4.2: Identification of archaeological eggshell

The eggshell of ratites (ostriches and allies) is readily identified in the hand due to its extraordinary thickness. The correct order can then usually be deduced on the basis of geographical location; the major extant and recently extinct ratite orders (following the phylogeny of Hackett et al., 2008) are (or were) endemic to distinct areas. The higher likelihood of preservation of ratite eggshell, combined with relative ease of identification, has led to it dominating archaeological and geochronological applications of eggshell (e.g. Brooks et al., 1990; Miller et al., 2000; Crisp et al., 2013). In many parts of the world ratites have never been coeval with humans (or have not remained so for long), and even where they were people may also have used non-ratite eggs. Non-ratite eggshell is more difficult to identify; differences in thickness between taxa are not as marked or as consistent as between ratites and non-ratites, and fragments are often extremely small and poorly preserved (e.g. Sidell, 1993).

The majority of eggshell fragments which are recovered from archaeological sites are placed in a bag marked 'avian eggshell' and archived - no further identification is usually attempted (e.g. Sidell, 1993). This is because there are few existing techniques for identification of such fragments. The most successful yet relies on scanning electron microscopy (SEM) to compare a range of parameters relating to the internal structure of eggshell fragments with a reference collection (Keepax, 1981; Sidell, 1993). Although this has been the only reasonably robust technique available for a long time, it is time and labour-intensive; high-throughput data generation is not possible. It is also subjective; interpretation may vary depending on the researcher performing the analysis. Recent research on moa eggshell suggests that (at least in some taxa) some of the parameters used are not in fact taxonomically diagnostic; for example, intra-species variation in shell thickness (one of the major taxonomic identifiers in the SEM technique) often exceeds inter-species variation (Oskam et al., 2011). Shell thickness also varies in domestic species (and also presumably in wild species) according to breed and diet, and between different parts of a single egg; this suggests that these internal parameters are not a reliable taxonomic indicator for heavily fragmented archaeological eggshell (Keepax, 1981; Sidell, 1993; Serjeantson, 2009).

In larger fragments shell curvature can be used, although this too is problematic; shell shape is fairly distinct in a few orders, but in most intra-specific variability can equate or exceed inter-specific variability (Keepax, 1981; Gill, 2000; Serjeantson, 2009). Where whole eggs are found measurement of shell length and diameter can be used to derive taxonomic identification, but the likelihood

of preservation of whole eggs is slight (although this has occurred in the U.K. on at least two occasions) (e.g. Keepax, 1981). Even where these are found, egg size provides low resolution; the degree of inter- and intra-specific variability is often too great to allow accurate identification (e.g. Keepax, 1981). The pigments in the shell cuticle tend not to survive in archaeological samples, and are in fact often lost fairly rapidly (100-150 years) from eggs held in museum collections (their survival in the ground has not been quantified). A range of approaches which have been tentatively described as useful in modern samples, but which are not applicable to archaeological samples as they assume no structural deterioration, have been reviewed in detail elsewhere (Keepax, 1981). Neural imaging software has also been used to identify eggshell fragments with a reasonable degree of accuracy (around 70%), but this technique faces similar problems relating to time and labour intensity and has not developed beyond the initial pilot study (Eastham & Gwynn, 1997).

1.1.4.3: Other applications of archaeological eggshell

Aside from taxonomic identification, amino acid dating has been the most predominant application of preserved ratite eggshell (e.g. Brooks et al., 1990; Miller et al., 2000). A range of other applications exist; for example, progressive changes in shell microstructure during embryogenesis have the potential to discriminate between eggshell originating from food waste and eggs which were allowed to hatch (Sidell, 1993; Beacham & Durand, 2007). These microstructural changes have been used to constrain the date of emergence of turkey husbandry in the American south-west (Beacham & Durand, 2007). The location of rhea eggshell deposits in Argentina in relation to the location of main

residential areas has also been used to argue for the importance of regional-scale archaeological research in order to understand subsistence patterns in that region (Medina et al., 2011). Environmental researchers have recently used preserved seabird eggshell to determine changes in oceanic mercury levels over the past 700 years (Xu et al., 2011). The potential of eggshell as a material for study is beginning to be realised in a diverse range of fields.

1.1.4.4: Eggs in the archaeological record: future applications

The major challenge facing archaeological eggshell research has always been that there is no robust, high-throughput identification system capable of analysing the large volume of material recovered from many deposits. The development of such a system is the core objective of this project. This system, once developed, can be applied to questions regarding wild and domestic resource use. For example, people may have taken eggs from a particular species of wild bird, while taking another species for meat (Serjeantson, 2009). In most cases, the relative contributions of wild and domestic eggs to the diet are completely unknown. The diversity of wild resources used by past societies can also potentially be addressed using this system. Seasonality of occupation at certain sites can also be addressed, as most birds have well defined breeding seasons (although the fact that eggs can be stored may cloud this issue to some extent). The relative status attached to different types of egg may be highlighted, and consideration of the favoured breeding sites of different birds may imply the use of certain types of technology in egg collection (Sidell, 1993). Most wild eggs are likely to have been collected within a relatively short distance of the site; the ecological preferences of different bird species may

then inform questions regarding the terrain surrounding the site (Sidell, 1993). The usefulness of amino acid racemisation dating has been established in ratite eggshell (e.g. Brooks et al., 1990; Miller et al., 2000), but the kinetics of diagenesis in eggshell biomolecules remains to be established for non-ratite species; only one pilot study has been published to date (Clarke et al., 2007). A subsequent chapter will investigate the potential of biomolecular diagenesis in the eggshell of domestic bird species (Chapter 5).

1.1.5: Conclusions on this section

This section has provided initial background information to the project, and has introduced the core objective and rationale of the work described in subsequent chapters. In order to properly use eggshell in an archaeological setting, an understanding of the material is required. The following section begins to consider the material under study by providing a description of the internal structure of eggshell.

Section 2: Eggshell Structure

1.2.1: Overview: the amniotic egg

The emergence of the amniotic egg was central to vertebrate colonisation of terrestrial habitats, and must be considered one of the most important events in evolutionary history. In birds, embryonic development takes place inside a structure which provides protection from the physical environment and microbial attack, regulates water and gas exchange with the external environment, and provides vital nutrients to the developing embryo (Simkiss, 1968; Arias et al., 1993; Nys et al., 2004). The eggshell must have sufficient mechanical strength to fulfil its protective functions, but this must be balanced with facilitating water and gas exchange, and ultimately allowing the emerging juvenile to break through it. This section will provide a summary of the basic structure of avian eggshell.

1.2.2: Eggshell structure

1.2.2.1: Basic structure

Eggshell consists of both organic and inorganic phases, and across diverse genera comprises around 10% of the total egg weight (Nys et al., 2004). The inorganic phase is formed almost exclusively of calcium carbonate ($CaCO_3$) in the form of calcite, which constitutes 95-96% of the shell weight (Arias et al., 1993; Nys et al., 2004). Although small quantities of aragonite and vaterite (other forms of $CaCO_3$) have also been observed occasionally in some species, these are considered anomalous (Becking, 1975; Sidell, 1993; Dennis et al., 1996). The calcite ultrastructure is stabilised by an organic matrix of proteins, many eggshell

proteins also play key roles in directing different stages of eggshell formation (Sections 1.3 & 1.4) (Gautron et al., 1997; Hincke et al., 2010). The organic fraction of the calcified layers (Figure 1.1) contains various proteins, glycoproteins and proteoglycans (discussed in detail in section 1.4), and comprises 2 - 3.5% of the total eggshell weight (Hincke et al., 1995; Nys et al., 2004; Mann & Siedler, 2006).

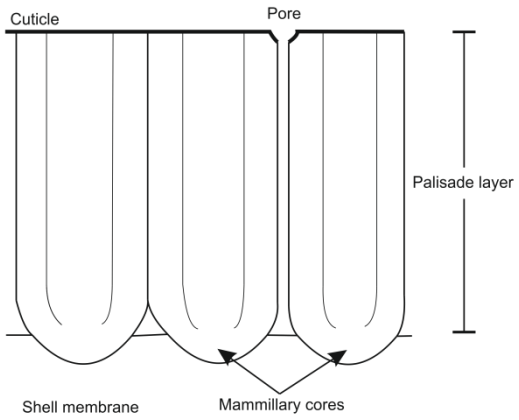

Figure 1.1: Generalised schematic of the internal structure of eggshell (adapted from Stewart et al., 2013). No measurements are given as these vary between taxa.

1.2.2.2: Shell membrane

At the innermost shell surface, a non-calcified bi-layered membrane separates the egg white from the calcified layers of the eggshell (Simkiss, 1968; Sidell, 1993; Mann, 2008). This membrane is deposited in the isthmus external to the peri-albumen layer which surrounds the egg white (Arias et al., 1991; Dennis et al., 1996; Nys et al., 2004). Most of the membrane (70 – 75%) is formed of proteins and glycoproteins (Wong et al., 1984: Arias et al., 1991). Lysozyme C (60% by weight) has been proposed as the major protein component of the

membrane, while collagen is also present (Wong et al., 1984; Mann, 2008). The outer shell membrane is strongly adhered to the inner surface of the shell, and is extremely difficult to remove mechanically (Sidell, 1993). Intact membranes are fundamental to successful eggshell deposition, as (among other functions) they prevent inward mineralisation (Nys et al., 2004). Recent high-throughput proteomic techniques have identified up to 137 protein constituents of these membranes in chicken eggshell, many of which also occur in other eggshell and body compartments (Mann, 2008). The function of these in the membrane (if any) mostly remains unknown (Mann, 2008). The shell membrane is occasionally preserved in extremely organic-rich deposits such as cesspits, but is not commonly recovered at most sites.

1.2.2.3: Mammillary layer

The most internal crystalline layer is deposited in a non-uniform manner on the outer shell membrane as small radially-arranged polycrystalline calcite clusters (mammillae) (Nys et al., 2004). These are deposited in a quasi-periodic manner on the external surface of the shell membrane, before small crystals grow upwards and outwards until they fuse and the membrane surface is covered; at this stage deposition of this layer terminates (Tyler & Fowler, 1978).

The mammillae comprise a base plate, calcium reserve body (CRB), a CRB cover and a crown; the CRB functions as a source of readily mobilised calcium for bone mineralisation in the developing embryo (Bond et al., 1988; Dennis et al., 1996; Reynolds et al., 2004; Beacham & Durand, 2007; Chien et al., 2009; Osterstrom

& Lilja, 2012). The fibres of the outer shell membrane penetrate the cones of the mammillary layer, which then act as sites of initial crystal nucleation for the large, oriented calcite crystals of the palisade layer (Robinson & King, 1963: Arias et al., 1993; Panheleux et al., 1999; Nys et al., 2004). The mammillary layer is divided by some authors into two components; the basal caps and the cone layer (e.g. Becking, 1975). It is possible that mammillary density correlates with breeding biology, with differences observed between altricial and precocial species; the mechanism(s) underlying this correlation are not clear (Osterstrom & Lilja, 2012).

1.2.2.4: Palisade layer, vertical crystal layer & cuticle

The palisade layer spans from the mammillary crowns to the internal surface of the cuticle (Figure 1.1), and comprises large calcite crystals aligned with their *c*-axes perpendicular to the shell surface (Nys et al., 2004). This layer provides the mechanical strength required by the eggshell. As the dense crystalline structure of the palisade layer precludes permeation of water and gases from the external environment, a porous network extends from the outer cuticle to the less dense mammillary layer (Simkiss, 1968; Nys et al., 2004). The dimensions and tortuosity of this system vary between taxa depending on their differing requirements, and can adapt rapidly to the differing conditions of novel environments (Tullet, 1975; Stein & Badyaev, 2011). In most taxa, a thin layer of small vertically aligned crystals lies at the external surface of the palisade layer (Keepax, 1981). Exterior to this is a non-calcified cuticle comprising proteins, lipids and polysaccharides (Simkiss, 1968; Lammie et al., 2006). This cuticle is the primary frontier of water and gas exchange between the embryo and the

environment, and is necessarily a region of very low chemical and biological activity (Nys et al., 2004). The eggs of some species (e.g. the gannet) do not have a cuticle but have a calcareous coating instead (Sidell, 1993).

1.2.3: Conclusions on section 2

This section has introduced the internal structure of the material which is the focus of this project. The following section summarises the current state of knowledge regarding the manner in which this material is deposited, and the role(s) of biomolecules in that process.

Section 3: Biomineralisation & Eggshell Deposition

1.3.1: Overview

$CaCO_3$ has been produced by metazoans for at least 550-600 million years (Grotzinger et al., 2000), and the oldest evidence for production by unicellular organisms dates to around 3.5 billion years ago (e.g. Lepot et al., 2008; Allwood et al., 2009). Today, a genetically diverse group of organisms, including many types of arthropods and molluscs, as well as fish and birds, produce some form of biogenic $CaCO_3$. This tends to form exoskeletal material in invertebrates, while in vertebrates it functions as the building material of the amniotic eggshell in birds and in the construction of the otolith. $CaCO_3$ can occur in three crystalline phases (calcite, aragonite and vaterite), and also as three hydrate non-crystalline phases (amorphous (ACC), monohydrate and hexahydrate) (Figure 1.2). The $CaCO_3$ in avian eggshell is almost exclusively calcite (Dennis et al., 1996; Fu et al., 2005; Xie et al., 2005). Mollusc shells tend to be composed of aragonite and/or calcite, while fish otoliths are composed of aragonite; in some marine arthropods and echinoderms amorphous $CaCO_3$ (ACC) is utilised, most commonly as a transient phase (Addadi et al., 2003; Weiner et al., 2005; Bentov et al., 2010).

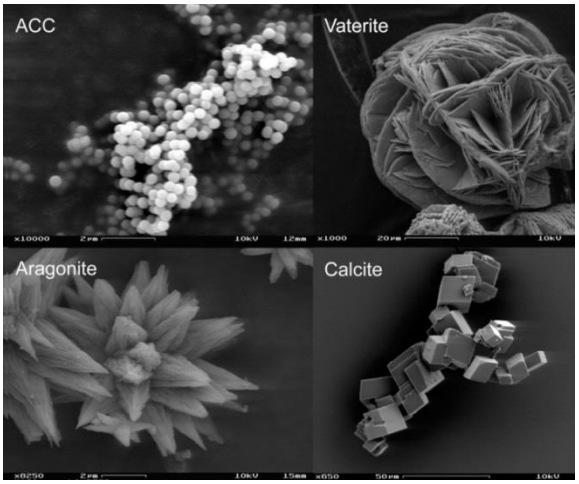

Figure 1.2: Transmission Electron Micrograph (TEM) images of crystalline calcium carbonate polymorphs, and amorphous calcium carbonate. Reproduced from: http://www.ruhr-uni-bochum.de/sediment/pictures/CaCO3_web1.jpg (accessed 01/08/13)

Although avian eggshell is the focus material of this project, research on biomineralisation processes in other organisms is potentially of direct relevance to understanding the processes of eggshell deposition; it is probable that at least some of the general pathways of biomineralisation are conserved across a range of genetically diverse organisms (Weiner & Addadi, 1997; Addadi et al., 2003; Fu et al., 2005; Weiner et al., 2005). The diversity of crystalline structures in nature highlights the fact that organic activity often plays a major role in their determination; structural differences between taxa are likely to derive from the effects of different organic fractions on growing crystals (e.g. Stolarski & Mazur, 2005; Weiner et al., 2005). Characterisation of the processes by which organisms exert exquisite control over crystal development is of potentially major significance to major growth industries such as nanotechnology and material science, as well as to biological sciences (Gower & Odom, 2000; Sarikaya et al.,

2003; Gebauer et al., 2008, 2009, 2010; Krattiger et al., 2010). A synopsis of the current state of understanding regarding crystal nucleation, morphogenesis and crystallisation in biominerals is provided below. All of this is potentially relevant to eggshell deposition, although biomineralisation in marine invertebrates has been more extensively studied (e.g. Aizenberg et al., 1996; Marin et al., 2005, 2008; Weiner et al., 2005). This section also introduces the intra-crystalline protein fraction, which is crucial to the methodologies developed and used in subsequent chapters.

1.3.2: Initial stages of crystallisation

1.3.2.1: The ACC phase

The nucleation of a transiently stable ACC phase is the initial stage of crystallisation in many biominerals (Addadi et al., 2003; Weiner et al., 2005; Pouget et al., 2009). Although ACC can exhibit short-range atomic order, lack of register between adjacent molecules precludes the emergence of long-range order and crystallinity (Aizenberg et al., 1996; Weiner et al., 2005; Gueta et al., 2007; Lam et al., 2007). ACC can be thermodynamically stable under certain conditions; however, at larger nanoparticle sizes and/or higher ion concentrations crystalline polymorphs become preferred (Freeman et al., 2010). ACC often occurs in nature as a transiently stable precursor to crystallisation, and there is evidence to suggest that this is the pathway leading to calcite precipitation in the mammillary layer of bird eggshell (Lakshminarayanan et al., 2006; Voinescu et al., 2007; Freeman et al., 2010; Gebauer et al., 2010). ACC also occurs as a metastable polymorph in its own right in a range of marine species

(Aizenberg et al., 1996; Addadi et al., 2003; Wilt et al., 2003; Weiner et al., 2005; Gebauer et al., 2008, 2010; Bentov et al., 2010).).

1.3.2.2: Metastable vs. transiently stable ACC: roles of organic molecules and hydration

The pathways leading to maintenance of metastable ACC (used as structural material in some marine invertebrates) and transient stabilisation of ACC during early crystallisation are different (Addadi et al., 2003; Weiner et al., 2005; Politi et al., 2008). Maintenance of metastable ACC is determined by the activity of specific ions and/or amino acid residues, although the precise mechanism(s) often remain unclear (e.g. Weiner et al., 2005). For example, specific phosphoprotein constituents inhibit crystallisation, and may play a crucial role in the maintenance of stable ACC in decapods (Bentov et al., 2010). Isolated phosphoserine and phosphothreonine residues exhibit high calcium binding capacity, which causes them to adsorb onto nascent nuclei and retard crystal growth until the ACC pool becomes supersaturated (Addadi et al., 2003; Bentov et al., 2010). At this point ACC will transform into more stable crystalline forms *in vitro* (Addadi et al., 2003).

Macromolecule concentrations *in vivo* are probably too low to support this as a general mechanism of ACC stabilisation; it is more likely that a microenvironment which does not favour crystal nucleation is formed (Addadi et al., 2003; *c.f.* Gebauer et al., 2009). As opposed to the transiently stable ACC phase observed in precipitation of crystalline phases, stable ACC is significantly

hydrated; the presence of water molecules around calcium ions may preclude their reconfiguration into anhydrous crystals (Addadi et al., 2003; Weiner et al., 2005). These examples serve to introduce the importance of the interaction between the mineral and molecular environments.

1.3.3: Crystal nucleation

1.3.3.1: Nucleation theory

Nucleation of both the polycrystalline aggregates of the mammillary layer and the large calcite crystals of the palisade layer are of fundamental importance to successful eggshell deposition. Understanding the mechanisms of crystal nucleation is crucial to determining potential organic roles in eggshell deposition. Two competing (but not entirely mutually exclusive) theories of crystal nucleation are currently in circulation, and the debate between them is not yet resolved. Traditional nucleation theory holds that nucleation proceeds by reversible addition of free ions to a pre-critical cluster until a critical size is reached and the cluster becomes stable (reviewed in Meldrum & Sear, 2008; Gebauer et al., 2009). This stable cluster will then act as a template for subsequent crystal growth. Recent research has identified long-lived pre-critical clusters which grow by collision and coalescence until critical size is reached and metastable ACC is formed (Colfen & Mann, 2003; Gebauer et al., 2008, 2009; Pouget et al., 2009; Song & Colfen, 2011). This has led to the emergence of the mesocrystal theory of crystal nucleation (e.g. Colfen & Antonietti, 2005; Song & Colfen, 2010), which demosntrates that pre-crystalline structures can aggregate by embedment in non-crystalline matrices. As these clusters have been observed to persist after the onset of nucleation, it has been proposed that it is

an increase in surface energy, rather than a simple increase in cluster size, which stabilises ACC and facilitates crystal nucleation (Pouget et al., 2009). This increase could result from impurities in solution (which are considered ubiquitous *in vivo* and all but impossible to fully eliminate *in vitro*) acting to stabilise $CaCO_3$ clusters, before their coalescence leads to formation of metastable ACC (Gebauer et al., 2008, 2009; Meldrum & Sear, 2008; Pouget et al., 2009). In this model, both free ions and pre-critical clusters play key roles in crystal nucleation, as do additives in solution (see below) (Gebauer et al., 2008, 2009).

1.3.3.2: Intra-crystalline proteins

The presence of organic molecules trapped within biominerals has been the subject of a large volume of research (e.g. Crenshaw, 1972; Towe & Thompson, 1972; Towe, 1980; Berman et al., 1988; Penkman et al., 2008, 2011). Until the conflict between different proposed pathways of mineral deposition is resolved, the precise spatial relationship between these 'intra-crystalline' proteins and the mineral phase cannot be fully understood. Recent transmission electron micrographs (TEM) of molluscan and algal biominerals reveal spherulitic Fresnel contrasts of a few nanometers, which probably correspond to these intra-crystalline proteins (Okumura et al., 2010, 2012). There is also evidence that it is possible for amino acids to be trapped within mineral crystals without disrupting lattice structure (Li et al., 2009). These proteins are not subject to the range of diagenetic influences which may be experienced by those exposed to the external environment (e.g. soil pH, hydrological conditions, microbial attack, infiltration by non-indigenous biomolecules). The significance of these intra-

crystalline proteins to this study is described in detail in a following chapter (Chapter 2).

1.3.4: Mineral deposition in eggshell

1.3.4.1: Overview

Avian eggshell deposition is one of the most rapid biomineralisation processes known; the complete shell (6 g of mineral in chicken eggshell) is deposited in the oviduct in less than 24 hours (e.g. Fu et al., 2005). $CaCO_3$ is precipitated from an acellular medium (uterine fluid) which contains the organic precursors of the shell matrix, and is saturated in calcium and bicarbonate ions (Nys et al., 2004). X-ray diffraction studies have demonstrated variable protein concentration within the calcified layers, with the highest concentrations observed in the mammillary layer and outer palisade (Lammie et al., 2006). This suggests major roles for the organic fraction in the initiation and termination of crystallisation, and also highlights the degree of temporally-controlled protein expression required within the oviduct for successful deposition of eggshell (Arias et al., 1993; Lammie et al., 2006).

During formation of a chicken egg the ionic and organic components of the uterine fluid change progressively during the three distinct phases of calcification, and are markedly different during initiation (5hr), growth (12hr) and termination (1.5hr) (Gautron et al., 1997; Hincke et al., 2010). This suggests major roles for the organic fraction of eggshell in regulating deposition (Gautron et al., 1997; Hincke et al., 2010). Initial crystal growth in the mammillary layer is

multidirectional, but only crystals growing perpendicular to the cuticle have space to continue growing in the palisade layer (Hincke et al., 2010). The organic component of eggshell limits growth space by inhibiting calcitic growth at faces parallel to the *c*-axis and thus promotes elongation of the calcite crystals which form the palisade layer (Hincke et al., 2010).

1.3.4.2: Calcite deposition: roles of the organic fraction

Calcite, the most stable crystalline form of $CaCO_3$, is formed when crystal growth is multidirectional (i.e. when no potential growth direction is thermodynamically favoured) (e.g. Silyn-Roberts & Sharp, 1986). Aragonite, which is formed when growth in a particular plane is thermodynamically favoured, is less stable than calcite but has mechanical properties which are often more useful (e.g. reduced brittleness); due to the thermodynamic favouring of calcite, aragonite requires some form of molecular input to favour its growth (e.g. Weiner & Addadi, 1997). For example, the presence of magnesium ions can favour aragonite precipitation, as can the activity of proteins which bind to specific crystal faces and thus favour unidirectional growth (Weiner & Addadi, 1997). The presence of calcite in avian eggshell could be viewed as a result of a lack of molecular activity favouring aragonitic growth. However, it is wrong to imply that the organic fraction of eggshell is not involved in directing calcitic growth. While the polycrystalline mammillary layer may be relatively free from molecular direction of crystal form during deposition, as spatial constraints on growing crystals seem to be the only factor favouring growth in particular planes, proteins are still heavily involved in promoting crystallisation in the mammillary layer (Freeman et al., 2010, 2011). The palisade layer exhibits greater molecular control of

crystal growth as it requires growth of large, anisotropic calcite crystals; it is likely that the organic fraction acts (in co-operation with spatial constraints) to favour crystal growth in a specific plane (Hincke et al., 2010). A range of proteins found in other mineralising systems can modulate calcite crystal nucleation and growth by binding to specific growth planes on calcite crystals (e.g. Albeck et al., 1993; Orme et al., 2001; Wang et al., 2009). The precise mechanisms of this in eggshell are unclear, but may relate to inhibition of growth at specific faces during nucleation and initial growth; spatial constraints (i.e. crystals only having one direction in which they can grow) probably become more important during later stages of calcification. This shows that the porosity and brittleness of eggshell, which are crucial to the material fulfilling its required functions (Section 1.2), are emergent properties of the pathways of crystal growth and are ultimately under molecular control.

1.3.4.3: Differences between eggshell and other mineralising systems

Although the general processes of ACC stabilisation and crystal nucleation may be conserved across diverse taxa (e.g. Addadi et al., 2003), there are important mechanistic differences between eggshell and other mineralising systems. This is not surprising; even within a single phylum (the Mollusca) a range of mechanisms by which calcite can be produced and shaped are known (Marie et al., 2009, 2011).

One major difference between eggshell and marine invertebrate calcification relates to the involvement of highly acidic proteins (pI < 4) in the latter (Belcher

et al., 1996; Falini et al., 1996; Weiner & Addadi, 1997; Feng et al., 2000; Aizenberg et al., 2002; Gotliv et al., 2003; Marin & Luquet, 2005; Marin et al., 2008). Although these affect the crystallising phase, the effects of particularly acidic amino acids on crystallisation cannot be reduced simply to pH effects; aspartic acid-rich proteins extracted from the calcitic phase of the spicules of marine poriferans and ascidians promote calcite formation, while glutamic acid-rich molecules are found in association with the ACC phase in the same organisms, and may even prevent crystallisation (Aizenberg et al., 1996). The most acidic protein yet discovered in mollusc shell has a theoretical pI of 1.67 and comprises 65% aspartic acid (Marin et al., 2008); the most acidic known eggshell protein has a theoretical pI of 4.53 and comprises 12% aspartic acid (Section 1.4 discusses eggshell proteins in detail). Furthermore, ovocleidin-17, which catalyses the early stages of calcification in chicken eggshell, is highly basic (theoretical pI = 9.78) (Reyes-Grajeda et al., 2004). Highly acidic proteins can affect calcite crystal morphology, act as nucleation templates for crystal growth when adsorbed onto a solid substrate, or inhibit crystal growth when free (Marin & Luquet, 2005). The failure to observe any such protein in avian eggshell deposition shows that the mechanisms in operation are probably different, although it remains possible that highly acidic proteins may be found. There are commonalities to be found; glycine and alanine-rich proteins are often involved in calcium binding in both eggshell and invertebrate mineral deposition (Hincke et al., 1999; Weiner et al., 2005).

1.3.5: Conclusions on biomineralisation and eggshell deposition

This section has summarised the state of knowledge regarding eggshell deposition at a molecular level. Although this remains incompletely understood, it is clear that organic molecules are key to this process, and also that there are important differences between deposition of eggshell and other biominerals. The following section will introduce the major proteins known to be present in eggshell, and discuss their roles in the deposition and regulation of the material; these are the molecules which will provide the basis of the eggshell identification technique developed in a later chapter (Chapter 2).

Section 4: Eggshell Proteins: Structure, Function & Diversity

1.4.1: Overview

This section describes some of the proteins known to perform roles in eggshell. Many of these are localised to specific regions, suggesting that they are involved in different stages of eggshell formation (Fernandez et al., 2003; Nys et al., 2004; Lammie et al., 2006). Unless stated otherwise, this summary will refer to the domestic chicken proteome, as this has been by far the most extensively researched. Data are generally lacking for most avian taxa, although some proteins are known from members of Struthioniformes (ostriches and allies), Anseriformes (geese and allies), and other members of Galliformes (turkeys). Major advances in understanding of other avian genomes and proteomes are expected in the near future (T. Gilbert, personal communication). Eggshell proteins have been subdivided into three general categories; those expressed in a range of bodily tissues; those present in the egg white as well as in the shell; and those unique to the shell (Nys et al., 2004).

1.4.2: Diversity of chicken eggshell proteins

Following the development of high-throughput proteomic techniques, up to 520 proteins have been found in chicken eggshell, 32 of which can be considered 'highly abundant' (Mann et al., 2006). The vast majority of these remain to be properly characterised, and their functions in the processes of biomineralisation and/or in the mature eggshell (if any) are as yet unknown. It is highly unlikely that the majority of these are actually integral to eggshell deposition and maintenance. Many proteins are apparently assimilated into eggshell in a non-

specific manner due to the rapid nature of mineralisation; proteins which are expressed in other parts of the oviduct may have diffused into the eggshell gland and thus been incorporated into the eggshell matrix, while others may represent breakdown products of other molecules present in the uterine fluid (Mann et al., 2006).

1.4.3: Eggshell proteins

1.4.3.1: Proteins expressed in diverse bodily tissues

Osteopontin is a moderately acidic protein (theoretical pI = 4.53) expressed (in different forms) in bone and eggshell, which functions as a phosphorylation-dependent inhibitor of $CaCO_3$ precipitation (Pines et al., 1995; Nys et al., 2004). It is found within the core of membrane fibres, the bases of mammillae, is highly concentrated in the palisade layer, and may modulate the speed of precipitation from uterine fluid and/or orient crystal growth (Nys et al., 2004; Hincke et al., 2010). Osteopontin is seemingly ubiquitous in tetrapod biomineralisation (Chien et al., 2009).

Clusterin is a moderately acidic (theoretical pI = 5.47) disulphide-bound heterodimeric glycoprotein expressed in a wide range of bodily tissues (Nys et al., 2004). Its structure and function is believed to be highly conserved across mammals and birds (Nys et al., 2004). Within eggshell, it is found throughout the mammillary and palisade layers, and probably acts as a chaperone to prevent premature aggregation and precipitation of other matrix components during

their incorporation into the rigid protein matrix (Mann et al., 2003; Nys et al., 2004).

Serum albumin has also been observed in association with clusterin in chicken and emu eggshell (Mann et al., 2003; Mann, 2004). A high proportion of **lipid-binding proteins** have been observed in the acid-soluble eggshell protein fraction, despite the very low relative abundance of lipids in eggshell; these probably exert antimicrobial effects (Lammie et al., 2006; Mann et al., 2006).

1.4.3.2: Egg-white and eggshell proteins

Ovalbumin is a moderately acidic (theoretical pI = 5.19) monomeric phosphoglycoprotein which is localised within the mammillary layer, and has been shown to be ubiquitous in a range of domestic species, as well as in reptile eggs (Mine, 1995; Panheleux et al., 1999; Nys et al., 2004; Miksik et al., 2007; Prajanban et al., 2012). Although ovalbumin was one of the first proteins to be isolated in pure form, its precise function remains unclear (Huntington & Stein, 2001). Despite sequence and structural homology with the serpin family of proteins, no inhibition of serine proteases has been observed (Huntington & Stein, 2001). The precise role of ovalbumin in egg-white and eggshell remains unclear, although its predominance (60-65% of total protein content) in egg-white may imply a major role as a storage protein (Huntington & Stein, 2001). It has also been shown to exhibit antimicrobial capacities (Pellegrini et al., 2004).

Ovotransferrin is a very slightly acidic (theoretical pI = 6.85) glycoprotein common to egg-white and eggshell. The transferrin family of proteins are heavily involved in iron binding, and are common as acute-phase reactants in plasma (Keung & Azari, 1982; Mine, 1995). This protein is therefore likely to be involved in iron transport in the developing embryo, and also in prevention of bacterial growth by iron sequestration (Keung & Azari, 1982; Mine, 1995).

Lysozyme is a highly basic (theoretical pI = 9.36) glycoside hydrolase found in a wide variety of species; it catalyses attacks on peptidoglycans in the cell wall of gram-negative bacteria, and therefore has strong anti-microbial capacities (Blake et al., 1965; Phillips, 1967; Mine, 1995; Vocadlo et al., 2001). Lysozyme occurs predominantly in the inter-crystalline protein fraction (Gautron et al., 2001), and is also a major component of the shell membrane (Mann, 2008). Although these proteins are apparently primarily involved in storage and transport, as well as in anti-microbial functions, they have been shown to affect calcite crystal morphology *in vitro* (Hincke et al., 2000, 2010; Gautron et al., 2001; Nys et al., 2004). Their roles in eggshell deposition may therefore be more complex than the current state of understanding suggests.

1.4.3.3: Eggshell-specific proteins

Proteins apparently unique to eggshell in *G. gallus domesticus* include the ovocalyxin family, which is thought to perform a variety of functions related to termination of mineralisation and shell formation (Gautron et al., 2001; Nys et al., 2004). **Ovocalyxin-32** (theoretical pI = 8.99) is localised to the outer palisade

layer, the vertical crystal layer and the cuticle, and is believed to play a key role in the regulation and termination of calcification (Gautron et al., 2001; Nys et al., 2004). This protein occurs in both the inter- and intra-crystalline protein fractions, and may also perform antimicrobial functions (Gautron et al., 2001; Xing et al., 2007). **Ovocalyxin-36** (theoretical pI = 5.61) is localised to the inner calcified layers and the shell membrane, and is highly up-regulated when the egg is in the uterus. It is therefore likely to be involved in the initial stages of shell formation (Nys et al., 2004; Gautron et al., 2007). High sequence homology with antimicrobial mammalian proteins suggests that ovocalyxin-36 may also exercise antimicrobial functions (Bingle & Craven, 2004; Tian et al., 2010).

Ovocleidin-116 (OC-116) is a mildly acidic (pI = 6.42) secretory calcium-binding phosphoprotein (SCPP); this family of proteins are ubiquitous in tetrapod mineralisation (Kawasaki & Weiss, 2003). This protein is at the core of a dermatan sulphate proteoglycan (see section 1.4.3.5) which is widely distributed throughout the palisade layer; it is also the most abundant (80µg/g shell) matrix protein (Hincke et al., 1999; Mann et al., 2002; Nys et al., 2004). The high proportion of glycine and alanine (24% of the total sequence) is consistent with a role as a framework protein (Weiner & Addadi, 1997; Hincke et al., 1999), while SCPP proteins bind calcium via acidic amino acid residues (e.g. Rose & Hincke, 2009). OC-116 is a member of the SIBLING (small integrin-binding ligand, N-linked glycoprotein) gene locus, where it is adjacent to osteopontin (Rose & Hincke, 2009). OC-116 is highly expressed during the formation of the shell matrix external to the mammillae, and has a strong effect on crystal orientation; it may modulate crystal growth during palisade formation by preventing crystal

growth along certain axes (Fernandez et al., 1997; Hincke et al., 1999; Nys et al., 2004). Also, the sulphated form could favour crystal elongation through electrostatic interactions (Rose & Hincke, 2009). Single nucleotide polymorphisms (SNPs) in this protein have been linked to changes in shell thickness, elasticity, and shape (Dunn et al., 2008); this protein is very important in both the developing and mature eggshell, and functional equivalents in other avian species should be targeted as a matter of priority. A role in bone mineralisation has also been proposed (Fernandez et al., 1997; Horvat-Gordon et al., 2008).

Protein	Mass (kDa)	pI	Length (AAs)	Type
Clusterin	51.3	5.47	448	Glycoprotein
Lysozyme	16.2	9.36	147	Glycoside Hydrolase
OC-116	76.9	6.42	743	Phosphoprotein
OC-17	15.3	9.78	142	C-type lectin
OCX-32	30.6	8.99	275	Unclear
Osteopontin	29.2	4.53	264	Phosphoprotein
Ovalbumin	42.9	5.19	386	Phosphoglycoprotein
Ovotransferrin	77.8	6.85	705	Glycoprotein

Table 1.1: Basic biochemical parameters of some known chicken eggshell proteins

1.4.3.4.1: C-type lectin proteins

C-type lectin-like (CTL) proteins (carbohydrate-binding proteins which require calcium to bind to their substrate and perform a wide variety of roles in nature) have been identified in a range of species, and have long been thought to play a key role in $CaCO_3$ precipitation and crystallisation. It is likely that these proteins exert biphasic effects on $CaCO_3$ precipitation, acting as promoters at low

concentrations and as inhibitors at high concentrations (Nys et al., 2004). **Ovocleidin-17 (OC-17**) is an abundant (40µg/g shell) phosphoprotein found in both glycosylated (OC-23) and non-glycosylated forms in the mammillary and palisade layers of galliform eggshell (Mann, 1999; Panheleux et al., 1999; Mann et al., 2002; Reyes-Grajeda et al., 2004; Portugal et al., 2010). Like OC-116, OC-17 has a very high component of alanine and glycine (28.9%); this is consistent with a role as a framework protein. OC-17, which is highly basic (theoretical pI = 9.78), is abundant in the intra-crystalline protein fraction (Gautron et al., 2001). **Ansocalcin** has been observed in the eggshell of anseriform species (geese, ducks and swans) (Lakshminarayanan et al., 2002), and exhibits 40% sequence identity to OC-17 (Nys et al., 2004). Both of these proteins have also been observed to bind bacterial polysaccharides and thus fulfil antimicrobial functions (Wellman-Labadie et al., 2008). In ostrich eggshell (OES), **struthiocalcin (SCA) 1 & 2** have been isolated (Mann, 2004). Emu (**dromaiocalcin (DCA) 1 & 2**) and rhea (**rheacalcin (RHA) 1 & 2**) eggshell also contain two different CTL-like proteins (Mann & Siedler, 2006).

1.4.3.4.2: Sequence homology and inter-relationships of CTL proteins

Within ratites (ostrich, emu and rhea), these proteins are easily grouped into two categories; Group 1 (SCA-1, DCA-1 & RCA-1) and Group 2 (SCA-2, DCA-2 & RCA-2) (Mann & Siedler, 2006). Group 1 proteins exhibit 70-77% sequence identity within ratites, no phosphorylation, and a variable number (7-9) of cysteines (Mann & Siedler, 2006). Ansocalcin (goose) exhibits 63-70% sequence identity with these proteins, no phosphorylation, and seven cysteines (Mann & Siedler, 2006). Group 2 proteins exhibit 78-85% sequence identity, 2-3

phosphorylated serines at almost identical locations, and a common set of 6 cysteines which forms the entirety of their cysteine content (Mann & Siedler, 2006). OC-17 is placed within this group on the strength of its two phosphorylated serines and six cysteines, despite only 42-47% sequence identity (Mann & Siedler, 1999, 2006).

The question of why ratites seem to require two CTLs while other taxa do not remains unanswered, but may be related to the extraordinary thickness and mechanical strength of ratite eggshell. It may also relate to gene loss in the Neognathae following divergence from the Palaeognathae (Figure 1.3) (Mann & Siedler, 2006). If this is the case, it is interesting that different genes have apparently been lost in the Galliformes and Anseriformes without affecting the ability of either to successfully mineralise eggshell; this would suggest that avian eggshell deposition might not be as dependent on tight molecular control as some other mineralising systems.

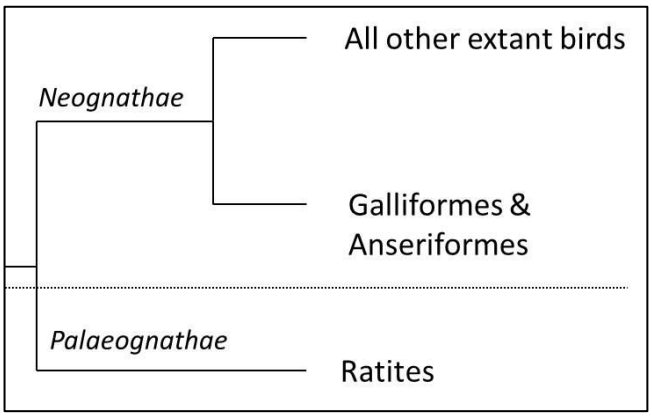

Figure 1.3: Simplified phylogenetic tree of Aves, following Hackett et al (2008)

1.4.3.4.3: Roles of CTL proteins in eggshell deposition

CTL proteins are believed to fulfil crucial roles as framework proteins in the precipitation of $CaCO_3$ during the initial stages of calcification in a range of phyla. Similar roles for these proteins have been proposed in molluscs, echinoderms and mammalian disease conditions (Wilt, 1999; Mann et al., 2000; De Reggi & Gharib 2001). The interaction between OC-17 and calcareous nanoparticles is size-dependent (Freeman et al., 2010). Computer models suggest that this protein is bound by arginine residues to smaller nanoparticles and promotes aggregation of these, leading to mineralisation (Freeman et al., 2010, 2011). Subsequent changes in nanoparticle structure and/or size dislodge the protein, which can then adsorb to smaller nanoparticles (Freeman et al., 2010, 2011). According to this model, OC-17 plays a catalytic role in $CaCO_3$ precipitation (Freeman et al., 2010, 2011). This process will continue until such time as the protein is surrounded by calcite and is occluded within the mineral (Freeman et al., 2010, 2011).

The role played by ansocalcin seems to be different (Lakshminarayanan et al., 2002, 2005; Reyes-Grajeda et al., 2004; Freeman et al., 2010). Although this is not necessarily surprising as these proteins align with different groups (see above), it begs the question of how the requirements of calcification are met in the two avian orders. The intra-crystalline component of ansocalcin has been observed to be high relative to that of OC-17 (Lakshminarayanan et al., 2005). This may be due to slightly different mechanisms of protein-driven mineralisation. Ansocalcin aggregates as crystal nucleation and growth progress (Lakshminarayanan et al., 2005), and may therefore act as a template for

mineralisation rather than as a catalyst. This would lead to a high degree of entrapment within the matrix relative to OC-17 if ansocalcin remains bound to its substrate while OC-17 dissociates after a critical size or specific structural conformation is reached. The functional role(s) of ratite CTL proteins in the developing and mature eggshell are not yet clear.

1.4.3.5: Effects of sugars on protein form and function

Many proteins are post-translationally modified by the addition of sugars. Pilot work on the role of proteoglycans and glycosaminoglycans in eggshell formation found that keratan and dermatan sulphate proteoglycans were differentially localised within the eggshell, hinting at different roles in the mineralisation process (Arias et al., 1991, 1992). Keratan sulphates are less anionic than dermatan sulphates, and are believed to be associated with the formation of the mammillae, in which initial crystal orientation is non-uniform; immunohistological studies have shown that deposition of the mammillae coincides with expression of keratan sulphate (Fernandez et al., 1997; Nys et al., 2004). *In vitro* studies have shown that both keratan sulphate and dermatan sulphate proteoglycans have high calcium affinity; it has been proposed that the main roles of keratan sulphates relate to maintenance of calcium reserve bodies, while dermatan sulphates affect the precipitation of $CaCO_3$ during shell formation (Fernandez et al., 1997). For example, OC-116 forms the core protein of a dermatan sulphate proteoglycan which functions as a framework protein during shell formation (Hincke et al., 1999).

Another example is provided by OC-17, which can occur in its phosphorylated form or as a glycosylated form (OC-23) (Mann, 1999). In this protein, phosphorylation and glycosylation appear to be mutually exclusive, and it is possible that the two forms of the protein play different roles in eggshell formation (Mann, 1999). While the mechanisms by which OC-17 affects mineralisation have been described (Freeman et al,. 2010, 2011), those of OC-23 remain unclear. Although the roles of sugars in the function of eggshell proteins remains largely unknown, it is clear that they have major effects on the roles played by these molecules.

1.4.4: Effects of eggshell proteins on calcite crystal morphology

It has been suggested that ansocalcin, but not OC-17, can affect calcite morphology and favour crystal aggregation at high concentrations (Lakshminarayanan et al., 2005). Purified OC-17 has been observed to have minimal effect on calcite crystal morphology at low concentrations (Nys et al., 2004), although at higher concentrations an effect has been observed (Reyes-Grajeda et al., 2004). It has been suggested that this difference in interaction between crystal morphology and protein may result from direct interaction between OC-17 (and SCA-1) and the carbonate anion, which has not been observed in the other proteins listed (Marin-Garcia et al., 2008; Freeman et al., 2011). This similarity in function is particularly interesting given that these proteins are not considered members of the same 'group' of CTL proteins (Section 1.4.3.4.2).

Osteopontin may play a role in determining crystal morphology and orientation in galliformes by inhibiting calcitic growth at specific crystallographic faces (Lavelin et al., 2000; Chien et al., 2008). The expression of this protein is induced by mechanical strain within the oviduct, and is limited to regions highly involved in the calcification process (Pines et al., 1995; Lavelin et al., 2000). Osteopontin contains phosphothreonine and phosphoserine residues (Lavelin et al., 2000) which make it a viable candidate for initial stabilisation of ACC (Bentov et al., 2010). It may also be involved in compartmentalisation of CRBs in combination with keratan sulphate (Fernandez et al., 2003). Ovotransferrin can cause reduction in crystal size, and favours elongation of calcite crystals at high concentrations, as do pure glycosaminoglycans (Nys et al., 2004). Lysozyme also affects crystal morphology by inhibiting growth at crystallographic faces parallel to the *c*-axis (Nys et al., 2004). All of these interactions may be governed by the hydrophobic and electrostatic interactions between biomolecules and calcitic surfaces (Nys et al., 2004).

1.4.5: Conclusions

1.4.5.1: Conclusions on section 4

This section has introduced some of the proteins present in eggshell, and described their roles in the developing and mature eggshell. Almost all of the proteins described have been shown to exert some effect on crystal morphology *in vitro*, and almost all can also perform antimicrobial functions. This diversity of effects on crystal growth goes some way towards explaining why characterising the precise interaction between the organic and inorganic phases of eggshell *in*

vivo is so challenging; the emergent products have been shaped by the synergistic activity of a wide range of molecules.

1.4.5.2: Conclusions on chapter 1

This chapter has explained the rationale underlying this research project. It has also provided a brief account of what is known about past egg use and of the archaeological challenge eggshell poses; of the physical structure of eggshell, and of the processes by which it is deposited; and of some of the roles which specific proteins play in deposition and maintenance of the eggshell. The following chapter describes the development of an identification system for archaeological eggshell.

Chapter 2: Early experimental work

Section 1: Isolating an intra-crystalline protein fraction in eggshell

2.1.1: Overview

2.1.1.1: Introduction

Before any of the major themes of this project can be developed, it must be demonstrated that a suitable target pool of proteins exists in avian eggshell, and that it can be isolated for analysis. This section details early attempts to isolate an intra-crystalline protein fraction in avian eggshell. Two approaches are described here: exposure of powdered eggshell to bleach in order to monitor retention of intra-crystalline proteins (the bleaching test), and monitoring of diffusive loss of amino acids at high temperature (the leaching test).

2.1.1.2: Closed biomineral systems: 'intra-crystalline' proteins

Biomineral proteins can occupy two positions relative to the mineral phase (e.g. Crenshaw, 1972; Towe & Thompson, 1972; Towe, 1980; Berman et al., 1988). These are referred to as inter-crystalline (i.e. occupying the spaces between mineral crystals) and intra-crystalline (i.e. occluded within the mineral phase) (e.g. Penkman et al., 2008). Intra-crystalline proteins are likely to result from occlusion of proteins during mesocrystalline assembly of mineral nanocrystals (e.g. Colfen & Antonietti, 2005; Song & Colfen, 2010), and have been isolated in a range of biominerals including the opercula of terrestrial gastropods, ratite and megapode eggshells, and certain brachiopods, foraminifera, and molluscs (Brooks et al., 1990; Stathoplos & Hare, 1993; Sykes et al., 1995; Walton, 1998;

Miller et al., 2000; Clarke et al., 2007; Penkman et al., 2008, 2011; Demarchi, 2009; Demarchi et al., 2011, 2013). There are major differences in the diagenetic pathways experienced by these different protein fractions, and the intra-crystalline fraction of some biominerals has been shown to function as an effectively closed system across the time-scales relevant to most archaeological and geochronological research (e.g. Towe, 1980; Nakahara et al., 1981; Brooks et al., 1990; Sykes et al., 1995; Miller et al., 2000; Penkman et al., 2008, 2011; Demarchi et al., 2011, 2013).

Although it has long been established that organic material can become occluded within mineral crystals (Nickl & Henisch, 1969; Henisch, 1970), the precise spatial relationship between 'intra-crystalline' proteins and the mineral phase remains incompletely understood due to an unresolved conflict between two proposed pathways of crystal nucleation (Section 1.3.3.1). The term 'intra-crystalline' has tended to be used in an operational context to refer to biomolecules isolated by oxidation, which have been shown to behave as if in an effectively closed system (e.g. Brooks et al., 1990; Penkman et al., 2008). As detailed previously (Section 1.3.3.2), proteins fully occluded within the mineral phase have recently been observed, and it is possible for amino acids to be trapped within mineral crystals without disrupting lattice structure (Li et al., 2009; Okumura et al., 2010, 2012). If these proteins are protected by the mineral phase, they are not subject to the range of diagenetic influences which may be experienced by those exposed to the external environment (e.g. soil pH effects, hydrological conditions, microbial attack, infiltration by non-indigenous biomolecules, contamination).

2.1.1.3: The importance of analysing a closed system

As the pathways of molecular diagenesis (detailed in chapter 5) differ between inter- and intra-crystalline proteins and amino acids, analysing a composite of both can produce misleading results and/or reaction kinetics which are impossible to model accurately (Collins & Riley, 2000). In order to circumvent this problem, removal of the inter-crystalline pool by exposure to a strong oxidating agent (bleach) can be used to isolate the intra-crystalline pool (e.g. Sykes et al., 1995; Penkman et al., 2008; Penkman, 2010; Demarchi et al., 2011, 2013; Crisp et al., 2013). In most biominerals this represents a small fraction of the total protein content. During natural diagenesis, the inter-crystalline fraction is exposed to environmental influences such as pH, changes in water availability, contamination, and microbial activity. There is also a high likelihood of diffusive loss (leaching) of components of the inter-crystalline fraction to the environment; this can have a significant effect on observed reaction rates, and has led to misleading reconstructions of diagenesis (e.g. Kimber et al., 1986; Kimber & Griffin, 1987; Collins & Riley, 2000). These influences can make reconstruction of the diagenetic pathways experienced by preserved biomolecules impossible (Collins & Riley, 2000); this necessitates isolation of a closed system of amino acids which can neither interact with the external environment nor escape from the mineral.

In a closed system, amino acids should only be lost by decomposition; free amino acids produced by hydrolysis will be retained within crystalline structures even after prolonged bleaching and heating (i.e. artificially induced diagenesis).

As they are partitioned away from direct environmental influences, diagenesis of intra-crystalline proteins should be affected only by time, temperature and the local molecular environment; non-indigenous molecules should be excluded (Sykes et al., 1995; Collins & Riley, 2000; Penkman et al., 2008). As all components should be retained within a closed system, observed reaction kinetics for diagenetic reactions should be more straightforward to model (e.g. Collins & Riley, 2000). For these reasons, intra-crystalline proteins provide the best available target for identification of taxonomically diagnostic peptides in eggshell (Section 2.2), and for quantification of diagenetic reactions in eggshell (Chapter 5).

2.1.2: Methods

2.1.2.1: Overview

The first step towards identifying potential closed system behaviour is to test for a residual amino acid fraction after exposure to bleach. This can demonstrate the existence of an amino acid pool resistant to oxidation, although isothermal leaching tests are required to demonstrate closed-system behaviour over a prolonged period of diagenesis. Use of diagenetic reactions to derive information about preserved biominerals requires that there has been effectively no interaction between preserved amino acids and the external environment. Only time, temperature and the local molecular environment should influence diagenesis. One way of testing the potential for interaction between amino acids in the intra-crystalline fraction and the environment is to monitor retention of amino acids under prolonged exposure to high temperature. In theory, if a sample proves resistant to leaching under short-

term high temperature conditions, it should also prove resistant over far longer periods at ambient temperatures, assuming (for example) that increased structural diagenesis over longer periods has no significant effect on leaching (e.g. Collins & Riley, 2000).

2.1.2.2: Sample preparation: bleaching test

Modern eggshell (chicken & duck) was obtained from commercial retailers. The shell membrane was carefully removed by hand, and the shell was washed using ultra-pure water and allowed to air dry. Each shell was then separately ground to a powder (≤500µm particle size) with a quartz pestle and mortar, and weighed into sterile 2ml Eppendorf tubes. Three laboratory replicates were prepared for each sample. The powder (typically 5-20mg per sample) was then exposed to strong NaOCl (w/v 12%, 50µl/mg) at room temperature (Penkman et al., 2008). Samples were agitated daily during bleaching to maximise exposure of the powder to the bleach. Upon expiry of a pre-determined length of time (0, 18, 24, 48, 72, 120 or 240 hours), the bleach was removed by repeated rinsing (x6) with ultra-pure water followed by brief suspension in HPLC-grade methanol. Samples were then air-dried before being separated into separate fractions for analysis of free (FAA) and total hydrolysable (THAA) amino acid composition. Unbleached controls were included in both analyses.

2.1.2.3: Sample preparation: leaching test

Chicken eggshell was powdered as described above and exposed to bleach for 48 hours. The bleach was then removed as described above. Samples were air

dried, then weighed out into sterile glass ampoules. 300µl of ultra-pure water was added before the ampoules were heat-sealed and placed in an oven at 140°C for varying time periods (0, 1, 2, 4, 6, 8, 24, or 48 hours). Samples were prepared in triplicate at each time point. When the ampoules were removed from the oven, 100µl of the water was removed and frozen for future reference. Another 100µl each was prepared for analysis of its FAA and THAA content. The powder was air-dried, before being weighed into sterile hydrolysis vials for demineralisation. Water blanks were included at each time point in order to derive background amino acid concentrations.

2.1.2.4: Demineralisation & rehydration

For analysis of the powder FAA content, subsamples (1-10mg) were weighed out into sterile 2ml glass hydrolysis vials. The samples were then completely demineralised using 2M hydrochloric acid (HCl). The smallest possible volume of acid was used in order to minimise hydrolysis of peptide bonds; the FAA analysis targeted only those residues which were free prior to acid demineralisation. The standard volume of HCl used was 20µl/mg sample, although in unbleached controls considerably more than this was required. Complete demineralisation required a higher standard volume than was used than in previous studies on mollusc shell, as eggshell seems more resistant to acid demineralisation (possibly due to the higher stability of calcite relative to aragonite) (Demarchi, 2009). 2M acid blanks were also taken at this stage. When demineralisation was complete, the samples were dried under vacuum in a centrifugal evaporator.

For analysis of the THAA content, samples were weighed into sterile vials (as above) and demineralised using 7M HCl at a standard volume of 20µl/mg sample. When demineralisation was complete the samples were flushed with nitrogen and heated at 110°C for 24 hours; this is the optimal temperature and time for hydrolysis of all peptide bonds combined with minimal artificially induced diagenesis (Hill, 1965; Kaufman & Manley, 1998). 7M acid blanks were also taken at this stage, and were treated in an identical manner to the samples. Upon removal from the oven, the samples were dried under vacuum in a centrifugal evaporator.

When dry, samples were rehydrated using a solution containing a known concentration of the non-protein amino acid L-Homoarginine (LHarg) (0.01M HCl and 1.5mM sodium azide, with 0.01 mM LHarg) in order to allow the subsequent quantification of individual enantiomeric concentrations. For the FAA analyses, a volume of 10µl/mg sample was sufficient in all cases. The volumes required for the THAA analyses were both higher and more variable. Unbleached controls typically required at least 160µl/mg, while the majority of bleached samples could be processed using 60µl/mg. After rehydration and thorough vortexing, an aliquot of each sample (13.5µl) was transferred to a sterile 2ml HPLC autosampler vial for analysis. LHarg blanks were also included at this stage. Acid blanks were rehydrated with an arbitrary volume of the LHarg solution. All water samples were rehydrated using 10µl/mg LHarg solution.

2.1.2.5: Analysis by reverse-phase HPLC

High performance liquid chromatography (HPLC) is commonly used for quantification of individual enantiomers; the enantiomeric separation is at least as good as in gas chromatography, and HPLC requires less sample and less exhaustive sample preparation (Kaufman & Manley, 1998). HPLC separates a mixture into individual components by dissolving it in a mobile phase, and then measuring the retention time of different components within a stationary phase (the 'column') (Figure 2.1). Retention time is affected by the chemical interactions between each component and the stationary phase. HPLC uses high pressure to force the analyte through a C_{18} stationary phase, before different enantiomers are identified in a fluorescence detector (see below) (Kaufman & Manley, 1998).

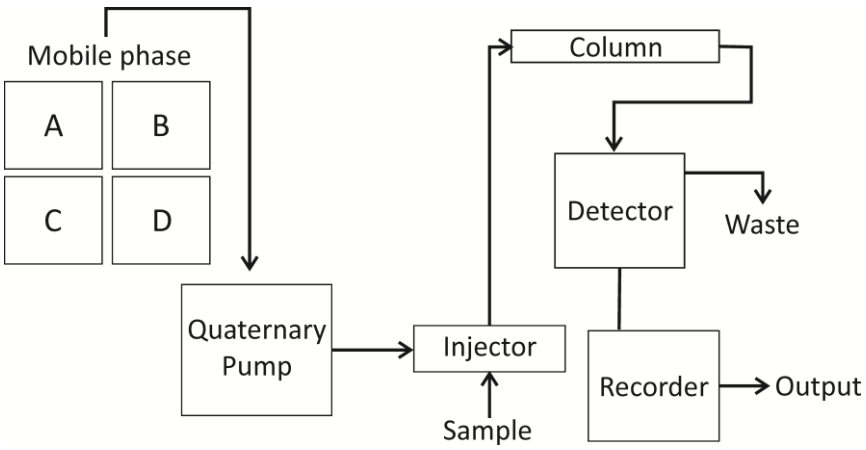

Figure 2.1: Schematic of the HPLC system used in this study. Mobile phase A = Na acetate buffer (see below); B = ultra-pure water; C = HPLC-grade methanol; D = acetonitrile.

This study utilises an automated reverse-phase HPLC (RP-HPLC) protocol, which separates amino acids by using a non-polar stationary phase and a polar mobile phase (Hare & Gil-Av, 1979; Kaufman & Manley, 1998; Penkman, 2005). Elution time for each amino acid is determined by mass, structure and hydrophobicity. This system allows enantiomeric separation in ten amino acids: aspartic acid/asparagine (Asx), glutamic acid/glutamine (Glx), serine (Ser), arginine (Arg), alanine (Ala), valine (Val), methionine (Met), phenylalanine (Phe), isoleucine (Ile) and leucine (Leu). Asx is used to denote combined aspartic acid and asparagine, while Glx denotes combined glutamic acid and glutamine; this is because asparagine and glutamine are rapidly and irreversibly deamidated during acid hydrolysis, and are subsequently indistinguishable from aspartic acid and glutamic acid (Hill, 1965). Other amino acids for which only one isomer is detected are L-threonine, L-histidine, glycine (which only has one isomer), and L-tyrosine. The major advantages of this system are the small volume of material required (<1mg) and the efficiency of the automated analytical procedure. The ability to detect enantiomers of a range of amino acids allows compromised specimens to be identified, as D/L ratios in different amino acids should covary strongly (Goodfriend, 1991; Kaufman & Manley, 1998).

The analytical protocol used is that developed by Penkman (2005). 2µl of sample was mixed online with 2.2µl of derivatising agent (OPA-IBLC - 260mM IBLC, 170mM OPA in 1M potassium borate buffer, adjusted to pH 10.4 with KOH) (e.g. Kaufman & Manley 1998; Fitznar et al., 1999). Derivatives were then separated on a Hypersil C_{18} BDS column (sphere diameter 5µm; 5x250mm column) at 25°C (elution time was around 115 minutes). A three-solvent linear gradient was used

for separation: sodium acetate buffer (23mM sodium acetate tri-hydrate, 1.3µM Na_2EDTA, 1.5mM sodium azide: adjusted to pH = 6 ± 0.01 using 10% acetic acid and 10M NaOH), HPLC-grade methanol, and acetonitrile. Enantiomeric detection took place in a fluorescence detector (Ex = 230nm, Em = 445nm). The elution time defines the enantiomer under consideration when compared to results obtained on standards of known composition (Figure 2.2). The peak area is directly proportional to the concentration of the enantiomer, and this is quantified by calibration against the internal standard (LHarg).

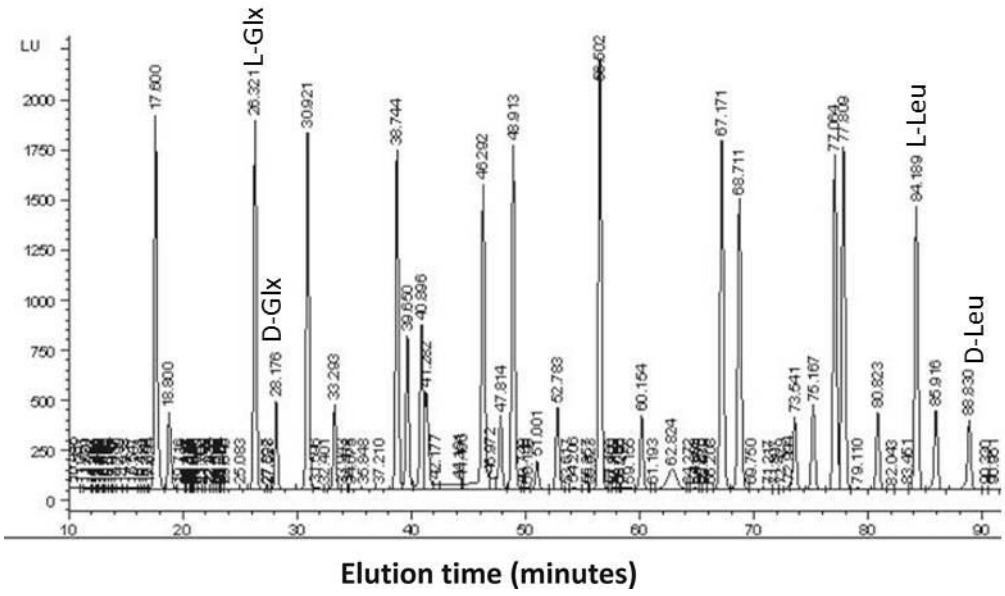

Figure 2.2: RP-HPLC output (standard solution) with examples of specific enantiomers given. Amino acids with negatively charged side chains (Asx, Glx) elute first, while those with large, hydrophobic side chains (e.g. Ile, Phe, Leu) elute last.

2.1.3: Results

2.1.3.1: THAA concentration

The THAA fraction within both chicken and duck eggshell (Figure 2.3) shows a dramatic decrease in total amino acid concentration after bleaching due to oxidation of inter-crystalline amino acids (Penkman et al., 2008). The total concentration appears to reach a plateau after 24 hours bleaching, after which no major decrease in concentration is observed (Figure 2.3). This suggests the retention of a fraction (around one third of the total protein content – Figure 2.3) which is resistant to oxidation, and is therefore probably intra-crystalline (Penkman et al., 2008; Demarchi, 2009). Inter-sample variability also decreases in both species after bleaching.

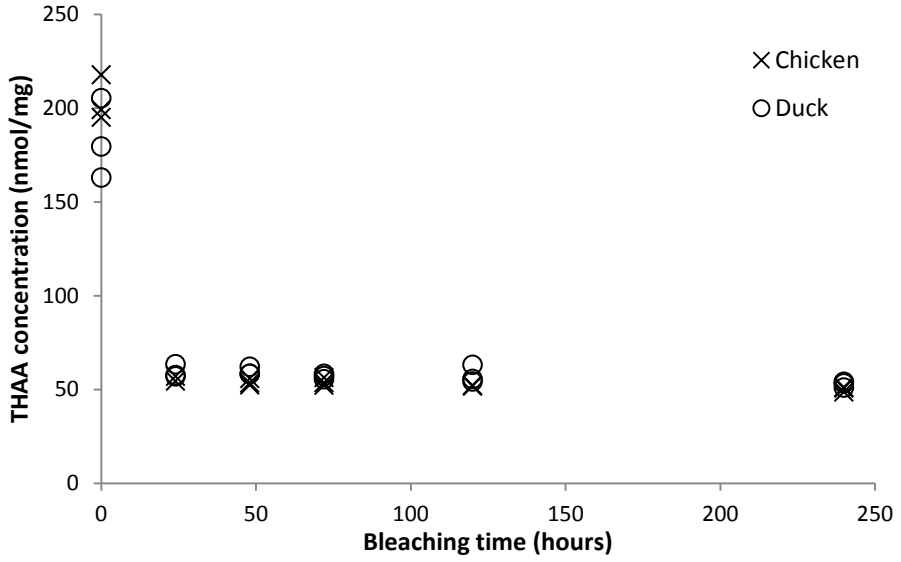

Figure 2.3: THAA concentration in powdered chicken and duck eggshell vs. bleaching time.

Despite the plateau observed in amino acid concentrations, these results do not confirm closed-system behaviour; minor reductions in total amino acid concentration are observed at most time points (Figure 2.4). However, the extent to which this results from natural heterogeneity between samples is not clear; the coefficient of variation (CV) of bleached chicken and duck shell often approximates the difference in concentration between time points (Figure 2.4). The minor reductions in amino acid concentration with bleaching time may suggest either that some residues may not be fully occluded within the mineral phase, as they are vulnerable to prolonged oxidation, or that the oxidating agent is causing structural damage to the mineral phase and thus exposing previously occluded amino acids. However, even at these time points the coefficient of variation closely approximates the concentration differences observed between time points (Figure 2.4). These results are promising but inconclusive; isothermal heating experiments are required to further test for closed system behaviour.

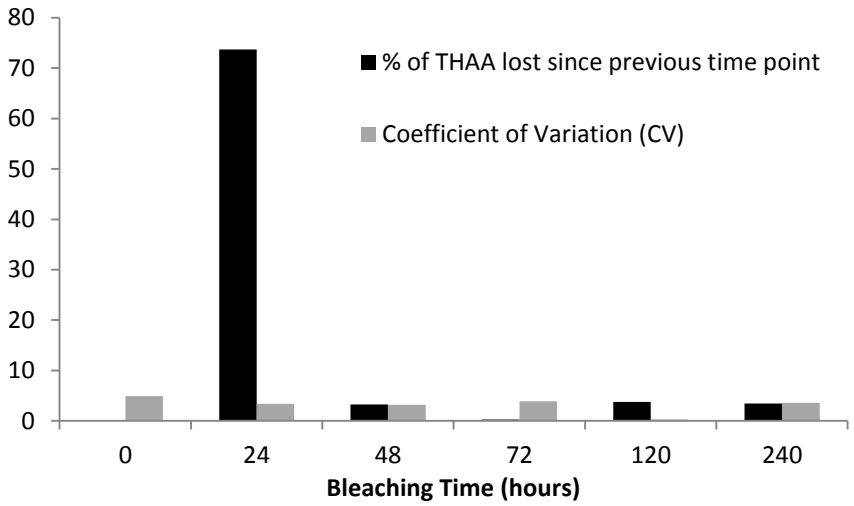

Figure 2.4: Amino acid loss with bleaching time in chicken eggshell. These results show minimal loss of amino acids with bleaching time in most cases.

2.1.3.2: FAA concentration

FAA concentration is extremely low in both species at all stages of the bleaching experiment (typically 0.3 – 0.8nmol/mg, maximum 1.27% of THAA). FAA concentrations were often close to the limit of detection of the techniques used; this introduces instrument error as a potentially major factor in interpretation of the FAA data. No clear patterns were expected, and none were observed (data not shown).

2.1.3.3: THAA composition

After exposure to bleach for 48 hours, around 26.4% of the total original amino acid pool is retained in chicken eggshell, and around 33.6% in duck eggshell. The proportion of individual amino acids retained is shown below (Figure 2.5). A higher proportion of every amino acid is retained in duck eggshell, suggesting that the intra-crystalline component is proportionally more significant than in chicken eggshell; this may relate to the different mechanisms by which duck and chicken eggshell proteins promote mineral deposition (Section 1.4.3.4.3).

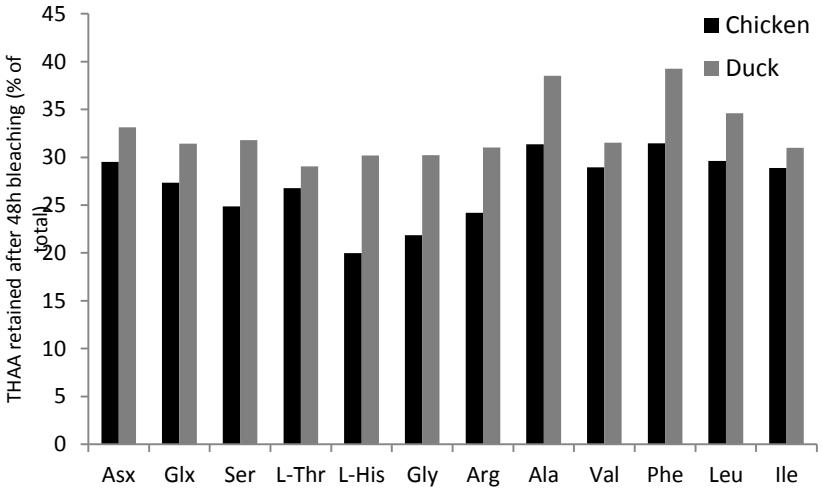

Figure 2.5: Intra-crystalline proportion of individual amino acids in chicken and duck eggshell. The general pattern of retention is similar in both species.

2.1.3.4: FAA composition

It is difficult to derive confident conclusions on the FAA data, as the data resolution is relatively poor; in chicken and duck eggshell, FAA concentrations are mostly in the 1-100 pmol/mg range (i.e. close to the limit of detection for the analytical technique used). If taken at face value, it appears that free glycine is proportionally over-represented relative to total glycine in both species (data not shown). This seems to suggest that glycine is preferentially released, possibly due to diketopiperazine formation (Steinberg & Bada, 1981, 1983); formation of particularly unstable peptide bonds; and/or as a breakdown product of other amino acids (Vallentyne, 1964; Bada et al., 1978; Bada & Man 1980; Walton 1998). There is a decrease in the proportion of FAA represented by serine after short-term bleaching which may be related to the increase in glycine representation.

2.1.3.5: THAA racemisation

A detailed description of the racemisation reaction is given in a subsequent chapter (Chapter 5). In the context of testing for closed-system behaviour, it was necessary to test whether the bleach affected racemisation in intra-crystalline amino acids. Bleach can dissociate to form hypochlorous acid, which can induce racemisation by facilitating proton abstraction (Neuberger, 1948). In a truly closed system, amino acids should be protected from this process, and a plateau

in D/L ratios should be observed when the intra-crystalline amino acids have been isolated. Asx, Glx and serine display an increase in D/L ratios with bleaching time in both chicken and duck shell (Figure 2.6). Only very slight increases are observed in slower-racemising residues.

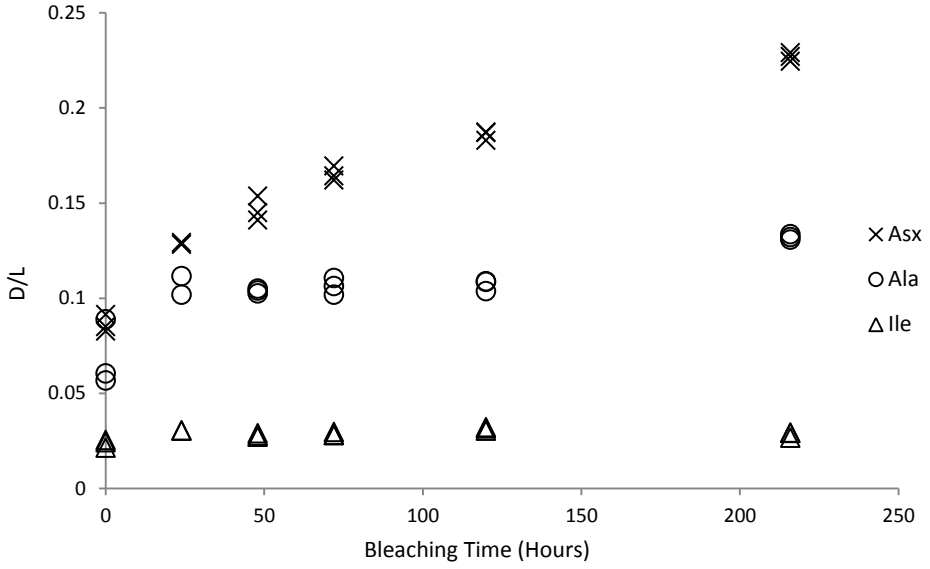

Figure 2.6: THAA D/L vs bleaching time in chicken eggshell. Data are shown for fast (Asx), medium (Ala) and slow (Ile) racemising amino acids. The same pattern is observed in duck eggshell.

The increase in D/L ratio with bleaching time in some amino acids confirms the need for more detailed tests for closed-system behaviour (next section). Although the hydrolysis protocol used (see above) does induce some racemisation (e.g. Kaufman & Manley, 1998), in a closed system this should be uniform for all samples. These results resemble those obtained from mollusc shell (e.g. Penkman et al., 2008; Demarchi, 2009; Demarchi et al., 2011); higher retention of racemised free amino acids in the intra-crystalline fraction can

explain the initial rise in D/L ratios, while the slower increase after bleaching suggests that at least some intra-crystalline proteins are not fully occluded within the mineral phase, and/or that the system becomes less closed with prolonged structural diagenesis. In a truly closed system, D/L ratios should rise initially with bleaching time, before either reaching a plateau or proceeding more slowly when the residual fraction has been isolated (Penkman et al., 2008). The potential for hydrolysis (and therefore for racemisation) in the closed system proteins should be relatively limited; the interior of the mineral phase in eggshell is largely anhydrous. However, water may be generated by decomposition of serine and threonine, or by condensation reactions which form melanoidins (Collins et al., 1992; Walton, 1998).

2.1.3.6: Diffusive loss of amino acids

Heating increases the rate of hydrolysis, leading to more rapid production of small peptides and free amino acids. In an open system, these can then be lost by leaching. Due to their smaller size and higher mobility, free amino acids are more likely to be lost than peptide-bound residues. In a closed system, the free amino acids should be retained, and the only cause of reduction in THAA concentration should be decomposition reactions. Under prolonged heating, there is a marked decrease in THAA in the powders, and a slight increase in the THAA present in the water (Figure 2.7). It is unclear from these results alone whether the decrease in powder THAA is due primarily to leaching or decomposition, but this confirms that at least a small amount of leaching is occurring (i.e. that the system is not completely closed).

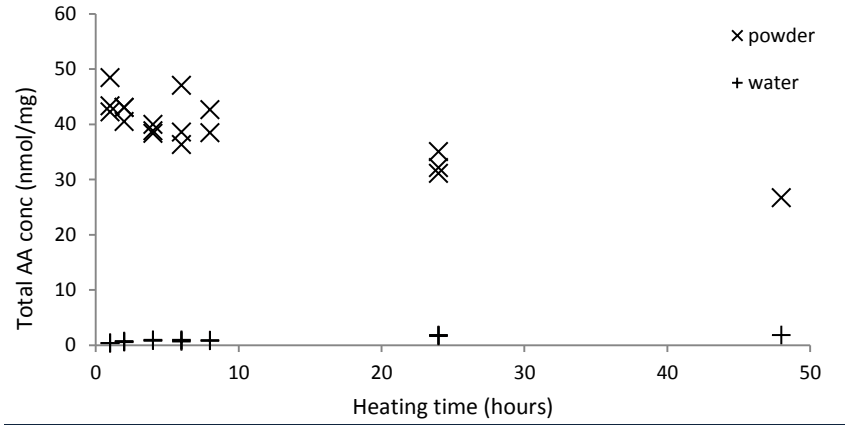

Figure 2.7: Changes in THAA concentration in powder and water with heating time (140°C)

In order to calculate the measurable extent of leaching, THAA concentrations present in the water were calculated as a percentage of the average THAA concentration of unheated, 48-hour bleached chicken eggshell. No routine background subtraction was employed, as levels of amino acids in the water blanks were generally low (15 – 115 pmol/mg in total: 1%-10% of THAA in the water; 0.03%–0.45% of THAA in the powder), although the maximum background amino acid concentrations overlap with the minimum THAA concentrations found in the water.

Only a small proportion of the THAA lost are present in the water (Figure 2.8). Despite over 50% loss from the powder after 48 hours heating, only 2.3% of the original amino acid pool is present in the water (Figure 2.8). 1% - 10% of this sub-pool is also likely to reflect background amino acids levels in the water. Pilot experiments with 72 hour bleached eggshell showed no significant difference in

either the proportion of amino acids lost or the proportion present in the water (data not shown).

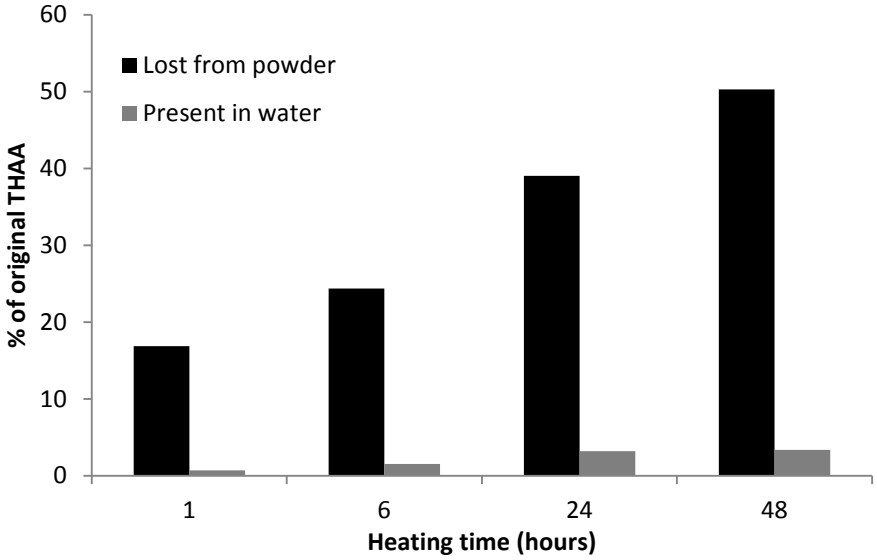

Figure 2.8: Proportion of original amino acid pool in powder and water following heating of chicken eggshell powder at 140°C.

Despite the fact that some amino acids are lost to the water during heating, these are not present in sufficiently high concentrations to fully explain the observed drop in THAA concentration (Figure 2.8). This suggests that a large proportion of amino acids are also lost by decomposition. This is consistent with the increase in proportional representation of alanine (and decrease in serine) with heating time; alanine is a decomposition product of serine (Vallentyne, 1964). It is also probable that some amino acids decomposed after leaching into the water, meaning that the leaching values observed may be misleadingly low. This may be supported by the increasing proportions of free alanine and glycine in the water during the early stages of heating (Figure 2.9), as these are both decomposition products of other amino acids, and are also the smallest and most likely to be lost by leaching.

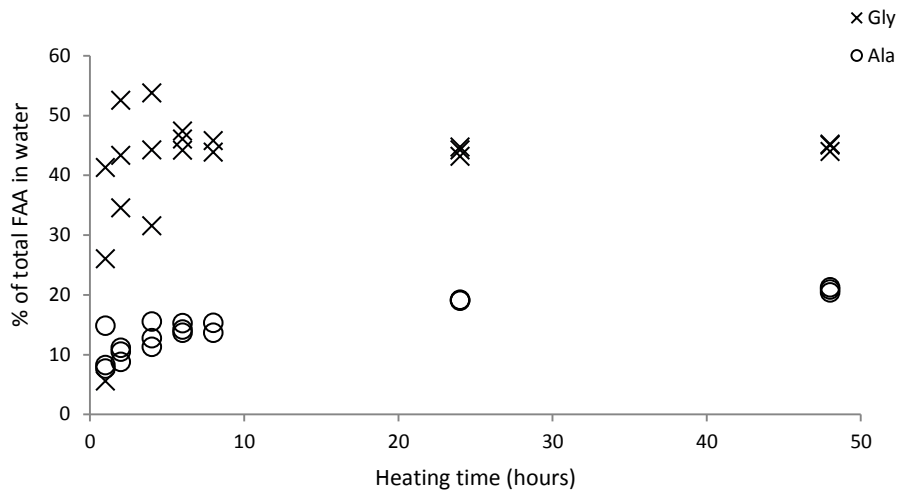

Figure 2.9: Proportion of FAA in water comprising glycine and alanine vs. heating time at 140°C

This does not address whether these residues were produced by decomposition within the mineral and then lost by leaching, or were products of other residues already lost from the mineral. Glycine and alanine are produced within the intra-crystalline fraction as decomposition products (e.g. Vallentyne, 1964; Walton, 1998). The proportion of free glycine and alanine residues lost by leaching decreases with heating time (due to rapidly increasing FAA concentrations in the powder) (Figure 2.10). A maximum of 7.25% (Gly) and 7.42% (Ala) is lost from the FAA fraction in the powder during early diagenesis: this in turn represents only 0.2% - 0.9% of total glycine, and 0.15% - 0.5% of total alanine. This limited diffusive loss of the smallest and most mobile amino acids suggests that while some leaching may occur, it is severely limited.

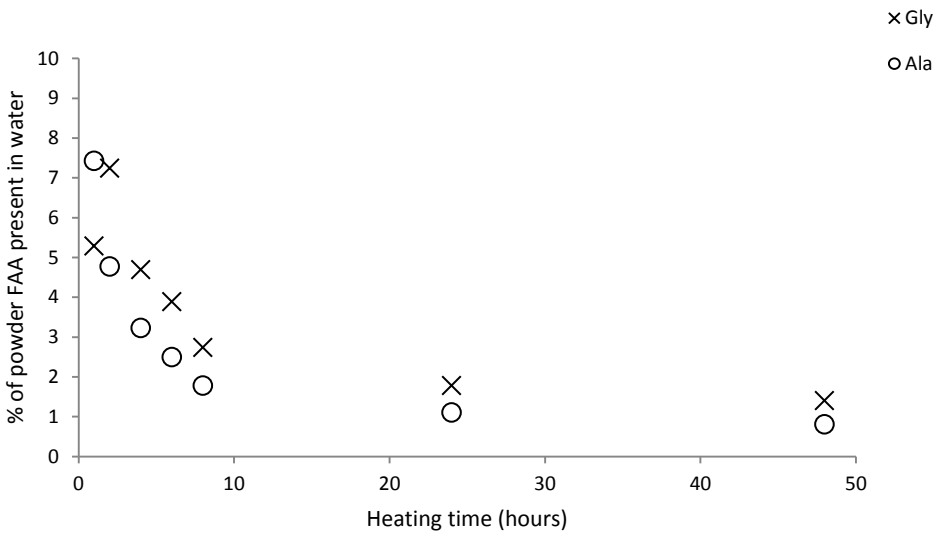

Figure 2.10: Free Glycine & Ala loss by leaching. Limited diffusive loss of the smallest and most mobile amino acids suggests that leaching from the intra-crystalline system is very limited.

2.1.4: Conclusions on bleaching and leaching tests

The apparent resistance to oxidation of a sizable proportion (25% - 35%) of the amino acids in both types of shell hints at possible closed system behaviour. However, the racemisation behaviour of this fraction has not been unequivocally shown to conform to expected closed-system behaviour. Despite this, the leaching test demonstrated pseudo-closed-system behaviour. Although it has been shown that the system is not completely closed, particularly during later diagenesis, diffusive loss of amino acids appears to be very low. This suggests a molecular system which is protected from external influence, and which can provide a relatively stable diagenetic environment. This allows the subsequent experimental work in this thesis to target this system. The following section details the development of an identification system for archaeological eggshell based on intra-crystalline proteins.

Section 2: Developing a taxonomic identification system for archaeological eggshell

2.2.1: Introduction

This section will describe the development of the core aim of this project; production of an analytical system for identification of eggshell fragments. A series of experiments sought to determine whether protein could be recovered from the intra-crystalline fraction of avian eggshell; to establish the analytical power and limits of detection of the instruments used; to test whether these peptides differed sufficiently between taxa to allow identification; and to optimise the laboratory preparation and analysis used in the technique to facilitate production of a high-throughput identification system.

2.2.2: Establishing limits of detection

2.2.2.1: Overview

Before archaeological eggshell could be considered for analysis, the protein concentrations present in the intra-crystalline fraction of eggshell had to be shown to be sufficient to allow detection using the mass spectrometry (MS) techniques used in this study. As a starting point, initial pilot experiments sought to determine the minimum weight of fresh chicken eggshell required to give an intelligible MS signal.

2.2.2.2.1: Methods: extraction of peptides

As this experiment was performed prior to optimisation of the preparation protocol, it used a standard protocol created for preparation of ostrich eggshell

(B. Demarchi, *pers. comm.*). Modern chicken eggshell (5 shells from the same batch) was washed by hand using deionised water, and the shell membrane was carefully removed. After the shells had air-dried, the shells were individually powdered using a quartz pestle and mortar (fragment size ≤ 500µm). Sixteen samples of four approximate weights (10 mg, 5 mg, 2.5 mg and 1 mg) were weighed out for initial analysis. These were exposed to strong bleach (50 µl/mg) for 48 hours in order to remove the inter-crystalline proteins and amino acids. The bleach was pipetted off, and the samples were rinsed several times in deionised water and briefly incubated in HPLC-grade methanol in order to remove any residual bleach before being left to air-dry.

Samples were demineralised using 2M HCl until demineralisation was complete. pH was then balanced to 7.5-8 using 1.5M NH_4OH and 2M HCl and pH strips. Proteins were denatured by addition of 1M dithiothreitol (DTT) (1µl per 200µl sample), which reduced disulphide bridges between cysteine residues. Samples were then incubated at 60°C for one hour, before cysteine residues were alkylated using 0.5M iodoacetamide (IAA) (1µl per 33µl sample) in order to prevent formation of new disulphide bridges. Samples were then left in the dark at room temperature for 45 minutes. pH was then re-balanced to 7-8 using 1.5M NH_4OH and 2M HCl.

4µl of trypsin solution (0.5µg/µl) was then added to the samples, along with acetonitrile (ACN) at a concentration of 8%. Samples were then incubated overnight at 37°C, before a further 2µl of trypsin solution was added the

following morning. Digestion was stopped using acidification; trifluoroaceteic acid (TFA) was added at a concentration of 0.5-1%. The samples were then ready for solid phase extraction. The use of trypsin (and other enzymes) to cleave protein chains into peptides (typically < 5kDa) within the preferred mass range of mass spectrometers is a common procedure (Olsen et al., 2004). Trypsin is commonly used because it only cleaves protein chains at particular basic residues (lysine and arginine), and the basic terminal residues of its product peptides sequester charge and help to prevent fragmentation of peptides upon exposure to the laser beam during MALDI ionisation (Wysocki et al., 2000; Olsen et al., 2004; Steen & Mann., 2004).

In initial experiments, solid phase extraction was performed using 100µl C_{18} Ziptips. 100µl of wetting/elution solution (50% ACN in 0.1% TFA) was aspirated and dispensed to waste; this was then repeated. 100µl of equilibration/washing solution (0.1% TFA in ultrapure water) was then aspirated and dispensed to waste; this was then repeated. The sample was then aspirated and dispensed, ensuring that the entire sample passed through the resin bed at least 3 times. Washing solution was then aspirated and dispensed to waste twice. Finally, the elution solution was aspirated and dispensed 10-15 times, before the sample was eluted into a clean vial.

2.2.2.2.2: Methods: analysis

1µl of sample was then combined with 1µl of matrix (α-cyano-4-hydroxycinnamic acid) on an MTP 384 ground steel Bruker MALDI target plate

and analysed on a Bruker Ultraflex MALDI-ToF mass spectrometer. Each analysis was performed in triplicate. MALDI (matrix-assisted laser desorption/ionisation) is particularly useful for fragile, non-volatile molecules such as proteins (Tanaka et al., 1988). This technique (Figure 2.11) uses a laser beam to ionise sample following co-crystallisation of the target peptides with an aromatic acid matrix (α-cyano-4-hydroxycinnamic acid in this project). The matrix absorbs most of the laser energy; this helps to ionise the analyte, and also prevents peptide fragmentation (Henzel et al., 1993). An electrical field accelerates all ions to the same kinetic energy; the time required for them to then traverse a field-free flight zone between a scintillator and the time-of-flight (TOF) detector is a function of their mass-to-charge ratio (m/z), which is a direct function of ion mass. A full description of the machine specifications used is provided alongside the optimised preparation procedure (Section 2.2.3.4.1).

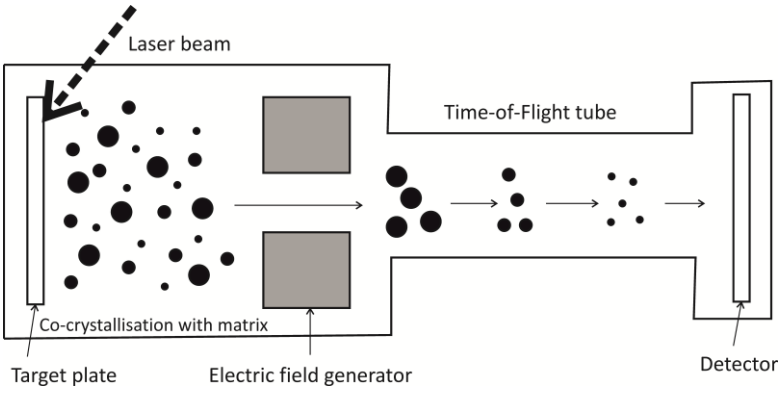

Figure 2.11: Basic mechanism of MALDI-ToF mass spectrometry. Smaller, lighter ions traverse the flight tube and reach the detector before larger, heavier ones.

2.2.2.3: Results

A minimum viable weight of shell was not determined, as even the smallest sample tested (1.16mg) gave a strong MS signal (intensity > $3x10^4$ units - Figure 2.12). Subsequent experiments using shell fragments instead of powder confirm that 1mg of eggshell is more than sufficient to provide clear spectra.

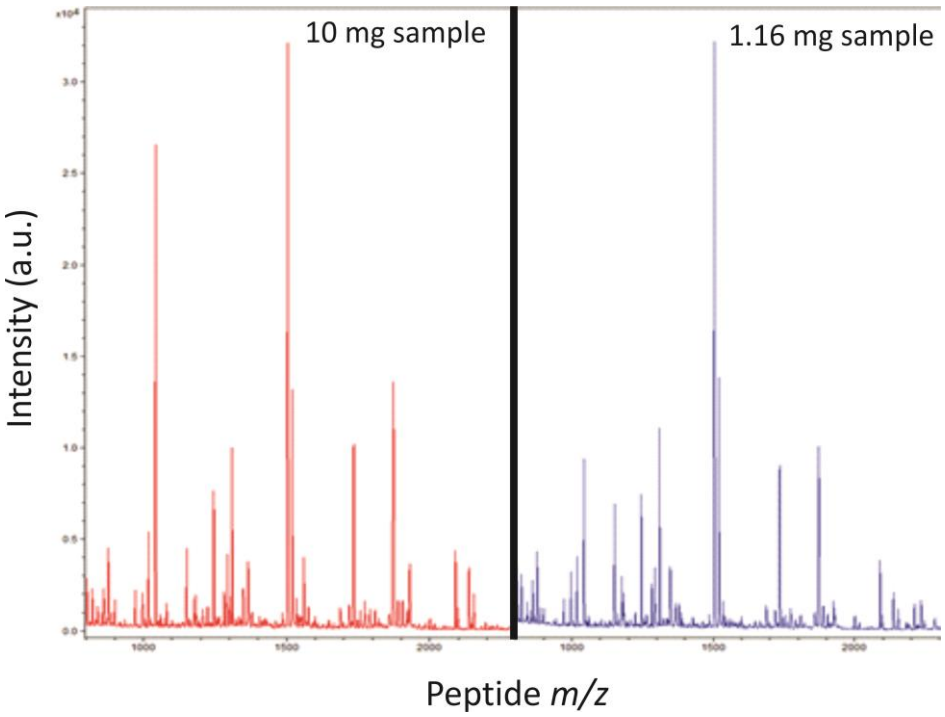

Figure 2.12: MS output from two samples of modern chicken eggshell. Note the similarity in peak intensity between the two samples.

It was decided not to proceed with smaller weights, as in most cases the likelihood of eggshell fragments smaller than 1mg in weight being recovered is slight. The signal strength obtained using 1mg powder and fragments suggests that considerably smaller weights (perhaps as low as 0.1mg) can be prepared using a protocol based on oxidation by bleaching, acid demineralisation and solid phase extraction. This has been subsequently confirmed by the successful

analysis of material from Pompeii, some of which weighed less than 0.2mg (Figure 2.13).

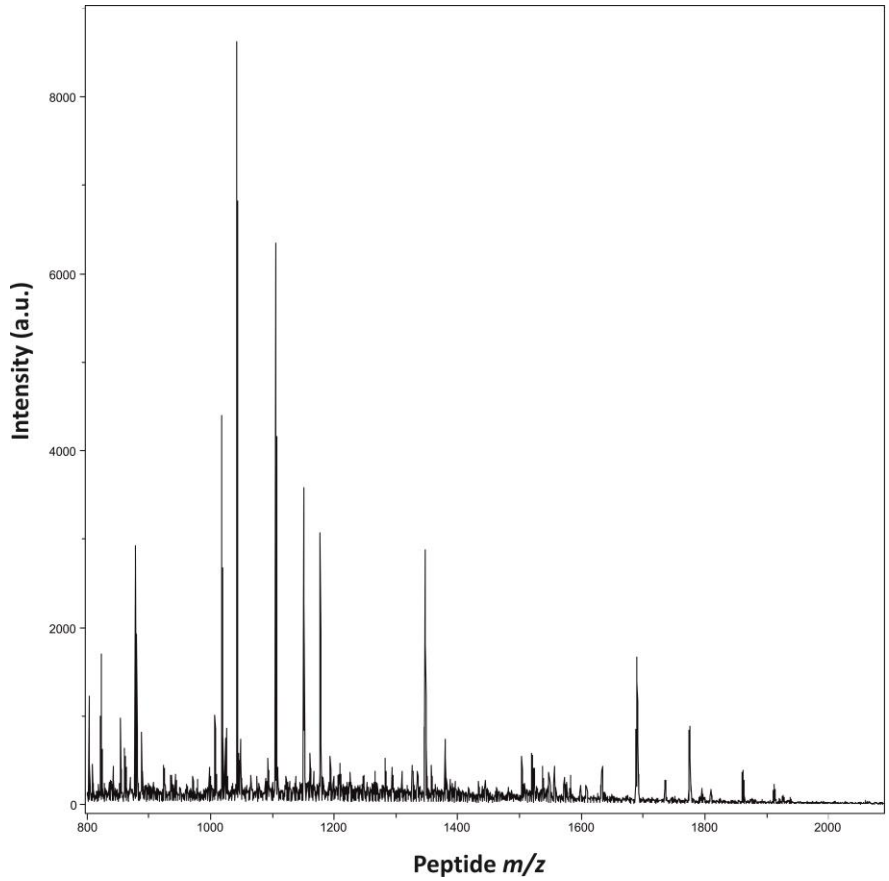

Figure 2.13: Mass spectrum obtained from minuscule (<0.5 mg) fragment of chicken eggshell from c.2000 year old sample from Pompeii. Material provided by J. Thompson (University of Bradford) and excavated during the 2000s.

2.2.3: Optimising the protocol

2.2.3.1: Overview

The core aim of this project is to produce a rapid, high-throughput means of identifying archaeological eggshell. The extraction protocol described above (Section 2.2.2.2.1) requires a large amount of time per sample; high-throughput

data generation is not possible. This necessitates the establishment of a more streamlined sample preparation protocol, in which time-consuming and unnecessary steps are discarded. Following establishment of limits of detection, subsequent pilot experiments sought to maximise the efficiency of the protocol by trialling a range of approaches. A simplified diagram of the process of sample cleaning, preparation and identification is provided below (Figure 2.14).

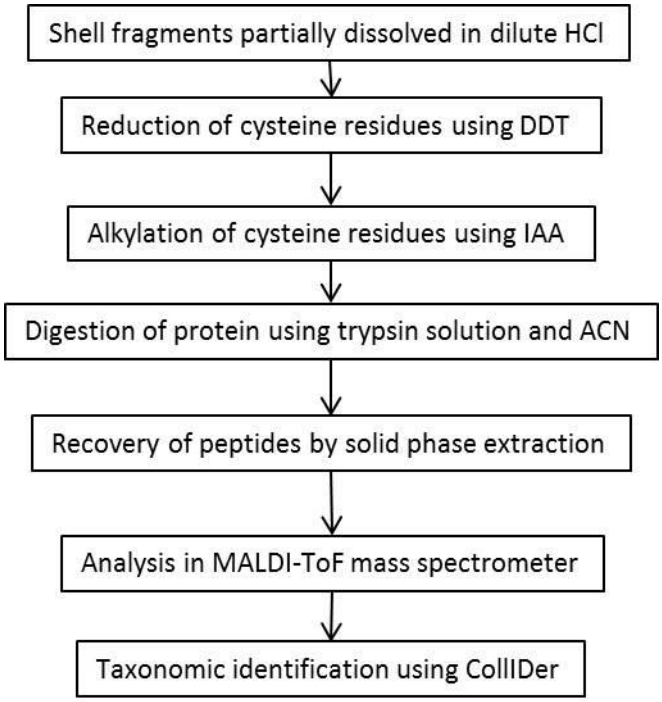

Figure 2.14: Simplified organigram of eggshell preparation and identification process.

2.2.3.2: Methods: protocols tested

The four protocols tested are summarised in the table below (Table 2.1). Four separate samples of modern chicken eggshell were prepared using each protocol. One of the major inefficiencies in the pre-existing protocol described above (method A) is the requirement for pH adjustment after demineralisation

and prior to addition of trypsin: the chemical (DTT – see above) used for reduction of disulphide bonds requires pH ≥ 7, while trypsin digestion requires that 7 ≤ pH ≤ 8. Performing this pH adjustment individually on multiple samples is time consuming, and precludes high-throughput data generation.

Method	Acid (HCl) Volume	Demin. time	pH balanced	Alkylation/ Reduction	Trypsin Digestion
A	10µl/mg (2M)	Instant	Y	Y	Y
B	10 µl/mg (0.6M)	7 days	N	Y	Y
C	10 µl/mg (0.6M)	7 days	N	N	Y
D	10 µl/mg (0.6M)	7 days	N	Y	N

Table 2.1: Protocols tested during method optimisation

In order to address this, the other trialled protocols (methods B, C and D) used dilute HCl (0.6M) to partially demineralise samples over 7 days at a low temperature (4°C). This partial demineralisation allowed the remaining mineral to buffer the solution at neutral pH, removing the need for pH adjustment. Although 48 hours of exposure to bleach is sufficient time to isolate the intra-crystalline fraction (Figure 2.3), these experiments (and subsequent ones) used 7 days of exposure as standard. Method B employed alkylation/reduction of cysteine residues and trypsin digestion; method C did not employ alkylation/reduction; and method D tested whether trypsin digestion was required.

2.2.3.3: Results

Method D, in which trypsin digestion was not used, gave very poor results. The few peaks which were recovered from modern material gave very weak MS spectra (i.e. with low intensity and signal-to-noise (S/N) ratios). None of the peptides which were subsequently identified as being taxonomically diagnostic

(Section 2.2.4.2) were recovered using this method. This shows that protein digestion is required for identification of archaeological material. However, it was also concluded on the basis of the signal strength obtained using methods A, B and C that a slightly lower concentration of trypsin (0.4µg/µl) was sufficient, and that the secondary addition of extra trypsin solution was not required.

In order to quantify the success of the two other methods, the mean signal-to-noise (S/N) ratio was calculated for peptides which had been successfully sequenced from chicken eggshell (described in detail in chapter 5). Using only these peptides allowed omission of peptides which might result from digestion of trypsin, and therefore gave a clearer picture of the relative efficacy of the methods. Method C, in which alkylation and reduction of cysteine residues was not performed, gave consistently weaker S/N ratios than method B (Figure 2.15). Although the same peptides were recovered using both methods, some were less consistently recovered using method C; none of the peptides tested were recovered only using method C.

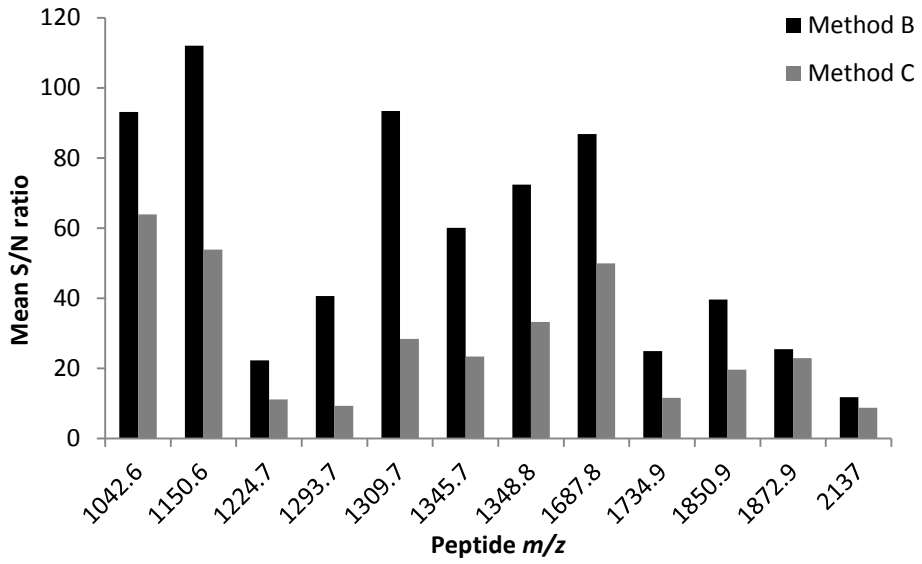

Figure 2.15: Difference in peptide intensity obtained with (method B) and without (method C) alkylation and reduction of cysteine residues.

Given these results, alkylation and reduction was used in all subsequent analyses. The fact that not alkylating and reducing saves a small amount of preparation time does not adequately compensate for the disparity in signal intensity. Even if target peptides do not contain cysteine, its presence elsewhere in the protein may cause unpredicted conformational difficulties if the tertiary structure of the protein prevents the active site of trypsin from reaching the side chain of lysine or arginine (Olsen et al., 2004). Method B was chosen as the optimal protocol, as it required far less laboratory time per sample than method A, but still gave equally clear results. The sample preparation protocol used in subsequent experiments is described in detail below.

2.2.3.4.1: Optimised protocol: sample preparation

In the first instance, this method was employed on modern chicken eggshell obtained from commercial retailers. It has since been shown to work on modern and museum shell in a range of species, and on archaeological eggshell (Chapters 3 & 4). The inner shell membrane was removed by hand, and the shells were washed in deionised water. Once dry, eggshell fragments were weighed, placed in sterile 2ml Eppendorf tubes, and incubated in strong household bleach at a concentration of 50µl/mg over 7 days. The bleach was then pipetted off, and the samples were rinsed several times with ultrapure water. They were then briefly incubated in HPLC-grade methanol in order to deactivate any residual bleach.

After air-drying, the fragments were partially demineralised over 7 days at 4°C using 0.6M HCl (10µl/mg sample). 0.01M dithiothreitol (DTT) was then added (1µl per 2µl sample), and samples were incubated for 1 hour at 60°C. 0.05M iodoacetamide (IAA - 1µl per 3.3µl sample) was then added, and samples were incubated for 45 minutes in the dark at room temperature. 4µl of sequencing grade porcine trypsin solution (0.4µg/µl) was then added to each sample, and acetonitrile (ACN) was added at a concentration of less than 10% of the total solution. The samples were incubated at 37°C overnight. The mechanism of trypsin digestion is described above. Digestion was stopped by acidification after 18-24 hours: trifluoroacetic acid (TFA) was added to the samples so that the final TFA concentration was 0.5-1%.

2.2.3.4.2: Optimised sample preparation protocol – solid phase extraction & analysis

Once a reliable protocol was established, solid phase extraction in all subsequent experiments was performed using BioVyon C18 10mg 96 well plates (Porvair, Fareham, UK). The process and solutions used were identical to those used in the solid phase extraction procedure described above, but the use of a 96 well cartridge and vacuum manifold to draw solutions through the resin bed massively increased the efficiency of the technique (a greater than ten-fold increase). Samples were analysed in positive mode on the Bruker Ultraflex III MALDI-ToF with the following parameter settings: ion source, 25 kV; ion source, 21.4 kV; lens voltage, 9 kV; laser intensity, 35-60%; and mass range, 800–4000 Da. Peptide masses below 650 Da were suppressed. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg^1-Bradykinin, *m/z* = 904.681; Angiotensin I, 1295.685; Glu^1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929). FlexAnalysis software 2.2 (Bruker Daltonics) was used to baseline subtract, normalize spectra and determine peak *m*/*z* values and intensities in the mass range of 800−4000 *m*/*z*.

2.2.4: Deriving taxonomic information from mass spectra

2.2.4.1: Overview

This section has described a method by which peptides can be recovered from eggshell. The next requirement was a system which can translate the mass spectra obtained into taxonomic information. Although this could theoretically be achieved by manually matching mass spectra to reference spectra, this has

the dual disadvantage of requiring a significant commitment of time to each sample, and of introducing subjectivity into the identification process. This subsection describes the compilation of the reference database (from the collection described in Chapter 3), and novel software developed in order to automate taxonomic identification by peptide mass fingerprinting (PMF).

2.2.4.2: Peptide mass fingerprinting

Peptide mass fingerprinting (PMF) is one of the most common applications of MALDI-TOF mass spectrometry. PMF was developed during the early 1990s, and facilitates protein identification by screening experimentally derived peptide masses against a database of known protein sequences (Henzel et al., 1993; James et al., 1993; Pappin et al., 1993). This provides a ‘mass profile’ or 'mass fingerprint' for any given protein, which can be sufficient to allow positive identification (James et al., 1993; Pappin et al., 1993; Hollemeyer at al., 2007, 2008; Buckley et al., 2009). The major limitation of PMF is that is provides a list of peptide masses, without sequence information (e.g. Bienvenut et al., 2002; Yergey et al., 2002). Although a good reference collection (Chapter 3) can allow identification of taxa not represented in current protein databases, the proteins themselves cannot be positively identified using PMF.

In order to construct a platform for PMF in this study, samples were prepared and analysed using the optimised preparation protocol described above. Reference spectra were successfully obtained for all species in the reference collection (Chapter 3); the reference database has been previously published

(Stewart et al., 2013), and upon final submission of this report will be made freely available online (along with the searching software). A Microsoft VB application (CollIDer) was then developed to screen a look-up table based upon the observed peptide masses (Stewart et al., 2013). The code and look-up tables for this application have been published (Stewart et al., 2013). In order to minimise bias, an equal number of spectra were used to represent each species in the look-up list. This application takes a two-pronged approach to PMF. The first approach is to screen the entire dataset and report the species or group of species which provide the closest match. This provides a first pass at identification, but must sometimes be treated with caution as different species have different numbers of peptides naturally present. For example, if a sample contained barnacle goose and arctic tern eggshell, the VB application would report barnacle goose as a much more likely match because there are more peaks to *be* matched in this species. This necessitates the use of a second approach to PMF which screens a list of taxonomically diagnostic masses and reports all matches. In order to identify peptide masses which could be considered taxonomically diagnostic, the entire reference collection (1442 separate peak *m/z* values) was screened (S/N ≥ 6). An example of the output provided by this method is given below (Table 2.2).

PMF Signature			
Species	**Peptides Observed**	**Total Peptides**	**% of Peptides**
Pochard	16	28	57.1
Eider	12	28	42.9
Pintail	11	28	39.3
Mallard/Domestic Duck	10	28	35.7
Barnacle Goose	8	28	28.6
Mute Swan	7	28	25.0
Biomarkers Identified			
Peak *m/z*		**Taxonomic ID**	
825.3		Anseriformes	
1306.6		Anseriformes	
1372.6		Goose/Duck	
1392.9		Anseriformes	
1528.7		Anseriformes	
1723.7		Duck	
1739.8		Duck/Swan	
1755.8		Goose/Duck	
1851.8		Duck/Swan	
1867.9		Duck/Swan	
2051.9		Anseriformes	
2221.0		Duck	
2237.0		Duck	
2378.2		Goose/Duck	
2452.3		Duck (Eider)	

Table 2.2: Example of output provided by software developed for eggshell identification. In this example of archaeological eggshell from Bornais (Section 4.2), the combination of PMF and biomarkers show that this represents duck eggshell. There is one biomarker suggesting that it is eider duck, but this is insufficient for a positive identification to species level in this case.

2.2.4.3: Resolution and temporal span of the new technique

The level of taxonomic resolution varies; for example, it is impossible to identify the very closely related members of the gull family (Laridae) to species using this

technique. Ducks and geese can be separated, but different species of each usually cannot; while chicken and turkey are readily distinguishable. An additional caveat is that the identifications are made from peptide masses alone, without knowledge of the protein(s) from which they are derived. As the proteins differ between species, their survival rates are bound to differ. For this reason, it is difficult to estimate the expected temporal span of this technique in the almost complete absence of sequence information. An attempt is made to model this in chicken eggshell in a following chapter (Chapter 5).

2.2.5: Conclusions on this section

This section has described the development of an accurate, high-throughput taxonomic identification system for small fragments of eggshell. The application of this technique to archaeological eggshell is described in a subsequent chapter (Chapter 4). The following chapter describes the construction of the reference collection which facilitates the subsequent analyses.

Chapter 3: Building the reference collection

3.1: Overview

Subsequent developments in this study are dependent upon the construction of the reference database using the techniques described in the previous chapter. This hinges on the availability for analysis of the eggs of a range of birds; this chapter describes the development of the reference collection which facilitates fulfilment of the core aim of this project. This was undertaken during the early stages of this project and involved cataloguing of the Yorkshire Museums Trust (YMT) egg collection, and working closely with museum curators and staff in order to identify and target specimens suitable for destructive analysis.

3.2: Introduction

The need for a research project to destructively sample biological samples creates a difficult situation for both museum curators and researchers. In many cases, particularly at smaller regional museums which may not function explicitly as research institutions, there is no established framework by which curators can judge requests for material. Also, many biological collections remain incompletely catalogued. This chapter describes the production of the reference collection for this project. Attention is drawn to the particular problems inherent in identifying specimens suitable for destructive analysis, and the steps taken to minimise those problems in this study are described. This chapter also further reinforces the scientific value of retaining data-poor biological materials for destructive analysis (e.g. Russell et al., 2010).

3.3: Museum collections as a research resource

For many types of biological specimen, museum collections represent the only viable, legal and ethical source of material for research. Differing attitudes to collection of biological specimens in the past, when collectors were largely not restricted by law or by conscience, have led to concentration of various types of biological materials in museum collections. Due to the more enlightened attitude towards collection of wild specimens prevalent in modern British society, these now represent an irreplaceable resource.

Egg collections provide a good illustration. Egg collecting was a popular pastime during the late Victorian and Edwardian eras, when many collections numbered in the thousands (e.g. Manson-Bahr, 1959; Lightman, 2000). The practice was outlawed in the 1950s (Protection of Birds Act, 1954), and in the early 1980s it became illegal for individuals to possess an egg collection, regardless of whether they had personally collected the material (Wildlife and Countryside Act, 1981). This led to concentration of egg collections in museums; these are now the most accessible source of material for any researcher wishing to study eggs of non-domestic species (e.g. Russell et al., 2010).

3.4: Destructive analysis of museum specimens

Many museums have a long tradition as research institutions, and most fulfil this role to the extent of their capabilities. The last few decades have witnessed significant evolution of scientific techniques; the rate of progress in molecular

technology has been particularly marked. This continuing progress has opened up a wide range of scientific possibilities, and has led to a concomitant increase in requests for destructive analysis of museum specimens.

The need to destroy museum specimens, in whole or in part, places curators in a difficult position. While they must aim to ensure the continued value of their collections, of which research potential forms a significant part, they also have a duty of care and must be certain that a request is appropriate before releasing material. Often, there is no established framework to assist them in making this decision.

Another problem has been that the factors driving research are often not easily compatible with rigorous collection care. For example, research is often driven by a desire (or need) for high-profile publications; many researchers move automatically towards the most high-impact specimens available. Not only does this lead to requests for destructive analysis of the most irreplaceable specimens in the museum collection, it often also leads to analysis being conducted prematurely in a bid to beat competitors to publication. When such requests are granted, the destructive sampling of the material has often been more detrimental than was necessary. When this approach to requesting and sampling material is combined with a failure to systematically feed back negative results to curators, a situation arises in which curators become hostile to the concept of destructive analysis.

3.5: Cataloguing the YMT egg collection

In order to tackle these problems, dialogue between the YMT curator (I. Gladstone) and the author began before even initial preparation of a sampling request had begun, and both were involved in shaping the request for material. Steps were taken to minimise the damage done to the museum specimens, and to give the curator full control over the decision making process; these are described below.

The first potential problem faced was that the YMT egg collection was incompletely catalogued. As the curator did not know the nature of most of the collection, she could not guarantee the availability of any material for destructive analysis. In order to address this problem, a full catalogue of the egg collection was required; this was produced during the first year of this project. Although this project required only taxonomic information, all available data were included in the catalogue in order to aid the curator in judging the suitability of specimens for destructive analysis. The catalogue of the collection includes a wide range of parameters, any combination of which may be present for any given specimen; these include taxonomic information, collection information (e.g. date, location, and collector), original clutch size, and condition. This provided the Museum with a catalogue of a previously un-quantified collection, and allowed the author and the curator to begin to shape a request for material. Taxonomic identifications provided alongside the material were cross-checked by comparison with identification guides and with conspecific material in the collection (where available).

3.6: Identifying specimens for destructive analysis

The production of the catalogue allowed the curator to begin to rank specimens in order of their suitability for destructive analysis, but further questions needed to be addressed before sampling could proceed.

The amount of material which was needed from each egg, and how much could be approved for destructive analysis, first had to be established. In order to address this, research was conducted using commercially available eggs to establish the smallest possible amount of material from which peptides could be recovered (Section 2.2). Fortunately, the protein content of eggshell is high, and sufficient concentrations were recovered from very small (<1 mg) fragments of shell (Figure 3.1).

Figure 3.1: Gull egg from YMT collection prior to and after destructive sampling. The small sample size required minimises damage to specimens.

Another issue which needed to be addressed was the number of samples needed per specimen, and the number of specimens needed per species. Based on

analyses of domestic species, it was found that proteomic content is remarkably consistent both between different eggs and in different parts of the eggshell of the species tested (Table 3.1). This minimises the number of specimens required; for specimens which are not common in the collection, and for which only very limited sampling is possible, a single sample from a single specimen is sufficient.

Peptide *m/z*	Standard Deviation	Samples
1042.6	0.037	48
997.5	0.041	23
1293.7	0.036	44

Table 3.1: Variability in the mass of selected peptides obtained from chicken eggshell. These represent a number of eggs from different batches, and a range of different parts of the eggshell. Although not all peptides can be shown, these low standard deviations are representative.

Several avenues were explored in an attempt to identify the most suitable means of actually removing samples from specimens in the museum collection. Eggshell is a very brittle, delicate material and so great care had to be taken when sampling. The process used in this project exploited the fact that eggs in the collection already had a hole through which the original contents had been blown by the collector (Figure 3.1). Using very fine scissors, a small piece of shell was removed from around this hole. This allowed sampling to be carried out with only a minimal impact on display potential.

There is also a spectrum of quality in the data attached to specimens in the YMT egg collection. At one end of this spectrum, some eggs have a wealth of information attached, including taxonomic information, the date of collection, the identity of the collector, the original clutch size, and the location of collection. At the other end, some specimens had only taxonomic information; in a few cases, there were no attached data at all. Only taxonomic information is required by this project. Also, the physical condition of the specimens is unimportant; this allowed prioritisation of eggs which were already cracked or broken.

This background work facilitated the production of a 'bottom-up' approach to identifying samples suitable for destructive analysis, whereby the curator could identify the least data-rich, worst-preserved specimens for each species when considering which material to release. The background work performed as part of the sampling request created a situation conducive to production (and subsequent approval for the most part) of a streamlined, well researched and realistic destructive sampling request.

3.7: The reference collection

The list of taxa represented in the reference collection developed for this project is given as an appendix (Appendix A). The majority of these were drawn from the YMT collections; some samples were also provided later by the Natural History Museum, the Zoological Museum of the University of Copenhagen, and by J. Sidell (English Heritage). The reference collection currently comprises 58 species representing 12 orders (Appendix A).

3.8: Conclusions on chapter 3

This chapter has described the development of a reference collection for this project. This is of fundamental importance to the successful execution of this project, and facilitates the developments described in subsequent chapters. This chapter shows the value of early dialogue between researchers and museum curators when attempting to source material for destructive analysis, and also further highlights the scientific value of museum collections, even when data-poor. The following chapter will describe the application of both this reference collection and the techniques described in the previous chapter to archaeological assemblages.

Chapter 4: Applying the new identification technique

Section 1: Egg use in Anglo-Scandinavian York

4.1.1: Introduction

4.1.1.1: Overview

This chapter describes the first application of the new technique developed in the previous chapters to archaeological eggshell assemblages. The focus of the first case study presented is Anglo-Scandinavian York; background information regarding the history of egg use and previous archaeological applications has been provided in a previous chapter (Chapter 1).

4.1.1.2: Anglo-Scandinavian York

Danish Vikings arrived in York during the mid-9th century A.D. and held the city until 954 A.D., when it came back under English control. The period between the mid-10th century and the arrival of the Normans (who themselves were second generation Norse settlers) in the mid-11th century is referred to as the Anglo-Scandinavian period; York was nominally under the control of an English king, but its culture was still heavily influenced by the Scandinavian settlers. This case study focusses on a particular aspect of the domestic economy of the city during this period; in order to put this in context, some brief details of the types of food used in the city must be provided.

4.1.1.3: Subsistence strategies

During the Anglo-Scandinavian period in York, the vast majority of the meat consumed was sourced from domestic species; mid-sized domesticates such as cattle, sheep and pigs formed the basis of the meat component of most diets, and cattle in particular were predominant (O'Connor, 1989, 2000). The surrounding Vale of York is very flat and fertile, and provides excellent conditions for growing many crops and particularly for raising cattle (O'Connor, 1989). Although a wide agricultural area provisioned the city, it may not have been a large enough market to drastically alter the nature of activity in the surrounding agrarian area; many of the bone remains suggest that the meat consumed in the city consisted in what was no longer needed in the wider area, rather than having been raised especially for consumption in York (O'Connor, 2000). Smaller species such as chickens, geese and pigs were probably raised within the city itself (O'Connor, 1989). There is very little evidence for use of wild foods such as venison, although a minor contribution was made by wild-fowling, and both riverine (e.g. eel) and marine (e.g. herring, cod) fish made significant dietary contributions (O'Connor, 2000). Dairy products were probably important, but evidence for cattle and goats within the city is limited (O'Connor, 1989). The importance of eggs in the economy of Anglo-Scandinavian York has only been inferred from the prevalence of chicken bones (O'Connor, 2000), and direct evidence is completely lacking; this case study will address egg use in the city during this period, and so contribute to understanding of the manner in which the city was provisioned. In order to achieve this, eggshell assemblages from two major Anglo-Scandinavian sites in York were analysed.

4.1.1.4.1: The Sites: Hungate

Hungate is a large, multi-phase site located near the centre of York, abutting the north bank of the River Foss (Figure 4.1). The Dig Hungate excavations conducted by York Archaeological Trust (YAT) began in late 2006 and will be completed in line with the Hungate (York) Regeneration Ltd. development schedule. The source of the material used in this study, Block H, was excavated between 2007 and 2011. Most of the contexts evaluated in this study are provisionally assigned to Anglo-Scandinavian age activity (unless stated otherwise). During the Anglo-Scandinavian period (late 9th – mid 11th centuries), the site is thought to have been relatively low-status compared with the contemporaneous Coppergate site (see following section). Over 2000 fragments of eggshell were recovered from the excavation by 5mm and 1mm sieving, which was performed routinely on samples of most types of deposit.

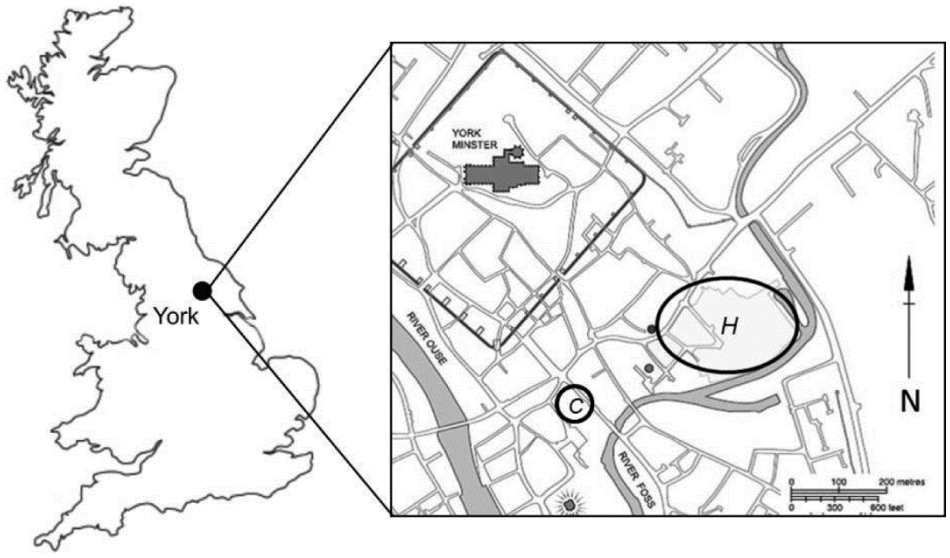

Figure 4.1: Location of Hungate (H) and Coppergate (C) sites. Reproduced courtesy of York Archaeological Trust. Based on the Ordnance Survey mapping © Crown copyright

4.1.1.4.2: The sites: Coppergate

Coppergate, which is located around 350 m to the south-west of Hungate (Figure 4.1), was excavated by YAT between 1976 and 1981. A site of activity during the Roman period, Coppergate was apparently deserted during the post-Roman period, and became active once more with the onset of the Viking/Anglo-Scandinavian period (mid-9th Century) (O'Connor, 1989). During the early Anglo-Scandinavian period, there is evidence for glass working and possible structures; these were definitely established at the site by the mid-10th Century (Hall, 1989). There is also evidence of iron working at the site during this period (Hall, 1989).The areas to the rear of these structures contained a large number of pits, in which organic preservation was often excellent (Hall, 1989). Relative to Hungate, Coppergate is considered a high-status (although not elite) site on the basis of the type of industrial activities, finds and structures excavated. 758 fragments of eggshell were recovered from the site by YAT excavators, and were analysed using the technique described in a previous chapter.

4.1.2: Methods

4.1.2.1: Analysis and identification

All eggshell fragments from Hungate and Coppergate (n > 2750) were analysed using the optimised protocol developed in a previous chapter (Section 2.2.3.4). Fragments were taxonomically identified using the Microsoft VB application described in the same chapter (Stewart et al., 2013). Multiple fragments were

routinely included in a single analysis; the technique is capable of detecting multiple taxa in any given sample.

4.1.2.2: Establishing a method of quantification

As for all heavily fragmented biomaterials, quantification of the resulting data presents a challenge. The percentage of eggshell fragments representing any given taxon is not a reliable indicator of the relative frequency of use, as it does not account for differential pathways of egg fragmentation, which are unknown and probably impossible to quantify with any degree of confidence. For example, chicken eggshell is thinner and more brittle than goose eggshell, and can be reasonably expected to fragment into a higher number of pieces. Assessing only the *presence* or *absence* of any given taxon in a context is a more appropriate method of quantification (e.g. Stewart et al., in press). For Hungate, where successful identifications were not made for all contexts, the percentage representation per context is calculated as the proportion of the contexts in which successful identification was achieved.

4.1.3: Results

4.1.3.1: Success of the new technique

In total, over 2750 separate eggshell fragments were analysed. Analysing this volume of material would not be practicable using previously available techniques (Sidell, 1993a, b). Successful taxonomic identifications were achieved for 35 of 39 contexts at Hungate (89.7%), and all 29 contexts at Coppergate. At Coppergate, the success rate of the technique by fragment was 98.3% (12 of 758 fragments

remain unidentified); comparative data are not available for Hungate, but this figure is lower. Where no identification was made, this was due to poor quality mass spectra rather than inability to match good spectra to the reference database.

4.1.3.2: Characterising the eggshell assemblages

Only two taxa were identified from the Anglo-Scandinavian deposits at Hungate: domestic chicken (present in 23 of 26 (88.5%) of contexts from which eggshell was identified) and goose (present in 4 of 26 (14.3%) contexts (Figure 4.2). No other taxa are observed either in the Anglo-Scandinavian levels or subsequent levels, and these figures are similar to those obtained when the whole Hungate assemblage is considered (Chicken: 32 of 35 contexts (91%); Goose: 5 of 35 contexts (14%)). This technique cannot presently distinguish between wild and domestic geese, or even between different goose species; this is due to the very high degree of protein conservation within the order Anseriformes. It can be stated that no unambiguously wild taxa are present in the eggshell assemblage, and given that domestic geese were kept in the city it seems reasonable to propose that at least the majority of the goose eggshell might represent domestic birds. Chicken is equally ubiquitous at Coppergate, where it is present in 27 of 29 eggshell-bearing contexts (93%) (Figure 4.2). Goose eggshell is far more prevalent at Coppergate (12 of 29 contexts, 41%) than at Hungate (Figure 4.2). The only appearance of duck eggshell in this study was also in a context from Coppergate (Figure 4.2); as for goose, it is unclear whether this represents wild or domestic duck. This represents the first successful application of a high-throughput proteomics based identification technique for archaeological eggshell.

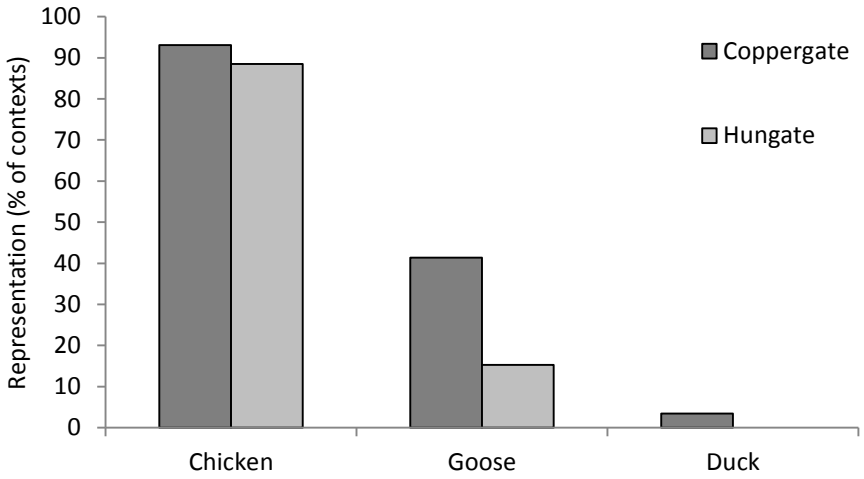

Figure 4.2: Taxa identified in Anglo-Scandinavian eggshell assemblages from Hungate and Coppergate, York.

4.1.4: Discussion

4.1.4.1: Success of the new technique

Although the success rate of the new technique is high (Section 4.1.3.1), a slight disparity was observed between the two assemblages. Several non-mutually-exclusive factors may explain this. At face value, one of the simplest possible explanations is a simple age effect, where the proteins present in older specimens are more degraded than those in younger specimens. This may seem a logical explanation, but some of the Hungate contexts from which no identification was made are among the youngest at the site (T. Kendall, pers comm). Other factors which may contribute to this disparity include better organic preservation in general at Coppergate; improvements in resolution and execution of the technique between the two analyses (the Hungate analysis was the first conducted using the new technique); possible burning of some shell fragments at

Hungate; and misidentification of very small fragments of plaster or mollusc shell as avian eggshell at Hungate (A. Jones, pers comm).

4.1.4.2: Advantages of the new technique

This project develops ZooMS as a robust, reliable, accessible and rapid means of identifying archaeological eggshell fragments. This is not the only means being developed for taxonomic identification of eggshell; the relative advantages and disadvantages of ZooMS compared to other techniques must be determined. Recent advances have facilitated recovery of DNA from preserved eggshell, and shown that it can be very useful as a taxonomic identifier; DNA analysis is potentially capable of a greater degree of resolution than the system developed in this thesis (Oskam et al., 2010, 2011; Coghlan et al., 2012). Nonetheless, there are several major advantages of the ZooMS technique. For one, it can provide taxonomic information for very small (<1mg) fragments, allowing analysis of entire eggshell assemblages and subsequently more robust quantification; current methods for isolation of DNA from eggshell require far more material, and many of the eggshell fragments analysed during this study (and subsequent studies) are simply too small to allow sufficient DNA recovery using current techniques (Egloff et al., 2009; Oskam et al., 2010). Unlike some promising DNA techniques (e.g. Lee & Prys-Jones, 2008), the ZooMS method does not require preservation of shell membrane or cuticle, both of which are often missing from archaeological specimens. It can process entire eggshell assemblages, and has a high success rate. Another major advantage is that, at present, this technique is less costly than DNA analysis. This technique may aid the development of DNA sequencing of

archaeological eggshell by identifying fragments which are most suitable for subsequent analysis.

4.1.4.3: Birds in the Anglo-Scandinavian domestic economy

Birds were a minor but significant component of the domestic economy of Anglo-Scandinavian York; their bones account for between 2% and 7.5% of the animal bones recovered from sites dating from this period (O'Connor, 1989, 2000). As there are taphonomic differences in preservation and recovery of bird and mammal bone (bird bone is generally smaller and lighter), it is possible that this slightly underestimates their actual contribution. Although it is difficult to quantify, it is clear that birds contributed significantly to the diets of people in York during this period. The majority of the recovered bird bones represent domestic chicken, while goose and duck bones are also reasonably abundant (O'Connor, 2000). Although distinguishing domestic and wild types is difficult (particularly in duck) it is thought that these largely represent domestic birds (O'Connor, 2000). In his analysis of the animal bone recovered from Anglo-Scandinavian deposits in York, O'Connor uses the sex distribution of chicken bones to suggest that egg production was probably the main focus of chicken husbandry during this period, rather than production of meat or male birds for cockfighting (O' Connor, 2000).

The relative contribution of wild and domestic species to the bone assemblages differs between urban York (where domestic chicken predominate) and many contemporaneous sites in coastal and island locations, in which seabirds tend to vastly outnumber domestic species (island sites are described in the following

section) (e.g. Sharples 2005, 2012; Harland et al., 2012). Non-domestic species recovered from Anglo-Scandinavian sites in York include a range of wetland birds such as wild ducks and geese, cranes, and a range of waders (O'Connor, 2000). These are commonly winter-flocking species, and were most likely exploited during winter when farming activities placed less demands on time (O'Connor, 2000). Other species represented include cliff-nesting seabirds (razorbill, guillemot), woodland species (woodcock, wood pigeon), and moorland species (golden plover, black grouse) (O'Connor, 1989).

4.1.4.4: The eggshell evidence

The eggshell evidence suggests that domestic chicken provided the majority of eggs used at both sites during the Anglo-Scandinavian period (Figure 4.2). This is consistent with the bone evidence, which suggests that these were the most common bird in avian bone assemblages (O'Connor, 2000). There is also a smaller contribution of goose eggshell at both sites, and a minimal contribution of duck eggshell at Coppergate (Figure 4.2). Although this technique cannot distinguish domestic and wild types, these results are consistent with domestic species providing at least the vast majority, and very possibly all, of the eggs consumed in the city, with at most a minimal contribution from wild species. There is a complete lack of demonstrably wild species in both of the eggshell assemblages (Figure 4.2). Ducks and geese are known to have been kept domestically in the city during the Anglo-Scandinavian period (O'Connor, 1989, 2000). It can therefore be cautiously proposed that the duck and goose eggshell represents domestic species, although this cannot be stated conclusively. The lack of demonstrably wild species in the eggshell assemblage supports the idea that wild-fowl may have been a focus of exploitation during winter (i.e. outside of their

breeding season), when many economic activities were slower than during other seasons and/or additional sources of food or income were required (O'Connor, 2000). The eggshell results from Hungate and Coppergate suggest that in terms of egg use, the economy in Anglo-Scandinavian York was entirely domestic. This provides an interesting contrast between the avian bone assemblages and the eggshell assemblages.

4.1.4.5: Beyond taxonomic identification: can eggshell inform questions of status?

A major difference was observed in the prevalence of goose and duck eggshell at the two sites. It is possible that this results from the difference in status between the sites. Coppergate is not considered a 'luxury' site based on the types of bone identified and the structures and finds (e.g. O'Connor, 1989), but is considered higher status than Hungate (although the degree of social stratification in Anglo-Scandinavian York is unclear). Although direct data for the Anglo-Scandinavian period are unavailable, it has been estimated that chickens in England were producing 70-200 eggs per year by the late medieval period (late 12th – early 14th century), and that members of every social stratum would have had access to these (Stone, 2006; Slavin, 2009). It seems reasonable to propose that these would also have been an everyday food item during the Anglo-Scandinavian period.

The results of this study might begin to suggest that the eggs of geese were higher status or more expensive items in Anglo-Scandinavian society. Direct evidence to

support the notion that goose eggs were a higher status food during the Anglo-Scandinavian period in England is lacking, but some support may be found in roughly contemporaneous accounts from Norse-influenced Ireland. The twelfth century Irish tale 'Fled Dúin na nGéd' suggests that goose eggs were considered higher status fare than chicken eggs (Mac Con Iomaire & Cully, 2007). Direct comparison between Anglo-Scandinavian northern England and Ireland is valid, as major cultural links between the Vikings and Ireland were well established by this stage (e.g. Ó Corráin, 1998, 2001). It seems reasonable to expect a degree of cultural overlap between these regions during the Anglo-Scandinavian period. In Ireland at least, the perception of goose eggs as a luxury food seems to have persisted into the Modern era (Mac Con Iomaire & Cully, 2007).

4.1.4.6: Value of museum collections

Another outcome of this case study is re-emphasis of the research value of museum natural science collections (Chapter 3). This case study adds to the sustained curatorial efforts to establish the scientific research value of museum egg collections which have occurred in recent years (e.g. Russell et al., 2010). In terms of the specific suitability of museum egg collections for proteomic research, one recent study urged caution as some proteins found in modern shell were not recovered from museum material (Portugal et al., 2010). However, this study did not account for the presence of distinct inter- and intra-crystalline protein pools (Chapter 2): it is probable that the proteins which were missing are not commonly located in the intra-crystalline fraction, and/or are particularly susceptible to rapid diagenesis. For the purposes of analyses such as this one, in which recovery of the

full suite of proteins found in modern eggshell is not essential, museum stores represent a unique research resource.

4.1.5: Conclusions

This initial case study on Anglo-Scandinavian York demonstrates that this new technique is viable for use on archaeological material, and also provides some interesting observations on the nature of egg exploitation in the city during this period. The main observation derived here is that the eggs exploited in the city were either entirely or overwhelmingly derived from domestic species, with no input from demonstrably wild taxa. This is consistent with previous assertions that the primary role of domestic chickens was as egg providers; the following section will build upon the proof of concept provided by this case study to conduct an in-depth analysis of egg use in Norse-era island sites in Scotland.

Section 2: Norse-era egg use in Scottish islands

4.2.1: Introduction

4.2.1.1: Overview

The previous section demonstrated the successful application of a new biomolecular technique to archaeological eggshell assemblages, and provided new information on egg use in Anglo-Scandinavian York. This section attempts to further explore the archaeological potential of eggshell by presenting an investigation of egg use in Norse-era sites in the Western and Northern Isles of Scotland. This section contains the bulk of the archaeological investigation conducted during the construction of this thesis. The main questions under consideration are the extent to which the occupants of these sites relied upon domestic and wild species for their eggs; whether patterns of egg use were consistent between different areas and periods in Scotland, and in the wider sphere of Norse influence; and whether Norse egg use differed significantly from that of previous inhabitants of these areas.

The term 'Norse-era' will occur throughout this section and must be defined. The Viking period in Scotland often refers to a period of raiding between the 8th and 9th centuries A.D., while the Norse era, which began in the early 9th century with the arrival of Scandinavian settlers, is often considered to have followed (Crawford, 1987; Graham-Campbell & Batey, 1998). It must be noted that this terminology is not universally accepted, and that there is often not a clear transition between the two periods (e.g. Graham-Campbell & Batey, 1998). By its nature, this study deals with the remains of settlement; the term 'Norse-era' is

adopted here to represent all Scandinavian settlement in Scotland in the period between *c.*800 A.D. and the mid-15th century, when the Northern Isles of Scotland (Orkney and Shetland) were granted by the king of Norway to the Scottish crown as part of a dowry payment. The material representing the 'Late Iron Age' in this section derives from 5th – 6th century A.D. activity.

4.2.1.2: Subsistence in the Norse era Scottish islands

Faunal assemblages and plant remains from Norse-era sites in the Scottish islands evidence a subsistence strategy based on a combination of arable and pastoral farming, with cereal cultivation and dairy production providing important dietary contributions (Crawford, 1987; Morris & Rackham, 1992; Graham-Campbell & Batey, 1998; Sharples, 2005, 2012; Mulville et al., 2005; Barrett, 2012). The most prevalent animals represented in bone assemblages are mid- to large-sized domesticates such as cattle, sheep and pigs (Bond, 2007; O'Connor, 2010; Harland, 2012). Bones of wild species, including marine mammals, deer, and a range of wild birds, also occur in Norse assemblages, but usually form a minor component of the recovered bones (Serjeantson 1988, 2001; Mulville, 2002, 2010; Sharples, 2005, 2012; Harland, 2012; Harland et al., 2012). Marine resources such as fish and shellfish were also exploited (Crawford, 1987; Milner et al., 2007; Harland & Barrett, 2012; Milner & Barrett, 2012). Little is known of the importance of domestic fowl in the Norse economy at sites such as these. Some bone assemblages from Norse-era island sites suggest that domestic fowl (and birds in general) were a relatively unimportant component of the diet compared to larger domestic mammals (e.g. Sharples, 2005, 2012; Harland et al, 2012).

4.2.1.3.1: Norse-era bird exploitation: the evidence from bones

In many Norse-era sites, domestic chicken are found alongside a variety of wild fowl; for example, the 8th – 11th century A.D. settlement of Haithabu in northern Germany produced a huge number of bird bones representing over 60 species, of which domestic chicken formed a relatively small component (13-18%) (Becker & Grupe, 2012). The mostly 9th century A.D. material from Kaupang in southern Norway shows a similar trend; roughly one third of the identified bird bones were found to represent chicken, while the rest of the assemblage comprises a range of wild species (Barrett et al., 2007). It must be noted that preservation of bone at Kaupang was exceptionally poor, and only a small number of bird bones (27) could be identified; this limits the interpretative potential of these data (Barrett et al., 2007). This is consistent with avian assemblages from contemporaneous Norse-era sites in Scotland and Scandinavia, in which wild taxa often predominate over domestic fowl, and in contrast to the urban setting of York, in which domestic chicken predominate (O'Connor 1989, 2000; Tyrberg, 2002; Sharples, 2005, 2012; Harland et al., 2012). For example, at Quoygrew in Westray, Orkney (Section 4.2.1.4.3), chicken bones account for only 4.5% of the identified bird bones (Harland et al., 2012). In the Norse levels at Mound 1, Bornais, South Uist (Section 4.2.1.4.2), only 1.5% of the identified bird bones represent domestic fowl, 10% represent ducks and geese, and 88.5% represent assorted wild taxa (Sharples, 2012). At Freswick Links, Caithness, only 11% of the recovered bones represent domestic fowl (Allison, 1995). All of these data suggest that wild-fowling made a greater contribution to the bird-meat consumed than domesticates (assuming that most of the bones derive from food waste).

In Norse assemblages from Iceland and Greenland, domestic fowl are apparently absent (McGovern et al., 1983, 2006). The absence of chickens from these locations may relate to unsuitable terrain and climate; to local dietary and economic preferences; to low productivity under the prevalent light regime (it is unclear whether the available breeds retained seasonal laying patterns); and/or to the widespread accessibility of wild birds and eggs rendering chickens irrelevant. This latter factor may have been important in parts of Iceland, in which inhabitants made widespread and extensive use of the eggs of certain wild species (McGovern et al., 2006). It might also relate to disposal practices; if bones were routinely cut into very small pieces (for example, to make stock), taphonomic biases may disfavour their preservation and recovery; although this is untestable and unprovable, it should be acknowledged as a possibility.

4.2.1.3.2: Norse egg use

For most pre- and post-Roman cultures, as well as almost all of the European cultures occurring in the period between Roman collapse and the Renaissance, very little is known about egg use (Stadelman, 2000). The information available for the Norse period is consistent with this observation in that it is patchy (both in time and in space), inconsistent, and generally scarce. In Iceland, where the domestic chicken seems to have been absent, eggs of wild species including ducks, geese and ptarmigan were extensively used (more so than adult birds) and could have sustained a large yield (up to 10000 eggs per year) if exploited correctly (McGovern et al., 2006). At Freswick Links on the northern Scottish mainland, a range of wild species were identified in the eggshell assemblage using scanning electron microscopy (Section 4.2.4.3) (Sidell, 1995; Graham-Campbell &

Batey, 1998). It must be noted that some of the parameters used for identification of eggshell in the Freswick Links study may not be valid, particularly when applied to degraded archaeological material. Patterns of egg use are unknown for most regions under Norse influence, and it is unclear whether the patterns observed in the limited number of studies published to date are consistent across the sphere of Norse influence. Unlike their wild progenitors, which retain seasonal laying regimes, modern laying breeds can produce upwards of 250 eggs per year; documentary evidence suggests that they may have done so in England since the late medieval period (12^{th} – 13^{th} century) (Stone, 2006). In the absence of documentary evidence, it is impossible to know at what point in history this ability to decouple laying patterns from light regimes occurred in different areas; for this reason, it is unclear what the egg yield from domestic fowl is likely to have been for the Norse.

Present understanding of egg consumption during the Norse period mostly comprises a series of open questions. Which species were exploited for their eggs? To what extent were domestic and wild resources used, and was this consistent across the geographical and temporal span of the Norse era in Scotland? Were these patterns different from what had gone before in Scotland, and from Norse activity in other areas? Questions such as these are the focus of this section.

4.2.1.4.1: The sites: overview

Four main sites provided the eggshell assemblages which were analysed in the course of this study (Figure 4.3). In total, over 12000 fragments of eggshell were analysed using the technique described in a previous chapter (Section 2.2). These sites provide a range of locations in the Western and Northern Isles of Scotland, and span almost the entire Norse era in the Scottish islands (early 9th century to 15th century). The assemblages from Quoygrew and Bornais (see below) are relatively large (especially the latter), and are also well phased and time-constrained (Barrett, 2012; Sharples, 2005, 2012); these form the backbone of subsequent discussion. The other sites provided smaller assemblages which are of interest as they expand the geographical range of the study, and provide a point of comparison with the larger assemblages.

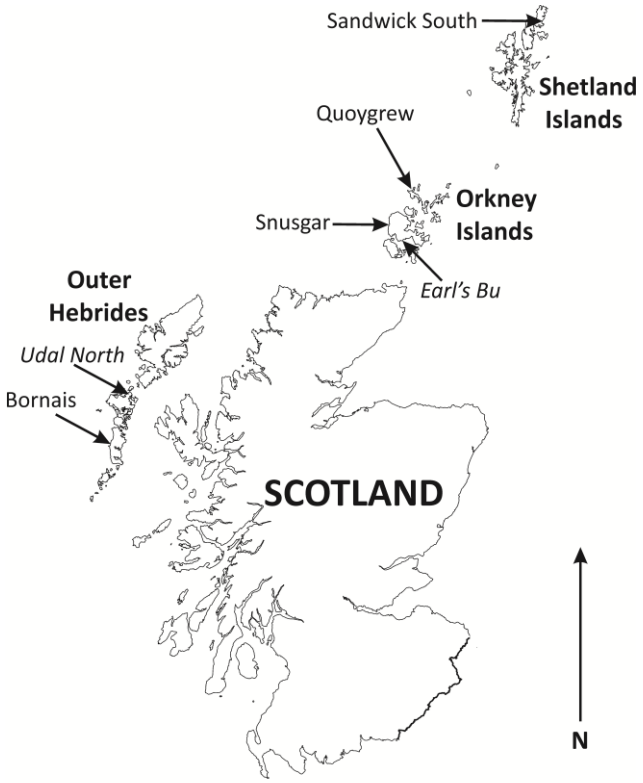

Figure 4.3: Location of sites in the northern and western isles which provided eggshell for analysis. Map modified from OS online resource.

4.2.1.4.2: Bornais, South Uist

Bornais is located in the centre of the machair (calcareous shell sand) plain that forms the west coast of South Uist (Figure 4.3), and was excavated during the late 1990s and early 2000s by a team from Cardiff University (Sharples, 2005, 2012). A very large quantity of eggshell (over 10500 fragments) was recovered and analysed during this study. Four separate areas, and four distinct phases of occupation, produced substantial eggshell assemblages (Sharples, 2005, 2012). Samples representing mound 1 represent two distinct phases of activity; some are associated with a Late Iron Age (LIA) wheelhouse dating to the 5th and 6th centuries A.D., others with Norse midden deposits dating to the 12th to 13th centuries A.D. Samples from mound 2 were recovered from floor deposits from a sequence of three large (and probably high status) houses that span the 10th to 14th centuries A.D. Samples recovered from mound 2A represent a lower status settlement mound and derive from middens and ancillary structures that cover the same period as mound 2. The samples from mound 3 represent another lower status settlement mound similar to mound 2A but are restricted to the final 13th and early 15th century occupation and come largely from house floors.

4.2.1.4.3: Quoygrew, Westray, Orkney

The settlement at Quoygrew, Westray (Figure 4.3) has been the focus of excavation and analysis during the last decade (Barrett, 2012), and has provided a relatively large eggshell assemblage (991 fragments) for analysis. The site was occupied from the early 10th century until the early 20th century; detailed accounts of the structure, phasing, excavation and history of the site have been published

(Barrett, 2012). Within the site, two major areas of Norse-era activity were highlighted (Barrett & Gerrard, 2012). In Areas A-F adjacent to a wave-cut bank on the foreshore, a dwelling structure dating to the 11th and 12th centuries (House 5) was excavated; this was replaced during the early 13th century with another dwelling structure (House 1) (Barrett & Gerrard, 2012). Related to these were a series of midden deposits (the 'fish middens') created during the 11th to 13th century occupation; the eggshell representing this part of the site was derived from these middens (Barrett & Gerrard, 2012). Further inland in Area G, fragments of other houses were excavated along with a series of related midden deposits, collectively forming a farm mound. Although some contexts in this area had been compromised by much later activity, many of the excavated deposits could be confidently assigned to Norse-era activity (Barrett & Gerrard, 2012). For the purposes of this study the material is predominantly from 11th to 13th century contexts of Areas A-F and 10th to 13th century contexts of Area G.

4.2.1.4.4: Snusgar, Orkney Mainland

Snusgar is located near the northern edge of the Bay of Skaill on the Orkney Mainland (Figure 4.3), and has been the focus of recent excavation (Griffiths & Harrison, 2011). One of the hypotheses being tested is that the Bay of Skaill formed an important central place during the Viking and Norse periods (Griffiths & Harrison, 2011). The eggshell (188 fragments) provided for analysis was recovered during the 2010 excavation of occupation layers and middens related to a large Norse longhouse (Griffiths & Harrison, 2011). These deposits have been radiocarbon and OSL (optically stimulated luminescence) dated to between the late 10th and early 13th centuries A.D. (Griffiths & Harrison, 2011). The large size of

the longhouse, the range of ancillary structures and activities performed on site (including metalworking), and the range of finds recovered suggest that the site represents a major farm (Griffiths & Harrison, 2011).

4.2.1.4.5: Sandwick South, Unst, Shetland

Sandwick South is located on the eastern machair of Unst, the most northerly of the Shetland Islands (Figure 4.3). Excavation of the site, which comprises a late Norse era (14th century) farmstead, began in the late 1970s (Bigelow, 1985). Sandwick South also incorporates some Iron Age features which predate Norse occupation, but the layers from which eggshell was recovered are considered secure and date to the early 14th century (Bigelow, 1985). Although the economic and social status of the occupants is not yet clear, the dwelling structure is small relative to the early Norse (9th – 10th century) high-status structures of Jarlshof on the Shetland mainland, and of Underhoull, also on Unst (Bigelow, 1985). The eggshell assemblage from this site is relatively small (172 fragments), but provides a unique insight into egg use in the northernmost of the Scottish islands.

4.2.1.4.6: Other sites

Two other sites also provided a small volume of material for analysis: the late Norse site of Earl's Bu, Orkney Mainland (C. Batey, pers comm), and the complex, multi-phase site of Udal North, North Uist (B. B. Smith, pers comm) (Figure 4.3). Although 79 fragments were provided from Udal North, the complexity of the phasing allowed only 6 to be confidently assigned to the period of Norse influence (B. B. Smith, pers comm). Earl's Bu provided material from only 4 contexts (C.

Batey, pers comm). Although these numbers are too low to allow meaningful intra- or inter-site analysis, the results are included in order to provide a point of comparison.

4.2.2: Methods

Eggshell fragments were analysed, identified and quantified in an identical manner to that described in a previous section (Section 4.1.2).

4.2.3: Results

4.2.3.1: Overview and method of quantification

The results obtained for each site are presented below (Figure 4.4); the data from Bornais (Figures 4.5 & 4.6) and Quoygrew (Figure 4.7) are also considered in more detail subsequently. In general, the success rate of the technique was high (over 92% in all sites); the success rate for the Iron Age deposits at Bornais was slightly lower than for the Norse levels (82%), consistent with the more advanced degradation of these older proteins. In the Norse levels, failure to identify taxa did not correlate with the age of deposits or with any other obvious parameter. The majority of failed analyses are likely to result from fragments being exposed to fire for prolonged periods; this alters the biomolecular composition of the eggshell and accelerates diagenesis (e.g. Crisp et al., 2013). In cases where fragments looked burnt, data were often not recovered (data not shown).

The identification technique used is qualitative; it detects only the presence or absence of different taxa in an assemblage, but does not calculate the relative contribution of each (Section 4.1.2.2). This is necessary as the fragmentation patterns of eggshell are unpredictable and complicated; it is not practicable (and may well not be possible) to reconstruct the number of actual eggs in a deposit from the fragments preserved in the archaeological record. These data are quantified as the *percentage* of eggshell-bearing deposits in which a given taxon is *present*.

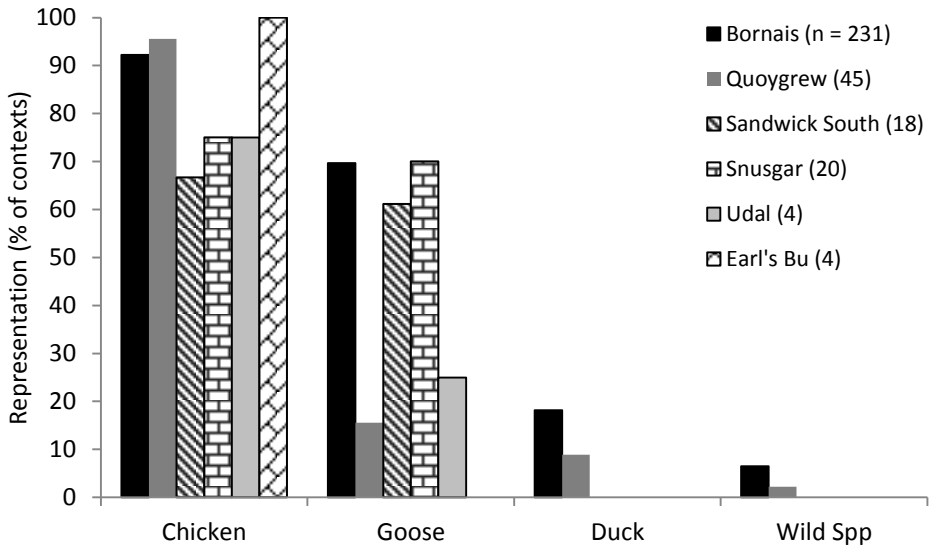

Figure 4.4: Combined data for all sites (n = number of eggshell bearing deposits). The only example of a wild species at Quoygrew relates to pre-Norse era activity.

4.2.3.2: Bornais

Although it provides a useful initial inter-site comparison, Figure 4.4 gives a very coarse-grained view of patterns of egg use. In order to draw more meaningful conclusions from these data, a more detailed breakdown of intra-site patterns is

required; this is only possible for the larger assemblages from Bornais (Figures 4.5 & 4.6) and Quoygrew (Figure 4.7).

When considered by phase, the Bornais data show a marked distinction between the Late Iron Age material from mound 1 and the Norse material from mounds 1, 2, 2a and 3 (Figure 4.5). The Iron Age material from mound 1 represents 33 contexts, and includes a range of wild species (gulls, alcids), ducks (whether wild or domestic is unclear), and small contributions from chicken and goose. Mound 1 also provided eggshell from 7 mid- Norse contexts; chicken and duck were present in 5 of these, geese in 3, and wild taxa (indeterminate charadriiformes) in one (data not shown). In the Norse levels representing all phases of occupation, chicken eggshell is predominant and almost ubiquitous (Figure 4.5). Goose eggshell is also very common, although it is not clear whether this represents wild or domestic species, while the contribution of duck eggshell diminishes through the Norse period (Figure 4.5). Perhaps most interesting is the striking difference in the contribution of demonstrably wild taxa between the Iron Age and Norse assemblages; these are present in only 4.3 – 8.6% of Norse era contexts, compared with over 50% of Iron Age contexts.

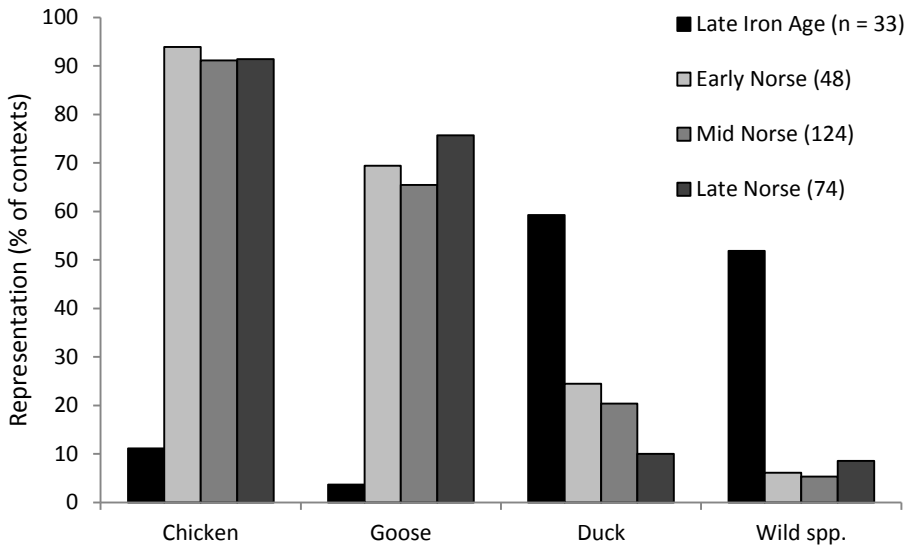

Figure 4.5: Bornais data by phase of occupation, with number of eggshell-bearing contexts given. The early Norse period at this site represents the mid-9th century – mid 10th century; middle Norse represents late 10th – mid 12th century; and the late Norse represents the 12th – 15th century.

Meaningful intra-mound comparisons were only possible for mounds 2 and 2A (Figure 4.6) due to the low number of eggshell-bearing Norse contexts in mounds 1 and 3; this also limits the meaningfulness of inter-mound comparisons and dictates that the data from these mounds is best considered in terms of phase of occupation (Figure 4.5).

In both mound 2 and mound 2A, chicken is present in all early Norse contexts which contain eggshell, and remains prevalent throughout most of the period of Norse occupation. The only exception is in mound 2A during the mid-Norse period (Figure 4.6), when goose eggshell becomes as prevalent as chicken eggshell; the representation of goose eggshell seems slightly higher in mound 2A than in

mound 2 during all phases. The contribution of duck eggshell to the mound 2 assemblage is higher than in mound 2A until the mid-Norse period, when it all but disappears from both assemblages: the mostly likely taxonomic affiliation is eider rather than domestic duck, but this has not been conclusively demonstrated. The only contribution of wild species in mound 2A is during the mid-Norse period, while in mound 2 this period has the lowest representation of wild species.

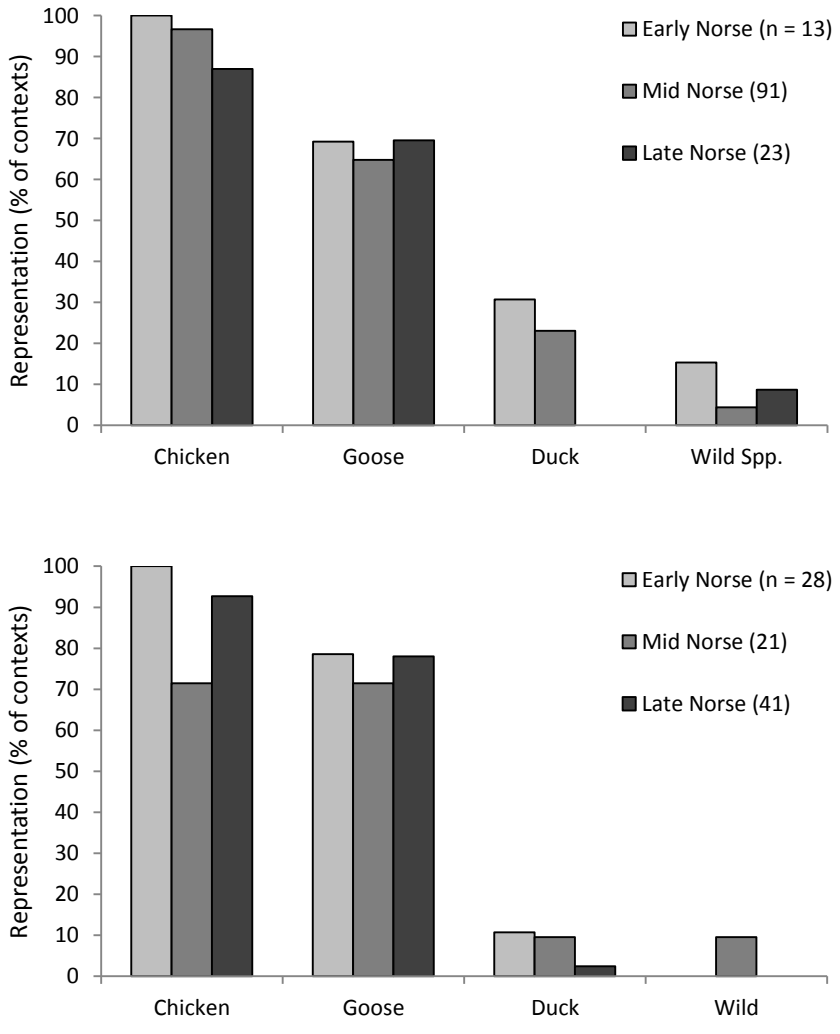

Figure 4.6: upper (Mound 2) and lower (Mound 2A): data presented by phase of occupation, with the number of eggshell-bearing contexts given.

4.2.3.3: Quoygrew

Most of the material from Quoygrew has been dated to between the 11th and 13th centuries A.D., and is drawn from the midden deposits related to the structure near the foreshore (Section 4.2.1.4.2). A small amount of material is drawn from the farm mound (Area G), and from earlier phases of occupation (10th century). Both sections of the site (Areas A-F and Area G) contain material representing both phases of occupation, but in both the main contributions are made by material from the later phase (11th to 13th centuries).

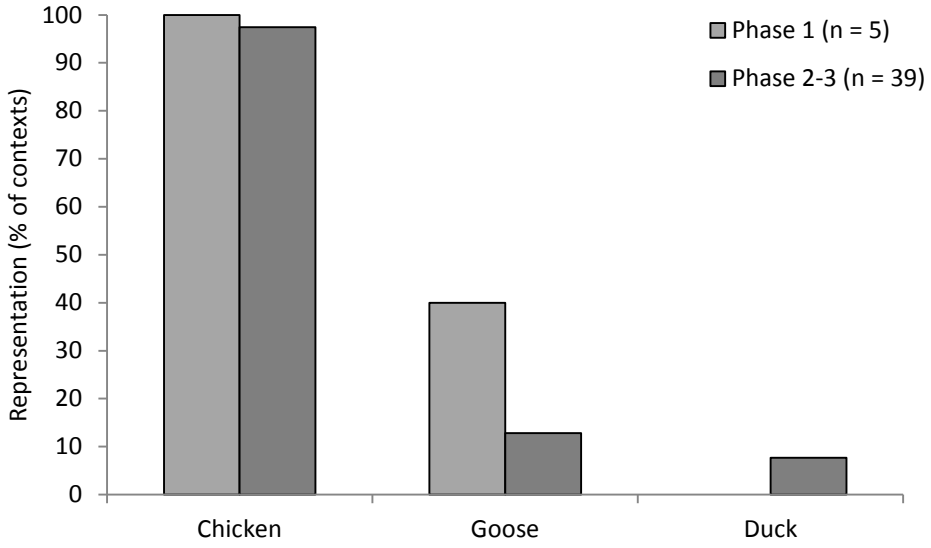

Figure 4.7: Quoygrew data by phase of occupation; note the very low number of eggshell-bearing contexts representing Phase 1 (10th century) relative to Phase 2-3 (11th – 13th century).

When considered by phase of occupation (Figure 4.7), chicken seems equally ubiquitous throughout the span represented by the eggshell assemblage. Goose appears to be far more prevalent during the earlier phase of occupation, but it is

impossible to discount the possibility that this relates (at least in part) to the very low number of contexts representing the earlier phase of occupation (n = 5); this pattern is not observed when the Area G material is considered alone (data not shown).

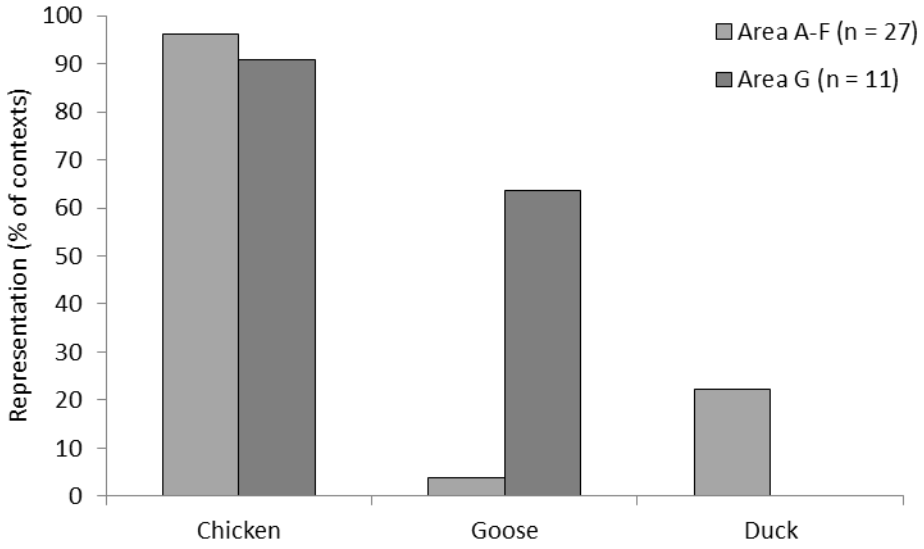

Figure 4.8: Quoygrew data presented by location within site, with number of eggshell-bearing deposits given. Note that all of the data pertaining to the fish middens (Area A-F) are aggregated; considering these individually does not alter the pattern observed.

These data can also be considered in terms of their location within the site (Figure 4.8), although this approach precludes separation by phase due to the very low proportional representation of the earlier phase of occupation (particularly in the 'fish middens' – Areas A-F). Again, chicken eggshell is almost ubiquitous; goose eggshell seems much more prevalent in the deposits relating to the farm mound (Figure 4.8). This is interesting in light of previous results from Anglo-Scandinavian

York and from Bornais and Snusgar (discussed in section 4.2.4.6). Duck only appears in the fish midden deposits (Figure 4.8).

4.2.3.4: Snusgar, Sandwick South, Udal North and Earl's Bu

At Snusgar (10th to 13th century A.D.) only chicken and goose were present (in roughly equal proportion), and no demonstrably wild species were recovered (Figure 4.4). This period is contemporaneous with the phases of occupation of Quoygrew represented in the eggshell assemblage, and the lack of wild species is consistent with the results from that site (Figures 4.7 and 4.8). At Sandwick South, which represents the late Norse period (early 14th century), the same pattern was observed (Figure 4.4). At Udal North, only chicken and goose are observed in the contexts which can be confidently assigned to Norse activity (Figure 4.4). It is interesting to note that a higher proportion of wild species are observed at this site prior to and after the Norse occupation (data not shown), but the complexity of the phasing at Udal North precludes a well-defined temporal separation and any patterns observed may be coincidental. All of the identified material from Earl's Bu was found to represent chicken (Figure 4.4).

4.2.4: Discussion

4.2.4.1: Spatial continuity in egg use at Norse-era sites

The general patterns of egg use highlighted here appear broadly consistent between most of the sites analysed (Figure 4.4). Chicken and goose are the dominant taxa during the Norse period, although the contribution of goose is lower at Quoygrew (Figure 4.4). Wild species are present at a low level or are

absent; reasons for this are considered below (Section 4.2.4.3). These data suggest that patterns of egg use were broadly consistent between different locations in the Scottish islands during the Norse period.

These general patterns are not consistent with contemporaneous Norse activity reported in other coastal and island settings. For example, Jane Sidell used scanning electron microscopy (Sidell, 1993a, b) to analyse the eggshell assemblage from the Norse settlement at Freswick Links, Caithness, and reported a range of cliff-nesting wild species (Graham-Campbell & Batey, 2008). Using the same technique, Sidell also analysed the 10th to 13th century eggshell assemblages from Myvatn in northern Iceland (McGovern et al., 2006). While waterfowl such as ducks and geese (which are also relatively common at Scottish island sites) predominated, it is not clear whether these represent wild or domestic types (McGovern et al., 2006). The eggs of ptarmigan (*Lagopus mutus*) were found to be the next most abundant, and eggs of seabirds were also present; domestic chicken are absent from both the bone and eggshell assemblages (McGovern et al., 2006). These results suggest a much greater emphasis on collection of wild resources than is observed at the island sites in this study; this may relate as much to the nature of the landscape and its suitability for farming and animal husbandry as to major cultural differences between areas (e.g. McGovern et al., 2006), although it must be noted that it is unclear whether the parameters used in identification in these studies are reliable.

4.2.4.2: Temporal continuity in egg use at Norse-era sites

One of the most striking observations made here is the major difference in exploitation patterns between the Late Iron Age inhabitants of Bornais Mound 1 and the Norse settlers (Figure 4.5). Although these are separated by a considerable temporal span, they suggest that Norse egg consumption was based almost entirely on domestic (and possibly domestic) species, with little input from wild species such as seabirds, or at least that consumption of wild and domestic eggs was not conducted in an identical manner. This is in stark contrast to the Iron Age material, in which a range of wild taxa (including swan, alcids, larids and curlew), are present in over half of the eggshell-bearing deposits (Figure 4.5). The presence of chicken in the Iron Age deposits shows that these people had access to domestic species, but also exploited the wild resources; the Norse, in contrast, provide no evidence of having used this resource to any great extent (or at all in the Northern Isles). Within the Norse period, patterns of egg use appear to remain broadly consistent.

4.2.4.3: Use of wild resources

The minimal use of wild species in the Norse-era material is a striking aspect of these data. Wild species (most commonly larids and alcids) make a minimal contribution to the eggshell assemblages recovered from Norse levels at sites in the Western Isles, and are not seen at all in sites from the Northern Isles (Figure 4.4). This lack of wild input becomes particularly interesting when compared with the avian bone assemblages. Previous analyses on Anglo-Scandinavian York showed that a wide range of bird species were represented in the bone assemblages, but this is not reflected in the eggshell (O'Connor, 1989, 2000;

Stewart et al., 2013; Stewart et al., submitted). This might be expected in an urban environment such as York; the low representation (or total lack) of wild species at the sites analysed in this study (Figure 4.4) is perhaps more of a surprise. All of these sites are located near a range of potentially ideal locations for exploitation of cliff and ground-nesting seabirds, and it had been considered likely that their occupants would have used the wild eggs available to them (Graham-Campbell & Batey, 1998; Sharples, 2005, 2012; Harland et al., 2012). The fact that these data suggest that they generally did not do so requires explanation.

The possibility of the lack of wild eggshell relating to a general lack of interaction with the locally available wild resources can be discounted. As mentioned previously, a range of wild birds are found in the bone assemblages from many of these sites (Sharples, 2005, 2012; Griffiths & Harrison, 2011; Harland et al., 2012; J. Best, pers comm). For example, in Mounds 1 and 3 at Bornais, at least 32 species representing 10 orders were identified in the avian bone assemblages (Sharples 2005, 2012). Of these, 24 species are at least to some extent considered seabirds or waterfowl. While a few species may not have actually been present as food waste, these remains are largely considered to derive from consumption (Sharples, 2005, 2012). At Quoygrew, at least 38 species were identified, similarly incorporating a wide range of seabirds and waterfowl (Harland et al., 2012). At Snusgar, there is evidence for systematic exploitation of gannets (Griffiths & Harrison, 2011), but wild birds are not represented in the eggshell assemblage.

Given the reported identification of wild eggshell from Freswick Links (Sidell, 1995; Graham-Campbell & Batey, 1998) and from sites in the Western Isles (Figure 4.4), a widespread cultural aversion to or neglect of wild eggs can also be discounted. It must be noted that different techniques were used in identification; attempts were made to secure the Freswick Links material (as well as material from other mainland sites) in order to test the correspondence between the techniques, but without success. It is possible that the range of wild species at Freswick Links has been overestimated.

One interesting possibility is that the general lack of wild eggshell in the Norse assemblages relates to systems of land ownership and access to wild resources. The Norse 'Udal' system of land ownership was complex (e.g. West, 1975), and made a strict distinction between privately owned and commonly held resources (G. Bigelow, pers comm). In brief, an Udal holder had absolute ownership over his land, which was gained by his family having held it over a number of generations; on his death, the land was divided among his descendants (both male and female), with the eldest son taking the main seat of residence (e.g. West, 1975; Tait, 2012). The rights to most terrestrial resources (which might include the nests of cliff- and ground-nesting birds) were most likely held by the main landowner (i.e. the Udal holder); in the Northern Isles at least, commonly held terrestrial resources were mostly limited to moorland (G. Bigelow, pers comm). In this scenario, it is possible that the sites considered here were not occupied by the main landowner, and therefore did not have legal rights to collection of wild eggs; given the uneven distribution of nesting colonies, the odds of analysing the eggshell assemblage from a site which held access rights to this resource are slim.

Of course, this does not suggest that such resources were never used illicitly or on a small scale, but it might explain the lack of exploitation in areas potentially rich in wild eggs. These laws would not interfere with exploitation of adult birds; these could be caught at sea, which was usually held to be an area where individual landholders had access to resources, or when within the territory of an individual farm even if it was not occupied by the main landowner (G. Bigelow, pers comm). An analogous situation has been observed in the mammal bone assemblage at Sandwick South; despite close proximity to seal-pupping sites, seal bones are all but absent from Sandwick South, while at the nearby high-status site of Underhoull (also on Unst) seal bones are abundant (G. Bigelow, pers comm). The most rational explanation of this disparity is that it arises from systems of legal access to these resources. This explanation can also tentatively reconcile the difference in wild egg use observed between Freswick Links (Section 4.2.4.1) and contemporaneous sites by placing the former as a central place with legal access to cliff-nesting birds (or a hub of illicit egg stealing), although there is little to suggest that this site was a centre of activity (Morris et al., 1995). This may also go some way towards explaining the observed difference in wild egg use between the Late Iron Age and Norse era settlers at Bornais (Figure 4.5).

The paucity of wild eggshell in the assemblages studied here raises a range of other possible explanations. One is that the inhabitants of these sites did not consider the eggs of these wild species as desirable as those of domestic species, and/or worth the collection effort required. Another (purely conjectural) possibility is that Norse seabird exploitation could be analogous to the still on-going regional tradition of targeting young seabirds; people may simply have

waited until the majority of the birds had hatched before collecting them. Relatively recent accounts from Welsh and Irish coastal islands tell of the seabird collectors using the eggs as a quick post-work snack, while taking the animals captured for sale or use on the mainland (Baldwin, 2009). This might lead to disposal patterns which are not conducive to the material being recovered; the evidence from Iceland, however, suggests that there was no specialised off-site disposal of wild eggs, at least in that area (McGovern et al., 2006). Seabirds might also have been captured as a secondary bonus of fishing activities, either by opportunistic capture in nets or by use of baited lines (O' Connor, pers comm). This would obviously bias exploitation of these species towards the adult birds only. The idea of capture as a secondary priority during the Norse period also fits well with the generally low contribution of bird bones to most of the overall bone assemblages at Bornais and Quoygrew (even after factoring in taphonomic effects); this does not appear suggestive of a major subsistence activity (Sharples, 2005, 2012; Harland et al., 2012). Another possibility is that the eggs of wild species were routinely used at these sites, but were disposed of differently and are therefore not found alongside normal food waste. This seems unlikely for two reasons. For one, only cultural factors could underlie such differential waste disposal, and the Norse were not renowned for being particularly fussy in their waste disposal habits (e.g. A. Jones, pers comm). The other reason is that a wide range of different types of deposit were excavated, particularly at Bornais and Quoygrew, and so even if wild eggshell was disposed of differently at these sites it should still stand an equal chance of being recovered.

4.2.4.4: Regional variability in egg use at Norse-era sites

Unfortunately it is not currently possible to compare these Scottish eggshell assemblages with findings from the wider area of Norse influence; there is presently an almost complete absence of data (with the exception of McGovern et al., 2006) from Norse eggshell assemblages in other regions. If future research can target assemblages from Scandinavia, mainland Scotland and Ireland, and other islands such as the Faeroes, Iceland and Greenland, our understanding of Norse egg exploitation will be greatly improved. The evidence provided by bird bones is broadly consistent between many areas in that they tend to exhibit a small contribution of domestic fowl, accompanied by a diverse array of wild taxa (e.g. Sharples, 2005, 2012; Becker & Grupe, 2012; Harland et al., 2012); in some areas, however, only wild taxa are observed (McGovern et al., 1983, 2006). The Scottish island data provided here suggest that there is not a clear relationship between the exploitation of species for meat and for eggs; this urges caution when considering patterns of egg exploitation in light of the species present in the bone assemblage.

4.2.4.5: The importance of the domestic chicken

The other interesting disparity between the eggshell and avian bone assemblages is in the representation of the domestic chicken. These are present at Quoygrew, but are outnumbered in the bone assemblage by a host of wild species including gulls, gannets, cormorants and shags (Harland et al., 2012). At Bornais, they are conspicuous by their inconsistent presence, which although important, does not dominate the avian bone assemblage. For example, no chicken bones were identified from the Iron Age deposits in Mound 1, while only small amounts were

identified from the Norse levels (Sharples, 2005). This is not reflected in the eggshell, in which chicken is the dominant taxon during the Norse period (Figure 4.4), and is present during the Iron Age (Figure 4.5). The difference in the prevalence of domestic chicken in the Iron Age and Norse levels at Bornais (Figure 4.5) suggests that although these were available to the Iron Age inhabitants, they were not as integral to the domestic economy as they were during the Norse period.

The disparity between the proportion of the avian bones representing chicken and the high representation in the eggshell assemblages suggests that egg production may have been the primary focus of Norse-era chicken husbandry, and may have been carried out by a relatively small number of hens, while wild-fowling provided the majority of avian meat. For example, the productivity of the breed(s) available at Bornais can be reasonably conjectured to have been around 200 eggs each per year; monastic records suggest that contemporaneous chickens in southern England produced 200+ eggs per year (Stone, 2006), and this is probably slightly higher than in Scottish islands due to differing light regimes. If 50-80 people (N. Sharples, pers comm) lived at the site, 25-40 laying chickens could have provided 100 eggs per person per year.

The relatively low representation of the chicken in many Norse bone assemblages may relate to their use in egg production being more important than in meat production; this is consistent with Norse avian assemblages elsewhere in Europe (Becker & Grupe, 2012). Although the ratio of males to females and the age

distribution of chickens at Norse sites is not clear, the prevalence of eggshell at the sites in this study, combined with the low representation of adult birds in the bone assemblages both in Scotland and abroad, suggest that egg production and not cockfighting or some other purpose was the main role played by chickens in Norse society. The widespread applicability of this hypothesis can only be tested by analysis of eggshell assemblages from Norse sites across a wider geographical area. It is interesting to note that in regions where domestic fowl were absent (e.g. Iceland), exploitation of wild eggs seems to have been more extensive than exploitation of the adult birds (McGovern et al., 2006). This suggests that eggs were considered a valuable and worthwhile resource by the Norse in these regions, even in the absence of a mass-producing domestic species.

4.2.4.6: Can eggshell identify status?

Another interesting speculation can be made regarding the potential application of eggshell in the archaeology of this period. Previous analyses on assemblages from York have shown a major discrepancy in the prevalence of goose eggshell between relatively low-status (Hungate – 15%) and relatively high status (Coppergate – 40%) sites of the same period (Section 4.1.4.5) (Stewart et al., in press). The results obtained in this study offer tantalising suggestions of a broader pattern; at Quoygrew, for example, goose eggshell is more prevalent in the deposits related to the farm house than in those related to the more modest dwellings on the foreshore, and the ratio between these is very similar to that observed in Anglo-Scandinavian York (Figure 4.8). Mounds 2 and 2A at Bornais are known to represent relatively high-status activity (N. Sharples, pers comm), and provide a very high representation of goose eggshell; the same can be said of

Snusgar, which may well represent another central, high-status site (Griffiths & Harrison, 2011). This allows the construction of a cautious hypothesis that the prevalence of goose eggshell may be crudely related to the relative status of a site (or an area within a site). In order to test this, more material is needed from relatively low-status sites; although this was available from Mound 3 at Bornais, there were not enough contexts represented (n = 6) to conduct a meaningful comparison. This possibility must therefore remain speculative for now. The idea of goose eggshell being considered higher status fare is, as previously stated, consistent with literary sources from Norse-influenced Ireland which are contemporary with the mid-late Norse period (12th century) (Ó Corráin, 2001; Mac Con Iomaire & Cully, 2007).

4.2.5: Conclusions on this chapter

Norse era Scotland has provided the setting for the first large, multi-site investigation of archaeological eggshell. This has raised a number of intriguing results and has begun to fill in a gap in our understanding of the subsistence strategies used during this period. Future research should widen the net to include coastal mainland sites and to increase the temporal span of the study. Following establishment of the new technique showcased in this chapter, it should become easier for researchers to construct wide-scale regional analyses of egg use in different areas around the world.

Chapter 5: Diagenesis of eggshell proteins and amino acids

Section 1: Background: amino acid & protein structure

5.1.1: Overview

The previous chapter described the application of a new technique for identification of archaeological eggshell based on proteins. This chapter will further explore the archaeological potential of these proteins and their constituent amino acids. Diagenesis of biomineral amino acids has also been successfully applied to archaeological and geochronological research (e.g. Penkman et al., 2008, 2011). Informed use of these in archaeological research requires an understanding of the manner in which they degrade over time. This chapter aims to explore and quantify the diagenetic processes operating on eggshell proteins and amino acids. The first section introduces the basic structure of amino acids and proteins; the second details the main diagenetic reactions experienced by these molecules in preserved biominerals; the third section attempts to describe these processes in eggshell by quantifying some of these diagenetic pathways, and also begins to compare observed patterns with those affecting archaeological samples; while the fourth section presents an exploration of some of the factors affecting diagenesis, and begins to test the correspondence between peptide recovery at high and low temperature.

5.1.2: Basic structure of amino acids

All amino acids share a common general structure (Figure 5.1): the α-carbon atom is attached to a carboxyl (COOH) group, an amino (NH_2) group, a hydrogen atom and a side chain (the R group) which is unique to each amino acid.

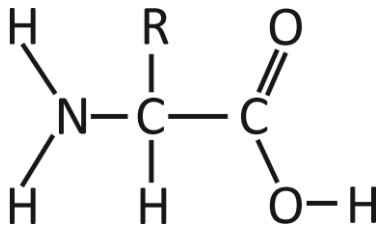

Figure 5.1: Basic amino acid structure. Aside from the R-group, this structure is common to all amino acids.

Although amino acids can potentially occur as a huge variety of forms, only a small subset (20) of these comprise the constituent materials of all proteins found in nature (Figure 5.2) (e.g. Bada, 1991). These are generally divided into four major groups based on their chemical characteristics: hydrophobic (non-polar), hydrophilic (polar), positively charged (basic) and negatively charged (acidic) (Figure 5.2).

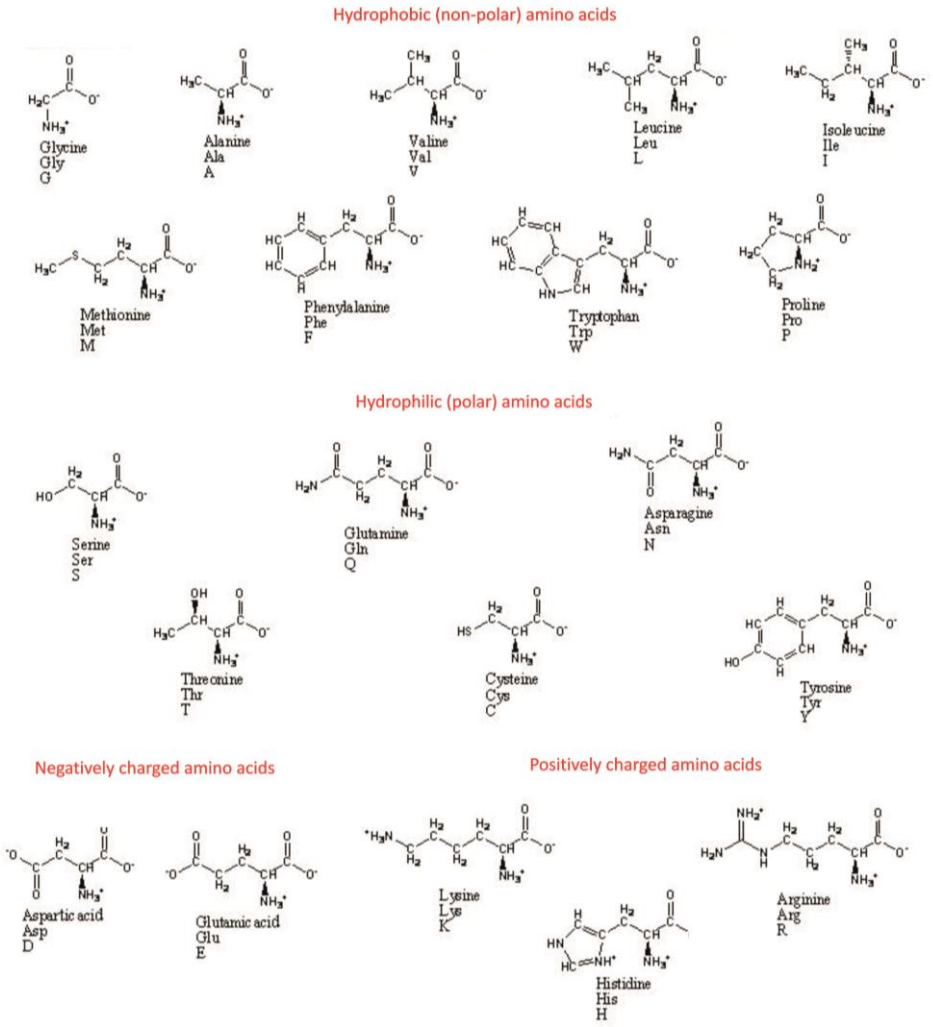

Figure 5.2: All 20 amino acids occurring in proteins, with commonly used abbreviations.

5.1.3: Chirality

5.1.3.1: Amino acid chirality

The α-carbon is chiral in all amino acids except glycine (the R group of glycine is a single hydrogen atom). These therefore have two possible configurations (enantiomers) which, although structurally identical, are non-superimposable mirror images of each other (Figure 5.3). These are commonly referred to as D

(Dextro) and L (Laevo) enantiomers (in reference to the direction in which they scatter polarised light). Some amino acids (isoleucine and threonine) are diastereomeric (i.e. they have two chiral centres).

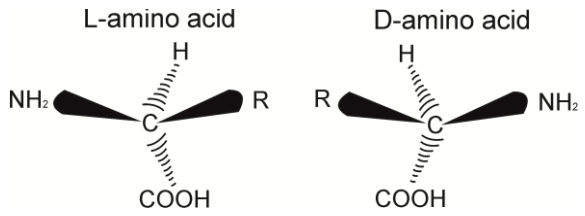

Figure 5.3: Chirality in a generic amino acid. The molecules are identical except for their configuration in space, in which they are mirror images of each other.

5.1.3.2: Homochirality in living organisms

With the exception of a few highly specialised proteins (e.g. Fujii & Saito, 2004), chiral amino acids occur almost exclusively as the L-enantiomer in living organisms. Although D-amino acids do also occur in a few specialised proteins in complex organisms, they have not been observed in any hard tissue (i.e. material which is likely to be archaeologically preserved) (Fujii & Saito, 2004). This discrimination occurs despite both configurations being equally favoured in thermodynamic terms. For example, the discrimination against D-tyrosine configuration during peptide synthesis has been estimated at around one part per 10^4 (Yamane et al., 1981). Disequilibrium is maintained by enzymes which stereoselectively use only the L-enantiomer. This discrimination is necessary; protein function is shape-dependent, and the molecular interactions which dictate conformation at subsequent levels of protein structure (Section 5.1.4) would be hampered if constituent amino acids were in a racemic mixture. There is no

obvious reason why L-configuration is more biologically viable than D-configuration and it is not clear why L-configuration came to dominate the living world; it has been argued that the all but total dominance of L-amino acids in nature is consistent with a single origin of life on Earth (e.g. Bada, 1991).

5.1.4: Peptide synthesis & protein structure

The amino acid sequence is the primary level of protein structure. At non-extreme pH (3.2 – 9.4) the amino group carries a negative charge, while the carboxyl group carries a positive charge. Peptide bonds are formed by condensation reactions between the amino and carboxyl groups of adjacent residues (Figure 5.4).

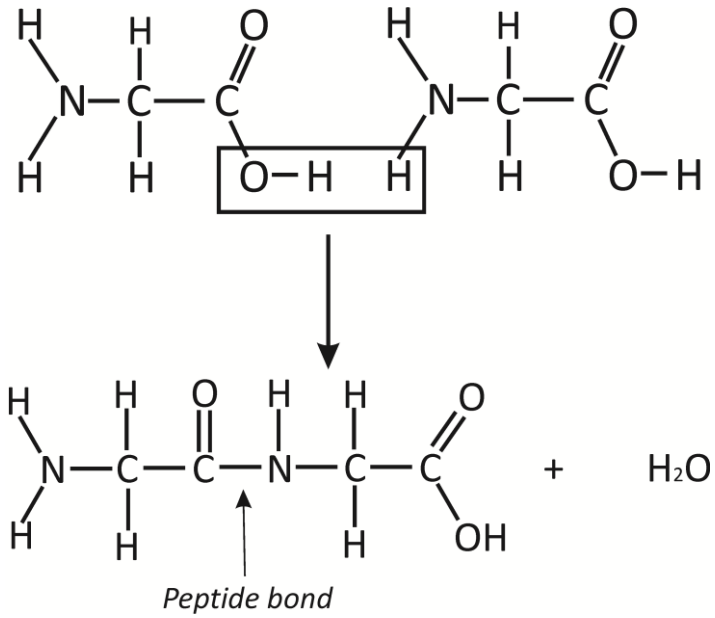

Figure 5.4: Condensation reaction between two glycine residues to form a peptide bond.

The secondary level of protein structure is the folding of peptide chains into 3-dimensional configurations (α-helix or β-sheet). This is governed mainly by hydrogen bonding between amide groups in the peptide chain; for example, large amino acids with aromatic side chains preferentially form β-sheet configuration. The tertiary level is the conformation of secondary structures into the final protein shape; this is governed by interactions between amino acids (including hydrogen bonding, salt bridges, disulphide bonds and non-polar hydrophobic reactions). Quaternary protein structure may involve the combination of individual monomers into a functional unit, and can also involve the attachment of non-protein molecules such as carbohydrates.

Section 2: Protein and amino acid diagenesis

5.2.1: Overview

The previous section outlined the basic structure of amino acids and the formation of proteins. This section will describe the main pathways by which these molecules degrade in biominerals; understanding these is central to archaeological and geochronological applications of biomolecular diagenesis. Three major diagenetic reactions are described: hydrolysis, amino acid racemisation, and decomposition. Specific emphasis will be placed on hydrolysis and racemisation, as these are directly measurable using the techniques available to this study. The main diagenetic reactions summarised in this section are briefly described below (Table 5.1).

Process	Mechanism	Product(s)
Hydrolysis	Scission of peptide bond by addition of water	Smaller peptides and free amino acids
Racemisation	Inter-conversion of D and L amino acids	Racemic mixture of D and L enantiomers
Decomposition	Various	Organic compounds including other amino acids
Succinimide formation	Deamidation of in-chain Asx	Cyclic dipeptide prone to rapid racemisation
Diketopiperazine formation	Aminolysis at N-terminal of in-chain residue	Cyclic dipeptide prone to rapid racemisation

Table 5.1: Summary of diagenetic mechanisms described in this section (Wilson & Cannan, 1937; Vallentyne, 1964; Hill, 1965; Steinberg & Bada, 1981, 1983; Geiger & Clarke, 1987; Walton, 1998)

5.2.2: Peptide bond hydrolysis

Hydrolysis is a reaction in which the addition of water causes fragmentation of molecules into smaller components. In the most commonly occurring hydrolytic

reaction in biomineral proteins, the peptide bond (Figure 5.4) is broken by the addition of water. The rate at which this occurs is affected by water availability, temperature and the hydrophobicity of amino acids adjacent to the bond (Bada, 1991; Walton, 1998; Collins & Riley, 2000). Over time, and assuming that water does not become limiting, protein chains will be progressively broken by hydrolysis into smaller peptides and ultimately into free amino acids (e.g. Hill, 1965; Bada, 1991). In most surface environments, proteins within biogenic carbonates are expected to be hydrolysed to smaller peptides and free amino acids within 10000 years (Bada et al., 1999).

The extent of hydrolysis can be quantified by measurement of the proportion of total hydrolysable amino acids (THAA) present as free amino acids (FAA). However, this provides only an approximation of the actual rates of hydrolysis occurring within the system, as the observed rate is a product of a vast number of underlying rate constants (Hill, 1965). There are 20 amino acids found in proteins. All of these differ in their affinity to water and chemical characteristics, and any pair can occur adjacent to each other; there are up to 400 separate rate constants underlying the observed rate of hydrolysis, assuming no effects caused by higher-order structures (Hill, 1965). Although measuring reaction parameters for individual amino acids improves the resolution of this approach to some extent, untangling this composite requires detailed knowledge of the molecular environment and a full understanding of the diagenetic process (e.g. Demarchi et al., 2013). As this is lacking for most eggshell proteins, and is beyond the capabilities of the analytical techniques used in this project, reaction parameters

observed for hydrolysis (Section 5.3) can only be considered accurate for the particular molecular environment from which they are derived.

5.2.3: Racemisation

5.2.3.1: Amino acid racemisation

The disequilibrium between D- and L-enantiomers (Section 5.1.3) is thermodynamically unfavoured, and is maintained only by stereoselective discrimination and continuous amino acid turnover (e.g. Johnson & Miller, 1997). When physiological interaction between the substrate and the organism ceases, the enantiomeric balance begins to shift towards equilibrium. This occurs after death in materials which are constantly renewed (e.g. bone), and immediately following deposition in materials which have no subsequent interaction with the producing organism (e.g. invertebrate shell and avian eggshell). This process is called amino acid racemisation (AAR), and is measured as the ratio of D- to L-enantiomers (D/L): in living tissues, D/L ≈ 0 in most amino acids. In diastereomeric amino acids, the process of reconfiguration around the α-carbon is often referred to as epimerisation (e.g. Johnson & Miller 1997). Racemisation also occurs around the β-carbon in diastereomeric amino acids, but this process is thermodynamically un-favoured relative to α-carbon epimerisation and its products are not recovered using the analytical techniques employed by this study (e.g. Bada et al., 1986). Isoleucine epimerisation, which leads to formation of the non-protein amino acid D-alloisoleucine, is referred to as racemisation hereafter.

5.2.3.2: Mechanism of racemisation

Racemisation is a spontaneous inter-conversion reaction driven by abstraction of the α-hydrogen, which leads to the formation of a planar carbanion (Figure 5.5) (Neuberger, 1948; Bada, 1991). A proton can then be re-added to either side of this carbanion, thus generating either the L- or the D-enantiomer. Over time, the D/L ratio will approach a dynamic equilibrium, where inter-conversion between enantiomers continues without altering the observed D/L ratio (e.g. Clarke & Murray-Wallace, 2006). In most amino acids, racemisation can only occur when a residue occupies a terminal position or a position in a cyclic dipeptide (Section 5.2.5.2.2); the rate at which hydrolysis produces new terminal positions is a major factor determining racemisation rate.

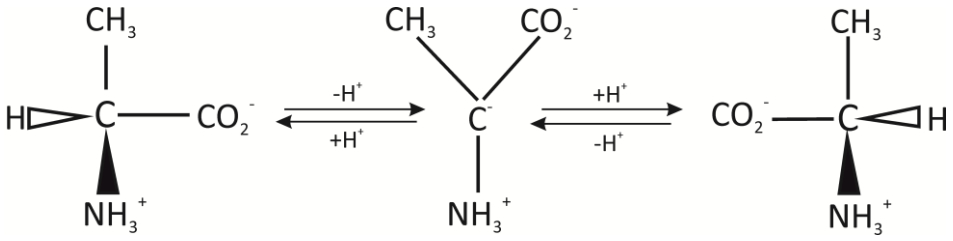

Figure 5.5: Mechanism of racemisation via a carbanion intermediate in free amino acids (after Demarchi, 2009)

5.2.3.3: Racemisation rates in free and peptide-bound amino acids

At low temperatures (0-25°C), the racemisation rate of free amino acids is pH-independent within a non-extreme pH range (3-9) (Bada & Shou, 1980; Bada, 1985). The fastest racemisation rates are observed in amino acids whose R groups have the greatest electron withdrawing capacity (e.g. serine & threonine), as the

intermediate carbanion is relatively stable in these (Bada & Shou, 1980; Smith & Evans, 1980). Conversely, the slowest racemisation rates are observed in amino acids with electron-donating side chains (e.g. valine & isoleucine) (Bada & Shou, 1980; Smith & Evans, 1980; Bada, 1985). The structure and size of the R group can also affect racemisation rates, which are generally slower in amino acids with large side chains as these can affect the rate of proton abstraction (Smith & Evans, 1980). These factors are considered the major determinants of racemisation rates, although they cannot explain all variation between different free amino acids in solution (Bada & Schroeder, 1975; Smith et al., 1978).

Use of free amino acids (FAA) in archaeology and geochronology has been relatively limited (but see Penkman et al., 2008; Demarchi et al., 2013). In many biominerals, free amino acids can be lost by diffusion to the external environment; this will skew the observed racemisation rate unless care is taken to isolate an effectively closed system of amino acids (Section 2.1) (e.g. Kimber & Griffin, 1987; Collins & Riley, 2000; Penkman et al., 2008; Demarchi et al., 2011, 2013). Also, kinetic patterns observed for FAA racemisation are ultimately dependent on the rate of release of racemised residues from peptides, which can be affected by a wide range of factors (see below) (Kriausakul & Mitterer, 1980; Wehmiller, 1980). This has necessitated focus on the bound fraction of amino acids, in which observed racemisation kinetics are usually more complex. Very little racemisation occurs in amino acids which are bound within intact peptide chains; most residues will only racemise at terminal positions (Kriausakul & Mitterer, 1978; Smith & Evans, 1980; Mitterer & Kriausakul, 1984; Kimber et al., 1987; Kaufman & Sejrup, 1995; Kaufman & Miller, 1995). Racemisation is faster in

N-terminal residues than C-terminal residues, and both of these are two orders of magnitude faster than the observed rates in peptide bound and free amino acids (Kriasaukul & Mitterer, 1978, 1980). A range of factors relating to the local molecular environment can affect racemisation rate. These include the position of an amino acid within a protein or peptide; the chemical and physical characteristics of the amino acid; the characteristics of adjacent residues; and the relative lability of the peptide bond between a residue and its neighbours (e.g. Kriausakul & Mitterer, 1980; Gaines & Bada, 1988). As racemisation rate is normally governed by hydrolysis rate, parameters derived for racemisation are also highly dependent upon the local molecular environment.

5.2.4: Amino acid decomposition

The other major diagenetic process experienced by biomineral amino acids is decomposition. Although this is not a major focus of this project, it is important to be aware of the potential effects of amino acid decomposition on observed reaction rates. While most amino acids will only decompose when free, those with hydroxyl aliphatic side chains (serine and threonine) can decompose while peptide-bound (Wehmiller, 1980; Walton, 1998). Decomposition products of amino acids include a wide range of organic compounds, in some cases including other amino acids (Abelson, 1954; Bada, 1971, 1991; Bada & Man, 1980; Bada et al., 1978; Brinton & Bada, 1995; Goodfriend, 1991; Hare & Mitterer, 1967; Hill, 1965; Murray et al., 1965; Vallentyne, 1964; Wilson & Cannan, 1937; Walton, 1998). Although patterns of amino acid degradation and loss have the potential to confound archaeological interpretation, attempts to directly model rates of decomposition have been generally lacking as they are difficult to measure

directly (but see Demarchi et al., 2013). In cases where the decomposition product of one amino acid is another amino acid, the potential effect on interpretation can be major. For example, serine decomposes to alanine and/or glycine, while glycine can also appear as a decomposition product of threonine (Bada et al., 1978; Bada & Man, 1980; Walton, 1998). This can limit the applicability of some amino acids to research on preserved biominerals, as it can be impossible to differentiate between residues resulting from decomposition and those in their original state.

5.2.5: The kinetics of diagenesis

5.2.5.1: The 'three-box' model

Modelling the often complex kinetics of diagenesis in a range of biominerals has been the focus of significant effort in recent years (e.g. Collins & Riley, 2000; Miller et al., 2000; Penkman et al., 2008; Crisp et al., 2013; Demarchi et al., 2011, 2013). Due to its potential applicability as a dating tool and palaeothermometer (e.g. Brooks et al., 1990; Miller et al., 2000; Penkman et al., 2011), much of this research has focussed on the racemisation reaction; racemisation rate is not constant between the starting point and measurable end point of the reaction (i.e. dynamic equilibrium). The progressive change in racemisation rate during diagenesis has been incorporated into the 'three-box' model of the kinetics of diagenesis, which also factors in the effects of hydrolysis and decomposition (Kriausakul & Mitterer, 1980; Wehmiller, 1980; Collins & Riley, 2000). This model posits that the interplay between the three major diagenetic reactions, combined with diffusive loss of FAA and small peptides from open systems, is responsible for

the observed changes in rate during the course of the reaction (summarised below).

During very early diagenesis, a high proportion of amino acids are peptide-bound and there are few terminal positions; racemisation should be relatively slow. As hydrolysis leads to production of smaller peptides, and therefore to more terminal positions, racemisation rate should then increase; during late diagenesis, the observed racemisation rate decreases (Kriausakul & Mitterer, 1980; Wehmiller, 1980; Collins & Riley, 2000). This decrease was considered to result from relatively highly racemised free amino acids and small peptides becoming predominant and being lost by diffusion to the external environment (Kriausakul & Mitterer, 1980; Wehmiller, 1980; Collins & Riley, 2000). This is an insufficient explanation, as the reduction in racemisation rate during late diagenesis has also been observed in closed-system amino acids (Section 5.3.3.4.2), suggesting that an alternative mechanism(s) must contribute to the observed rate dampening (e.g. Demarchi et al., 2013). The apparent reduction in racemisation rate during late diagenesis may relate to slow hydrolysis of a residual hydrophobic peptide-bound fraction, and/or to increased loss of racemised FAA by decomposition (e.g. Collins & Riley, 2000).

5.2.5.2.1: Complicating factors: other influences on diagenesis

The general pattern described by the three-box model, as well as most of the mathematical models used to describe diagenetic reactions (reviewed in Clarke & Murray-Wallace, 2006), assume that diagenesis consists only in the three

reactions described above. However, there are other diagenetic mechanisms in operation which can affect observed reaction rates, and these can be difficult to observe directly. Although measurement of these is beyond the scope of this project, it is important to be aware of the effects they can exert on observed reaction rates; two other potentially major diagenetic pathways are summarised below.

5.2.5.2.2: Diketopiperazine formation

As proteins are progressively broken down into smaller peptides and then FAA, dipeptides which cyclise by aminolysis at the N-terminus to form diketopiperazines (DKP) can be produced (Steinberg & Bada, 1981, 1983). This occurs most readily in small peptides with a high proportion of hydrophobic amino acids, and when glycine occupies the third position from the N-terminus (Steinberg & Bada, 1981, 1983; Sepetov et al., 1991). Dipeptide stability is determined by the order and chemical characteristics of constituent amino acids, and is a major factor in determining the rate of DKP formation (Gaines & Bada, 1988). The prevalence of this reaction in dipeptides, tripeptides and hexapeptides at neutral pH (6-8) suggests that it could be a major diagenetic pathway experienced by proteins in preserved biomaterials (Purdie & Benoiton, 1973; Steinberg & Bada, 1983). DKP formation can have strong effects on patterns of amino acid racemisation; constituents of DKPs racemise faster at neutral pH than terminal residues (Steinberg & Bada, 1981, 1983; Gaines & Bada, 1988). Following hydrolysis of the most labile peptide bonds, DKP formation may become one of the most prevalent diagenetic mechanisms. The formation of DKPs from dipeptides is reversible, but the original order of the residues in the dipeptide may

become inverted (e.g. Steinberg & Bada, 1981; Gaines & Bada, 1988). The order of the residues in the original dipeptide can be important in determining rates of DKP formation and subsequent racemisation. If two possible configurations exist (e.g. Gly-Leu & Leu-Gly), one configuration may cyclise more readily. This will lead to a concomitant increase in racemisation as more time is spent in the DKP state before hydrolysis produces either the original or the inverted dipeptide, and ultimately FAA (Gaines & Bada, 1988).

5.2.5.2.3: Asx diagenesis: succinimide formation

Asparagine is rapidly deamidated to aspartic acid during acid demineralisation (Hill, 1965). Following this, the relative contribution of each to the observed D/L ratio for combined asparagine and aspartic acid (Asx) cannot be quantified (Brinton & Bada, 1985). Asx racemisation displays a fast and a slow component; D/L ratios increase rapidly, but plateau before reaching equilibrium. The overall racemisation rate observed in Asx does not have a linear relationship with racemisation rates of other amino acids under identical conditions (Collins et al., 1999). This may be due to faster racemisation of hydrophilic asparagine than aspartic acid; the former will cause rapid initial racemisation, while the dominance of the latter in later stages will lead to an apparent deceleration (Goodfriend, 1991). Asx can also be rapidly removed from peptide chains as deamidation leads to the formation of cyclic succinimides (Asu) (Geiger & Clarke, 1987). This can lead to rapid loss of these residues by diffusion during early diagenesis in open systems (Bada, 1991). This deamidation reaction is relatively rapid if the Asx residue is adjacent to an amino acid with low steric hindrance, but in some cases Asu formation may be limited by structural constraints in intact polypeptides (Geiger & Clarke, 1987; van Duin & Collins, 1998; Collins et al., 1999). Fast initial rates of

Asx racemisation are often observed during early diagenesis; Asu can racemise far faster than peptide-bound or free residues before being transformed to D-Asx (Geiger & Clarke, 1987; Smith & Baum, 1987; Stephenson & Clarke, 1989; Radkievicz et al., 1996; Collins et al., 1999).

Section 3: Quantifying diagenesis in eggshell proteins

5.3.1: Overview: isothermal heating and the Arrhenius equation

All of the diagenetic reactions described above are far too slow to be monitored in real time at the temperature ranges experienced by preserved biominerals. Isothermal heating experiments at elevated temperature can be used to induce artificially rapid diagenesis. Reaction parameters are typically calculated at a range of temperatures (80, 110 and 140°C in this study) and then extrapolated back to the deposition temperatures experienced by preserved samples. This allows reactions which may take hundreds of millennia to proceed to equilibrium under geochemical conditions to be monitored in a far shorter period of time (e.g. Bada & Shou, 1980; Kriasuakul & Mitterer, 1980). One major assumption of this approach is that the relationship between temperature and reaction kinetics is constant across a wide temperature range. For racemisation, another assumption is that the inter-relationship between racemisation rate and all of the diagenetic reactions which affect it remains constant (i.e. that their activation energies are closely matched). Although there is experimental support for this in some biomineral systems (e.g. Miller et al., 2000), it cannot be assumed that this is universally the case (Demarchi et al., 2013; Tomiak et al., 2013). Derivation of Arrhenius parameters for different diagenetic reactions allows these assumptions to be tested in non-ratite eggshell.

5.3.2: Methods

5.3.2.1: Sample preparation and analysis

Free-range chicken, duck and turkey eggshell was sourced from commercial retailers. Fragments of shell were cleaned by sonication, the outer shell membrane removed, and crushed to a powder (grain size ≤ 500µm). Samples of this powder were weighed into sterile 2 ml Eppendorf tubes and exposed to bleach (NaOCl 12% w/v) at a concentration of 50µl/mg for 48 hours in order to isolate intra-crystalline proteins (Section 2.1). Samples were then rinsed using ultra-pure water, briefly suspended in HPLC-grade methanol, and left to dry. Subsamples were then weighed into sterile glass ampoules, and the powders suspended in ultra-pure water, heat-sealed, and heated at one of three temperatures (80, 110 and 140°C) for a pre-defined length of time (Table 5.2). Unheated controls were also prepared. At the end of the heating period, the water was removed and stored, the sample air-dried, and subsamples were weighed into sterile hydrolysis vials for analysis of both the free (FAA) and total hydrolysable (THAA) amino acid fractions; both were prepared in triplicate.

Temperature (°C)	Heating time points (hours)
80	0, 24, 96, 240, 720, 840, 960, 1200. 1440, 2160, 3600, 5760
110	0, 24, 120, 240, 360, 504, 720, 840, 960, 1200, *1500, 1800*
140	0, 1, 2, 4, 6, 8, 24, 48, 96, 120, 240

Table 5.2: Heating times used in kinetic experiments. Time points given in italics were not used in all experiments.

FAA samples were demineralised in 20µl/mg 2M hydrochloric acid (HCl), before being dried under vacuum in a centrifugal evaporator. THAA samples were

demineralised in 20µl/mg 7M HCl, flushed with nitrogen, and heated at 110°C for 24 hours (Kaufman & Manley, 1998), before being dried in a centrifugal evaporator. Samples were then rehydrated using a solution containing the non-protein amino acid L-homo-arginine (LhArg) (0.01M HCl and 1.5mM sodium azide, with 0.01 mM LHarg). For the FAA analyses, a volume of 10-20µl/mg sample was used; for the THAA analyses, a volume of 60µl/mg was used. After rehydration the samples were vortexed, and an aliquot of each (13.5µl) was transferred to a sterile 2ml HPLC autosampler vial for analysis. LHarg blanks were also included at this stage. Acid blanks were rehydrated with an average volume of the LHarg solution. Analysis was conducted using the reverse-phase HPLC procedure described in a previous chapter (Section 2.1). Enantiomers were identified using a Microsoft VB application designed by R. Allen (Peakachoose v2.6).

5.3.2.2: Mathematical approaches to modelling diagenesis: first-order kinetics

Diagenesis has been modelled using a wide range of mathematical approaches, all of which have relative strengths and weaknesses; these have been comprehensively reviewed (Clarke & Murray-Wallace, 2006). This subsection will describe the approaches used to model diagenesis used in this study.

As it has been shown that racemisation of free amino acids in solution generally conforms to a first-order model (e.g. Bada & Shou, 1980), these have often been used to model diagenetic reactions (e.g. Bada & Schroeder, 1972; Miller et al., 1992; Clarke & Murray-Wallace, 2006). A condition of the accurate application of these models is that the log-transformed data (see below) at each temperature conform to a linear (or pseudo-linear: $R^2 > 0.95$) model in relation to heating time.

An irreversible first order kinetic (IFOK) model can be used to log-transform hydrolysis data (equation 1):

Equation 1: $\ln([Bound]/[Total]) = -kt,$

where *([Bound]/[Total])* is the proportion of the total pool of a given amino acid which is peptide-bound at a given time point, *t* is time (s) and *k* is the rate constant (s^{-1}) at a given temperature (Miller et al., 1992).

For racemisation, a reversible first order kinetic (RFOK) transformation can be used (equation 2):

Equation 2: $\ln[(1 + D/L)/(1 - KD/L) = (1+K)k_1t + c,$

where $K = k_1/k_2$ (i.e. the forward rate constant/the reverse rate constant – these are equal for most amino acids so $K = 1$, but in isoleucine $K = 0.77$), and c is a constant used to offset racemisation induced during sample preparation (Bada & Schroeder, 1972).

The transformed data can then be modelled using the natural logarithm (equation 4) of the Arrhenius equation (equation 3):

Equation 3: $k = Ae^{-Ea/(RT)}$

Equation 4: $\ln(k) = (-E_a/R) \times (1/T) + \ln(A)$

where *k* = rate constant, *A* = frequency factor (s^{-1}), *Ea* = activation energy (kJ/mol), *R* = gas constant (8.8314621), and *T* = temperature (K) (e.g. Brooks et al., 1990;

Goodfriend & Meyer, 1991; Miller et al., 1992, 2000; Clarke & Murray-Wallace, 2006). Calculation of the rate constants at different temperatures allows construction of an Arrhenius plot (ln(k) vs. 1/T) (Figure 5.6). This gives a straight line (equation 5):

Equation 5: y = mx + c,

where m = $-E_a/R$ and c = Ln(A). The expected reaction rate at a given deposition temperature can then be extrapolated (Figure 5.6).

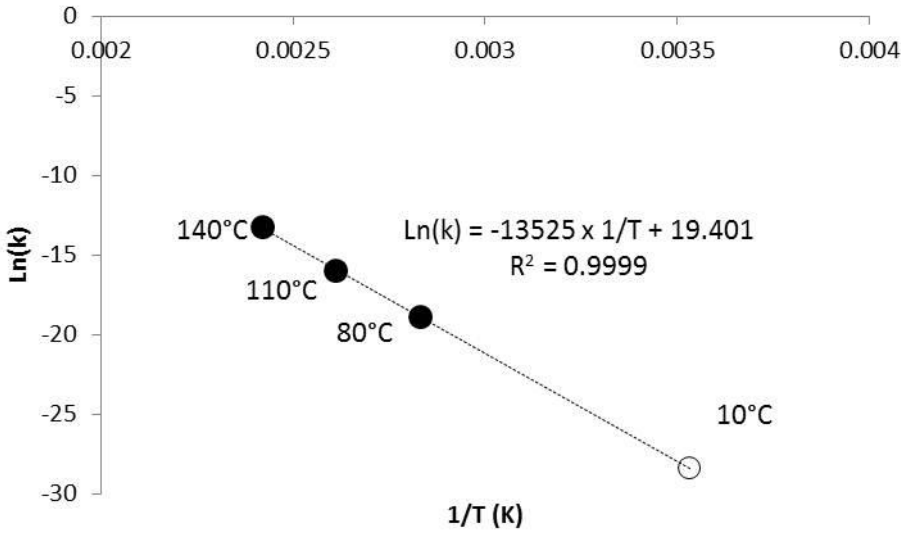

Figure 5.6: Arrhenius plot showing extrapolation of observed racemisation rates derived using a reversible first-order kinetic model (RFOK) to lower deposition temperatures. This example shows rates for isoleucine epimerisation in chicken eggshell. Closed circles represent experimental data, while the open circle represents the extrapolated value for the given temperature. The formula and R^2 value apply only to the experimental data; the accuracy of this projection is tested in a later section (Section 5.3.3.6).

5.3.2.3: Potential problems with first-order models

There are potential problems inherent in first order models. For one, in many cases data do not linearise well using the log-transformation required for FOK models. Another potential problem is the failure of FOK models to accommodate the reduction in observed racemisation rate during late diagenesis (Section 5.2.5.1); this tends to produce reaction parameters which are not applicable to late diagenesis. For racemisation of bound amino acids, as described above, the observed rate is strongly affected by other diagenetic processes; it is not a product of a simple first-order reaction (e.g. Kriausakul & Mitterer, 1980; Wehmiller, 1980; Collins & Riley, 2000). These potential problems necessitate the adoption of a second approach to deriving Arrhenius parameters for diagenetic processes.

5.3.2.4: Mathematical approaches to modelling diagenesis: 3rd-order scaling

A 3rd-order polynomial scaling method is also used to estimate the relative rate of diagenetic reactions at different temperatures. The major advantage of this approach is that it does not require linearization or transformation of the raw data, and therefore does not bias interpretation towards end points to the same extent as FOK models (Demarchi et al., 2013). Instead, the time axis is log-transformed, and a scaling factor applied to the 3rd order polynomial across the range over which the datasets for the different temperatures overlie (Figure 5.7) (Demarchi et al., 2013). A relative rate constant of 1 is assumed for the 110°C data, and the 80° and 140°C data are overlaid; the scaling factor can therefore also be defined as the relative rate (Demarchi et al., 2013). The scaling factor is derived using a Generalised Reduced Gradient Algorithm (Microsoft Solver) which

minimises the least-squares difference between the two polynomial functions (representing the 80 and 140°C data) over the selected range; another advantage of this approach is that it allows greater control of the ranges over which parameters are calculated (Demarchi et al., 2013). Arrhenius parameters can then be derived using an Arrhenius plot as described above (Figure 5.6).

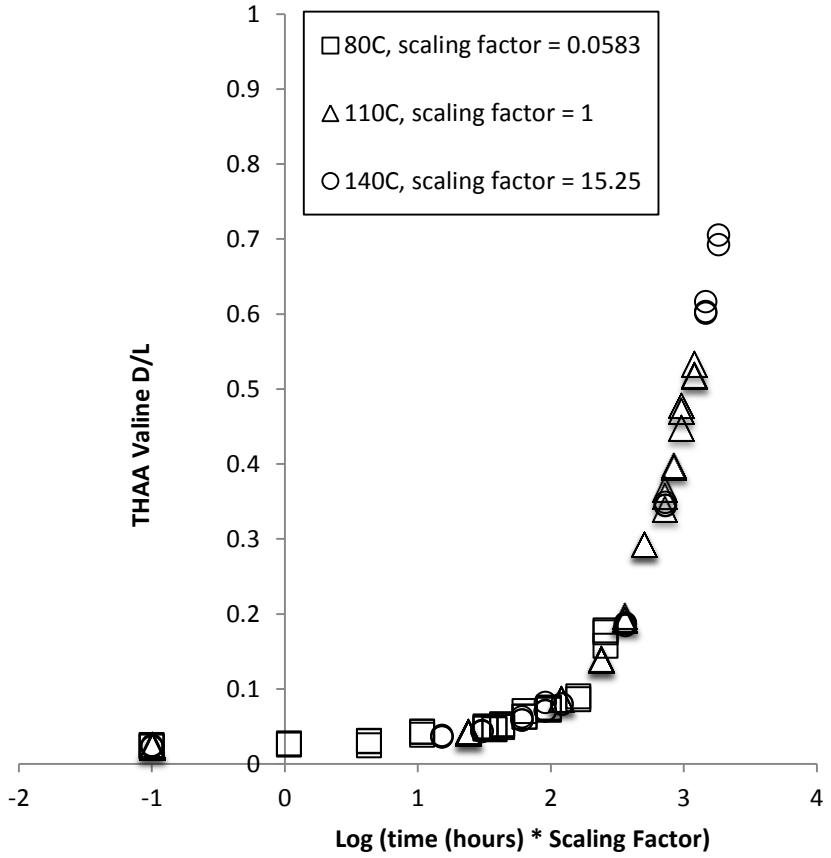

Figure 5.7: Plot of valine racemisation rate in heated chicken eggshell vs. log_{10} scaled heating time. The scaling factor for each temperature is derived by minimising the least-squares difference between the polynomial functions.

As the data at different temperatures accurately represent different stages of diagenesis (the 80°C data capture the early stages in far more detail than the 140°C data, which give a clearer picture of the latter stages), overlaying the data

in this way allows analysis of diagenesis across most stages without biasing interpretation towards the latest time points (Demarchi et al., 2013). One of the main disadvantages of this approach is that the Arrhenius parameters derived are highly sensitive to the range selected, and the process of range selection is partially subjective; for this reason, ranges must be provided alongside Arrhenius parameters, and the raw data made available on publication (Crisp et al., 2013; Demarchi et al., 2013; Tomiak et al., 2013). The other main disadvantage is that this method is purely mathematical, and has no underlying chemical rationale.

5.3.3: Results

5.3.3.1: THAA composition

For the sake of economy, amino acids are grouped here according to their chemical characteristics. Only the data from the chicken eggshell experiment are presented here; data for other species follow the same pattern.

While no significant linear ($R^2 \geq 0.95$) relationships were observed between heating time and the relative contribution of individual amino acids to the total intra-crystalline pool, pronounced changes did occur in the relative contributions of different types of amino acid (e.g. hydrophobic, hydrophilic). For example, as expected, the relative contribution of hydrophilic amino acids (serine and threonine; asparagine and glutamine could not be included due to deamidation during acid hydrolysis) diminishes with heating time; their peptide bonds are more amenable to hydrolysis, and they therefore become susceptible to decomposition (or loss by leaching) relatively rapidly. A concomitant increase in

the proportional representation of hydrophobic amino acids (valine, leucine, phenylalanine and isoleucine) was observed (Figure 5.8).

Although it is hydrophobic, alanine was not included with these as it (along with glycine) represents a breakdown product of other amino acids (e.g. Bada et al., 1978); distinguishing the original alanine pool from that produced by serine decomposition is beyond the scope of this project, and may be impossible. Although no increase in alanine or glycine concentration is observed with heating time during early diagenesis, it may be possible to detect whether the observed decrease in concentration appears artificially slow relative to other amino acids. Here, the rate of decrease in alanine concentration (measured as the slope of the line relating concentration to heating time) is among the slowest, and is consistent with that observed for other hydrophobic residues despite alanine having a lower normalised consensus hydrophobicity score (and thus being less resistant to hydrolysis) (Eisenberg, 1984). This may suggest a contribution of decomposed serine to the overall alanine pool, although this is not conclusive. Conversely, the rate of glycine decrease is among the highest observed, despite its moderate hydrophobicity (Eisenberg et al., 1984). This may suggest that there is a mechanism favouring rapid release of glycine (DKP formation (section 5.2.5.2.2) may be a candidate as it preferentially places glycine at terminal positions – Sepetov et al., 1991). Also, due to irreversible deamidation during sample preparation (Hill, 1965), the contribution of asparagine and glutamine to the hydrophilic amino acid pool could not be quantified.

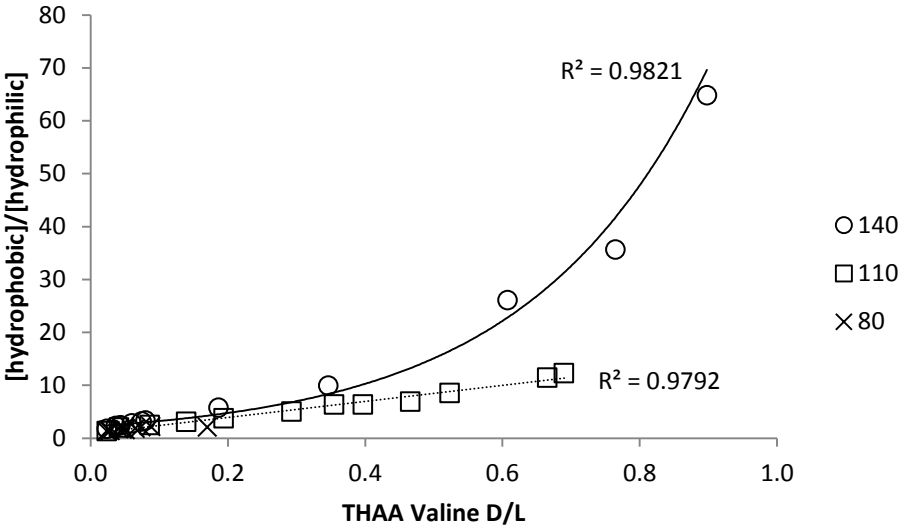

Figure 5.8: Increase in ratio of relative contribution of hydrophobic amino acids (not including alanine) to that of hydrophilic residues in heated chicken eggshell. Normalised against valine racemisation; this was chosen as THAA D/L vs. heating time follows a linear trajectory ($R^2 > 0.95$) at all temperatures, with the exception of the latest stages of diagenesis at 140°C. Note that the trajectories of the 110°C and 140°C data diverge after early diagenesis, with the latter following an exponential rather than linear trajectory.

5.3.3.2: FAA composition

Two reactions contribute to observed changes in FAA composition; hydrolysis of peptide bonds to form FAA, and subsequent decomposition into organic compounds (including, in some cases, other amino acids) (e.g. Vallentyne, 1964). The interplay of these at different temperatures may not always be consistent (e.g. Demarchi et al., 2013; Tomiak et al., 2013). Changes in FAA composition with

heating time track those in THAA composition; as heating time increases, hydrophobic amino acids increasingly dominate the FAA fraction (Figure 5.9).

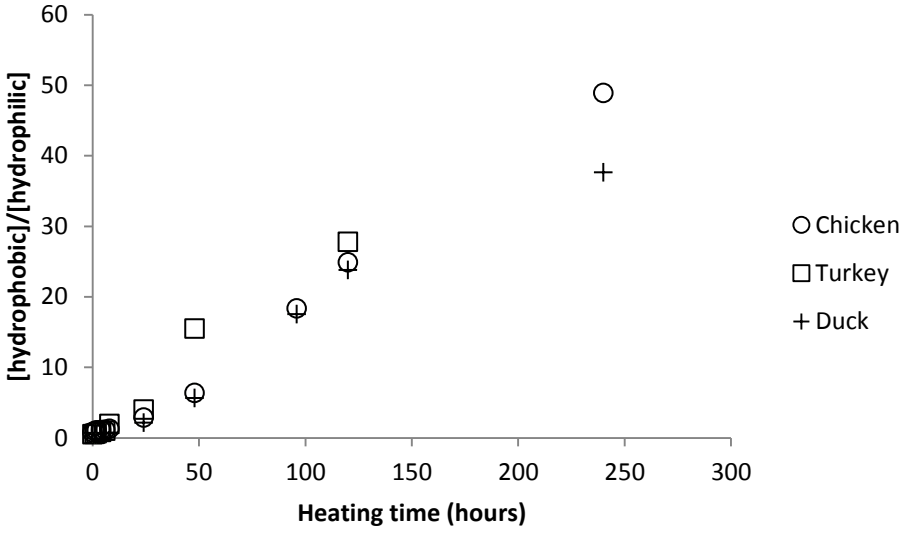

Figure 5.9: increase in ratio of relative contribution of hydrophobic FAA (not including alanine) to that of hydrophilic FAA with heating time at 140°C. The same pattern is observed at 110°C, but not at 80°C; as above, this probably relates to the limited diagenesis occurring at the lowest temperature. The relationship between heating time at 140°C and the ratio of hydrophobic (excluding alanine) to hydrophilic amino acids conforms to a linear model ($R^2 > 0.97$) in all species tested; the trajectory of the data is roughly consistent between species.

This may suggest that hydrophilic FAA are being lost to decomposition relatively rapidly. Decomposition is difficult to model based on these data. Serine FAA concentration increases (non-significantly – $0.9 < R^2 < 0.95$) during the 80°C experiment (Figure 5.10). This is expected; hydrophilic residues are most susceptible to hydrolysis, and decomposition should be relatively slow at this temperature. However, the rate of release of alanine is higher than would be

expected based on hydrophobicity alone (e.g. Eisenberg et al., 1984), both in terms of concentration (Figure 5.10) and relative contribution to the FAA fraction.

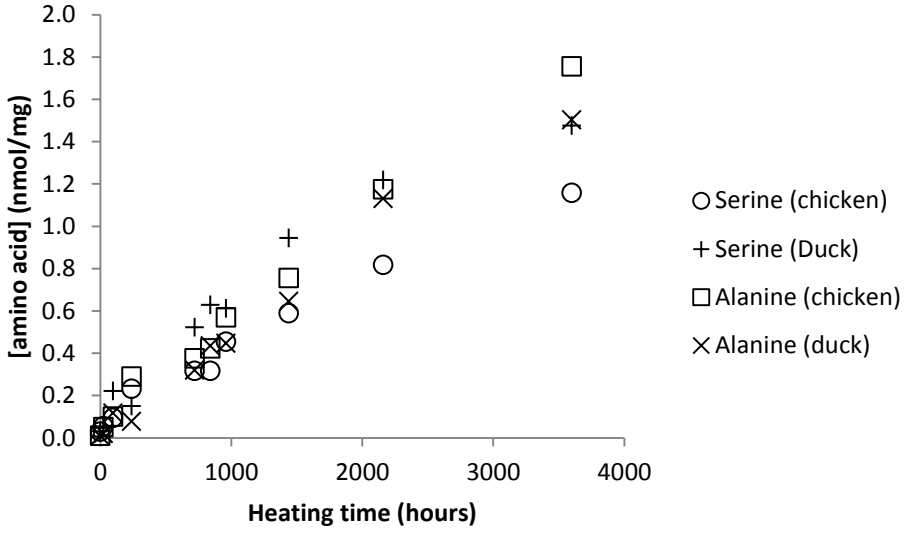

Figure 5.10: Increases in serine and alanine FAA concentration during heating at 80°C. Although the data are not presented, turkey and goose eggshell follow similar trajectories. The increase in alanine probably results from serine decomposition.

This relates to the decomposition pathways in operation; serine is unstable, and decomposes rapidly into alanine and glycine (as well as other organic compounds) (e.g. Bada et al., 1978; Walton, 1998). Alanine is hydrophobic, but less so than the other hydrophobic residues targeted in this study; glycine is either very slightly hydrophobic or very slightly hydrophilic (depending on the hydrophobicity scale used). It is therefore probable that the increases in alanine and glycine concentration and relative contribution in the FAA fraction during early diagenesis mostly represent serine decomposition (Figure 5.11). If this is the case, the composition of the FAA fraction during early diagenesis is conforming to the

expected model (i.e. early increase in hydrophilic residues, followed by domination by hydrophobic residues after more prolonged diagenesis). Once decomposition becomes a major factor, it is difficult to reconstruct the relationship between FAA composition and heating time.

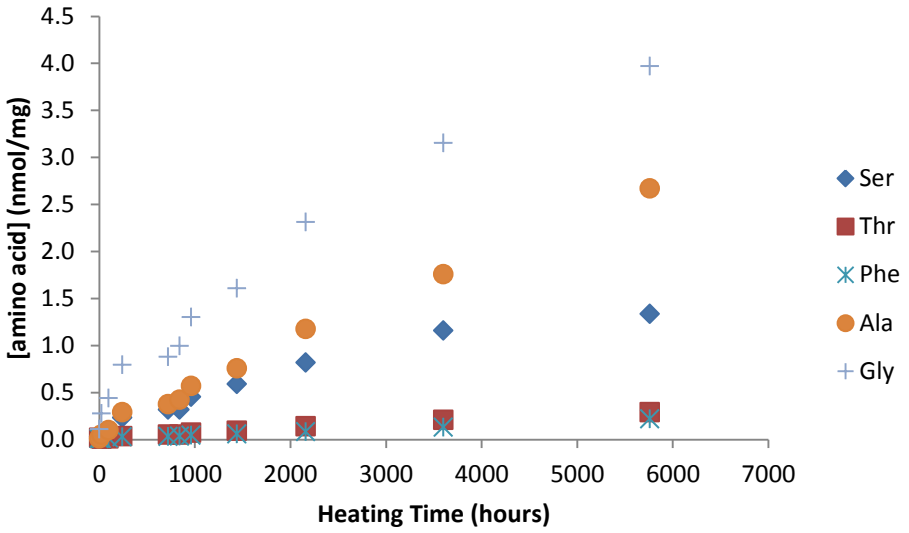

Figure 5.11 Increase in FAA concentration of amino acids with heating time at 80°C. The high rate of serine release suggests that peptide bonds adjacent to serine are particularly labile. The rates of release of alanine and glycine are higher even than that of serine; based on the relative hydrophobicity of amino acids, and known diagenetic pathways, such rapid release of these amino acids is unexpected. For reasons given above, it is probable that decomposition of serine is mostly responsible for the contribution of alanine to the FAA fraction during early diagenesis, while the factors governing rapid release of glycine are less clear.

5.3.3.3.1: Hydrolysis: general patterns

In theory, the extent of hydrolysis should increase from close to 0% FAA (little or no hydrolysis) in unheated samples towards 100% FAA at later points. In this experiment, complete (or almost complete) hydrolysis (95 – 100% FAA) is only observed in a few amino acids in chicken eggshell (Asx, threonine, glycine, and alanine), and in these only at the highest temperature (140^{o}C). In all species tested, a residual bound fraction of amino acids was observed at the end of the experiment (Figure 5.12). As would be expected, the slowest release of FAA occurs in hydrophobic amino acids and fastest in relatively hydrophilic residues (Figure 5.12). Hydrolysis did not reach completion in any amino acid in duck or turkey eggshell.

The retention of a residual bound fraction impacts upon rates of racemisation and decomposition, as in all but a few amino acids hydrolysis is a pre-requisite for these reactions to occur (e.g. Hare, 1971). When the rate of hydrolysis is considered (here using %FAA as a proxy), there are visible changes in rate during different stages of diagenesis. Following rapid hydrolysis during early diagenesis, there is a reduction in rate at later stages (Figure 5.13). This is consistent with data from a range of biominerals and amino acids, and most likely results from the retention of a hydrolysis-resistant fraction of proteins (Hoering, 1980; Walton, 1998; Collins & Riley, 2000; Miller et al., 2000; Penkman et al., 2008; Demarchi, 2009). This fraction may persist due to water becoming limiting during late diagenesis (e.g. Walton, 1998).

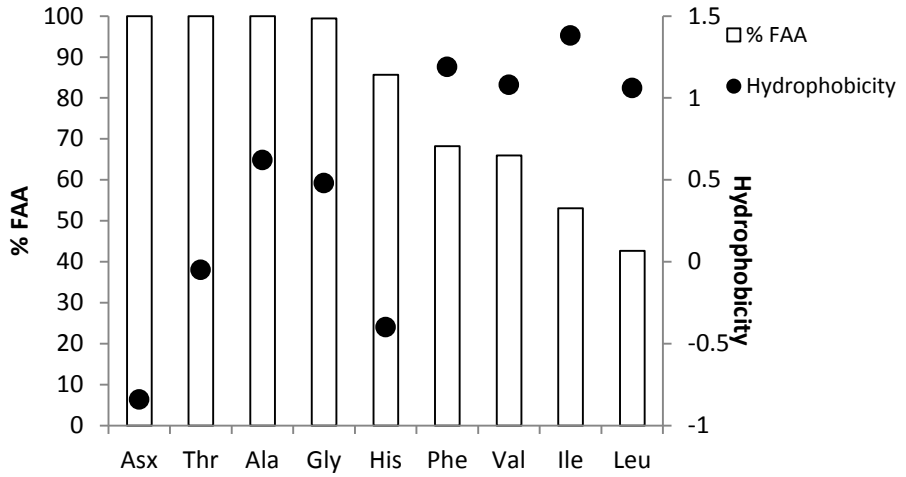

Figure 5.12: Extent of hydrolysis at last measured time point in heated (140°C) chicken eggshell, with hydrophobicity values (Eisenberg et al., 1984) given for each. As expected, rates of hydrolysis are clearly lower in the more hydrophobic residues.

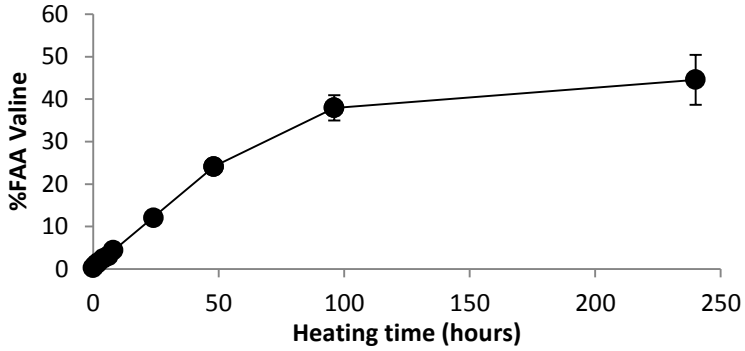

Figure 5.13: Changes in valine hydrolysis rate with heating time (140°C) in chicken eggshell. Similar patterns are observed in other species and amino acids (data not shown). The reduction in rate during late diagenesis is likely to result from water becoming limiting in the intra-crystalline fraction, although some water can be generated by other diagenetic processes such as decomposition and melanoidin formation (Bada et al., 1978; Rafalska et al., 1991). Error bars represent one standard deviation from the mean of three laboratory replicates.

5.3.3.3.2: Modelling hydrolysis using first-order kinetics

Arrhenius parameters for hydrolysis in five amino acids (valine, phenylalanine, leucine, isoleucine and Asx) were derived using an irreversible first-order kinetic (IFOK) model (Section 5.3.2.2) (table 5.3). Other amino acids recovered using the RP-HPLC method are not included; serine is very unstable, glycine and alanine represent a breakdown product of serine, Glx is poorly recovered as FAA due to formation of a stable lactam (Vallentyne, 1964), and in other amino acids the transformed data do not linearise well enough with heating time. As the relationship between hydrolysis and heating time is not linear across the whole observed range, only the initial linear portion of the reaction is considered (Figure 5.13). Subsequent stages could not be considered as the 80^{o}C data do not reach the first break of slope; these data (Tables 5.3 & 5.4) are therefore only applicable to early diagenesis. In most of the amino acids for which regression equations (Table 5.3) and Arrhenius parameters (Table 5.4) were determined, the log-transformed data gave a good linear fit ($R^2 \geq 0.95$ in most cases; $R^2 \geq 0.9$ in all) with heating time at all temperatures (Table 5.3).

All activation energies derived for hydrophobic amino acids (all except Asx) fall within the range of 113-120.5 kJ/mol (Table 5.3). This is higher than values observed in model peptides (83-99 kJ/mol), and is comparable to the range observed for these amino acids in limpet (*Patella vulgata*) shell (107 – 115 kJ/mol), and for valine and isoleucine in ostrich eggshell (114 kJ/mol) (Kriausakul & Mitterer, 1980; Qian et al., 1993; Demarchi, 2009; Crisp et al., 2013). As in *Patella* and other biominerals, the higher activation energies relative to model peptides result from the effect of protein conformation on hydrolysis; this does

not affect diagenesis of small peptides (e.g. Demarchi, 2009). As higher temperatures provide increased conformational relaxation, hydrolysis in biomineral proteins is likely to be highly affected by temperature (Collins & Riley, 2000). This may mean that the interaction between hydrolysis and other diagenetic processes is not the same across a range of temperatures; this can be tested by deriving Arrhenius parameters for racemisation across the same diagenetic range (Section 5.3.3.4).

Chicken			
	Regression equation	R^2	%FAA range
Asx	*ln(k) = -12607.3*(1/T) + 25.7*	*0.9977*	*2.1 – 31.7*
Isoleucine	ln(k) = -13792.0*(1/T) + 27.5	0.9664	0.1 – 6.3
Leucine	ln(k) = -13298.2*(1/T) + 26.7	0.9991	0.3 – 13.2
Phenylalanine	ln(k) = -13089.5*(1/T) + 26.2	0.9991	0.6 – 47.2
Valine	ln(k) = -13439.0*(1/T) + 18.7	0.9995	0.2 – 36.0
Turkey			
	Regression equation	R^2	%FAA range
Asx	ln(k) = -11439.2*(1/T) + 22.1	0.9997	1.8 – 23.7
Isoleucine	*ln(k) = -14135.3*(1/T) + 28.6*	*0.9917*	*1.7 – 6.5*
Leucine	ln(k) = -13096.4*(1/T) + 26.3	0.9985	0.3 – 12.3
Phenylalanine	*ln(k) = -13125.4*(1/T) + 26.3*	*0.9992*	*0.6 – 11.6*
Valine	*ln(k) = -13086.3*(1/T) + 26.1*	*0.9912*	*0.3 – 9.7*
Duck			
	Regression equation	R^2	%FAA range
Asx	ln(k) = -12455.4*(1/T) + 25.1	0.9996	0.7 – 40.7
Isoleucine	ln(k) = -13645.4*(1/T) + 27.2	0.9999	0.3 – 26.0
Leucine	ln(k) = -12955.4*(1/T) + 26.0	0.9979	0.2 – 27.8
Phenylalanine	ln(k) = -12885.8*(1/T) + 25.8	0.9992	0.5 – 27.2
Valine	*ln(k) = -13315.6*(1/T) + 26.9*	*0.9977*	*0.2 – 26.5*

Table 5.3: Regression equations for hydrolysis in five amino acids derived using IFOK model. Italicised entries represent cases where the linear fit between the log-transformed data and heating time was sub-optimal ($0.9 < R^2 < 0.95$) at one or more temperatures. The range of %FAA values over which these are applicable is given below (Table 5.4). The %FAA range is calculated as the range of %FAA values

represented at all three temperatures (i.e. at the end-point of the 80°C experiment).

Chicken		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Asx	*111.3*	*25.7*
Isoleucine	121.8	27.5
Leucine	117.4	26.7
Phenylalanine	115.6	26.2
Valine	118.7	25.6
Turkey		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Asx	101.0	22.1
Isoleucine	*125.8*	*28.6*
Leucine	115.7	26.3
Phenylalanine	*115.9*	*26.3*
Valine	*115.6*	*26.1*
Duck		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Asx	110.0	25.1
Isoleucine	120.5	27.2
Leucine	115.4	26.0
Phenylalanine	113.8	25.8
Valine	*117.6*	*26.9*

Table 5.4: Arrhenius parameters for hydrolysis of five amino acids during the initial linear derived using an IFOK model. Italicised entries represent cases where the linear fit between the log-transformed data and heating time gave a sub-optimal linear fit ($0.9 < R^2 < 0.95$) at one or more temperatures.

In most cases, activation energies for hydrolysis were closely matched in different species across the diagenetic range considered (Table 5.3). In valine, phenylalanine and leucine, all data fall within the expected analytical error range (given here as ± 6 kJ/mol); the range of values for isoleucine is also fairly tight

(8.4 kJ/mol) (Table 5.3). In all three species, the relative order of amino acids followed the same pattern; Asx returned the lowest activation energies, and isoleucine the highest (Table 5.3). The other amino acids gave almost identical activation energies (range = 3.8kJ/mol).

5.3.3.3.3: Modelling hydrolysis using a 3rd order scaling method

Regression equations (Table 5.5) and Arrhenius parameters (Table 5.6) were also derived for the same amino acids using a 3rd-order scaling method (Section 5.3.2.4). These parameters were calculated across the same range of hydrolysis as the 1st-order model described above in order to allow comparison of the two techniques. Although in theory the two techniques should provide the same results when applied to the same range, different Arrhenius parameters are derived using this method (Table 5.6); these fall within a range (103 – 114 kJ/mol) which is consistent with those observed in *Patella* shell, ostrich eggshell (with the exception of phenylalanine), and in other biominerals (Demarchi, 2009; Collins & Riley, 2000; Crisp et al., 2013).

Chicken			
	Regression equation	R^2	%FAA range
Asx	ln(k) = -13000.7*(1/T) + 33.9	0.9999	2.0 – 33.0
Isoleucine	ln(k) = -12465.5*(1/T) + 32.6	0.9998	0.0 – 5.0
Leucine	ln(k) = -12659.8*(1/T) + 33.1	0.9999	0.0 – 13.5
Phenylalanine	ln(k) = -12577.6*(1/T) + 32.8	0.9999	1.0 – 13.0
Valine	ln(k) = -13043.6*(1/T) + 35.2	0.9988	0.0 – 10.0
Turkey			
	Regression equation	R^2	%FAA range
Asx	ln(k) = -12528.9*(1/T) + 33.0	0.9927	1.8 – 25.0
Isoleucine	ln(k) = -13322.8*(1/T) + 34.7	0.9993	0.0 – 6.4
Leucine	ln(k) = -12981.6*(1/T) + 33.8	0.9998	0.0 -12.3
Phenylalanine	ln(k) = -12360.3*(1/T) + 32.3	0.9999	0.6 – 11.6
Valine	ln(k) = -13175.4*(1/T) + 35.6	0.9967	0.0 – 9.7
Duck			
	Regression equation	R^2	%FAA range
Asx	ln(k) = -12708.9*(1/T) + 33.0	0.9948	0.7 – 15.2
Isoleucine	ln(k) = -12619.4*(1/T) + 33.1	0.9989	0.0 – 5.3
Leucine	ln(k) = -12758.2*(1/T) + 33.4	0.9992	0.0 – 10.4
Phenylalanine	ln(k) = -12716.5*(1/T) + 33.3	0.9996	0.4 – 10.0
Valine	ln(k) = -13442.7*(1/T) + 35.0	0.9998	0.2 – 9.8

Table 5.5: Regression equations for hydrolysis in five amino acids derived using 3rd order scaling model. The maximum %FAA for each amino acid to which these apply is also given. The %FAA range is calculated as the range of %FAA values represented at all three temperatures (i.e. at the end-point of the 80^{o}C experiment).

Chicken		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Asx	108.1	33.9
Isoleucine	107.2	33.67
Leucine	107.5	33.77
Phenylalanine	103.5	32.5
Valine	108.5	35.15
Turkey		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Asx	105.2	33.01
Isoleucine	110.8	35.2
Leucine	107.9	33.84
Phenylalanine	102.8	31.83
Valine	109.5	35.56
Duck		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Asx	105.7	32.95
Isoleucine	114.0	35.89
Leucine	114.0	35.83
Phenylalanine	111.2	35.96
Valine	111.8	35.04

Table 5.6: Arrhenius parameters for hydrolysis during early diagenesis calculated using 3rd-order scaling method.

5.3.3.3.4: Comparing the models

When the two techniques are compared, the IFOK model gives systematically higher activation energies for hydrolysis than the 3rd order scaling method (Figure 5.14). While some variability between the techniques is expected due to their application over slightly different ranges (Tables 5.3 and 5.4), the fact that this seems to be a systematic difference highlights the need for careful model selection when considering kinetic data. The correspondence between the two

methods is best In duck eggshell, in which the scaling method gives the most similar (within 5%) activation energies to the 1st-order model (Table 5.6). In chicken and turkey eggshell, the activation energies derived using the scaling method are consistently between 3.5% and 12% lower than those derived using the IFOK model (Table 5.6). The only exception is Asx hydrolysis in turkey eggshell, which has a slightly higher (3.2%) activation energy using the scaling method. The data for chicken and turkey are very similar; the maximum difference between the species in a single amino acid is 3.6 kJ/mol (Table 5.6). This reflects the fact that the molecular environments in these species are very similar. The difference between these data and the duck data (Table 5.6) probably relates to differences between the molecular environments.

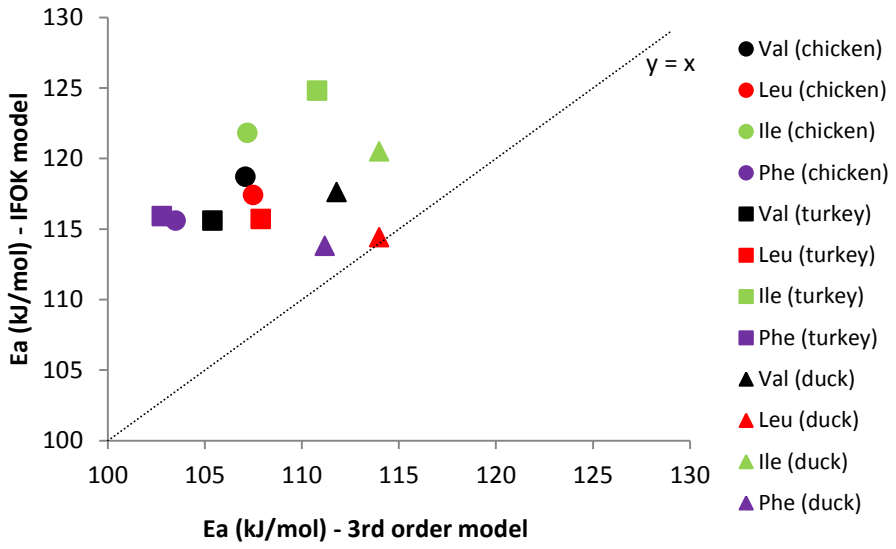

Figure 5.14: Activation energies for hydrolysis of four amino acids derived using IFOK and 3rd order models. The IFOK model seems to give systematically higher activation energies. The correspondence between the two models is closest in duck eggshell.

5.3.3.4.1: Racemisation: variability

All subsequent data for racemisation are derived from averaging of three laboratory replicates; the variability of DL ratios between samples must first be quantified. In eggshell, the coefficient of variation (CV) is below 10% in almost all cases; the only exceptions are found during early diagenesis, where DL ratios in some amino acids are very low and D-enantiomers are difficult to detect. It is likely that the higher CV values seen in these cases relate to instrument error as much as to genuine variability. Serine also becomes highly variable during late diagenesis. Unlike in *Patella* shell (Demarchi, 2009) there is no clear relationship between the level of variability and the relative abundance of different amino acids.

CV (%)	Asx	Glx	Ser	Ala	Val	Phe	Leu	Ile
Chicken	2.1	2.2	7.3	2.9	3.0	2.9	3.2	5.9
Duck	2.5	3.7	5.2	5.1	8.0	3.8	5.2	5.9
Turkey	3.0	5.5	5.8	5.7	9.7	6.5	5.4	8.3

Table 5.7: Mean coefficient of variation (CV) values for amino acid THAA DL ratios, across all temperatures and heating times. These data are consistent with the level of variability observed in Patella *shell (Demarchi, 2009), and slightly higher than those observed for isoleucine epimerisation in ostrich eggshell (Miller et al., 1992) and bivalve shell (Miller & Brigham-Grette, 1989).*

5.3.3.4.2: Racemisation: general patterns

According to the three-box model of diagenesis, racemisation patterns in most amino acids should track those observed in hydrolysis (e.g. Hare, 1971; Collins & Riley, 2000). During the earliest stages of diagenesis, racemisation (measured as

D/L) should be close to zero, as most peptides are intact. Also, in non-closed systems, racemised FAA may have been lost from the system by diffusion. As hydrolysis proceeds, D/L values should increase until they reach a dynamic equilibrium, after which interconversion continues but is not detectable. In a closed system like the intra-crystalline fraction of eggshell, diffusive loss of racemised FAA should not be a major contributor to observed changes in D/L (Chapter 2); decomposition of these is likely to have a much stronger effect.

In the intra-crystalline fraction of eggshell amino acids, racemisation initially proceeds at a fast rate as hydrolysis leads to production of new terminal positions, and racemised FAA are retained within the system. During later diagenesis, the observed rate decreases markedly (Figure 5.15). This may result from decomposition of racemised FAA creating an artificially low DL ratio, and/or from slow hydrolysis of a residual peptide-bound fraction of amino acids (e.g. Collins & Riley, 2000). In Asx, the initial rate of racemisation is extremely rapid, probably due to rapid racemisation via a succinimide intermediate (Section 5.2.5.2.3) (Geiger & Clarke, 1987). DL ratios in serine appear to decrease during late diagenesis due to rapid decomposition (Walton, 1998).

The relative order of racemisation rate in the amino acids under consideration is consistent between all of the species tested: Asx (and Ser during early diagenesis) > Ala/Phe > Glx > Leu > Ile > Val (fastest to slowest). Except for the position of phenylalanine, this is consistent with the relative order observed in

heated *Patella* shell (Demarchi, 2009) and in free amino acids in solution (Smith et al., 1978).

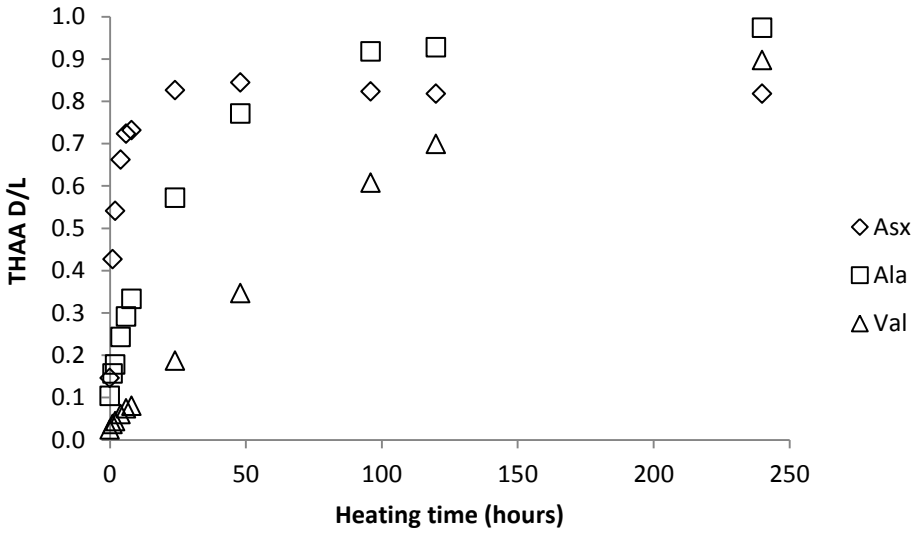

Figure 5.15: Racemisation rate in three amino acids in chicken eggshell, heated at 140°C. These are representative of the general patterns observed in all species tested, and in all amino acids (with the exception of serine). These are averages values obtained from triplicates; standard deviation < 0.06 in all cases.

5.3.3.4.3: Modelling racemisation using first-order kinetics

When compared to data derived from *Patella* shell (Demarchi, 2009) and ratite eggshell (Miller et al., 2000), in which at least one break of slope is observed, the log-transformed data (Section 5.3.2.2) for some amino acids in avian eggshell gives a relatively good linear fit with heating time (Figure 5.16). In some cases, the linear fit was only good over part of the range of D/L values observed (Table 5.8); this relates in some cases (Phe, Leu) to the inability of the equation to deal with D/L values over 1 (except in diastereomeric isoleucine, for which the expression is modified), in some (Ala) to an increasing contribution of racemised

FAA resulting from decomposition of other residues, and in others to genuinely different racemisation patterns (Asx).

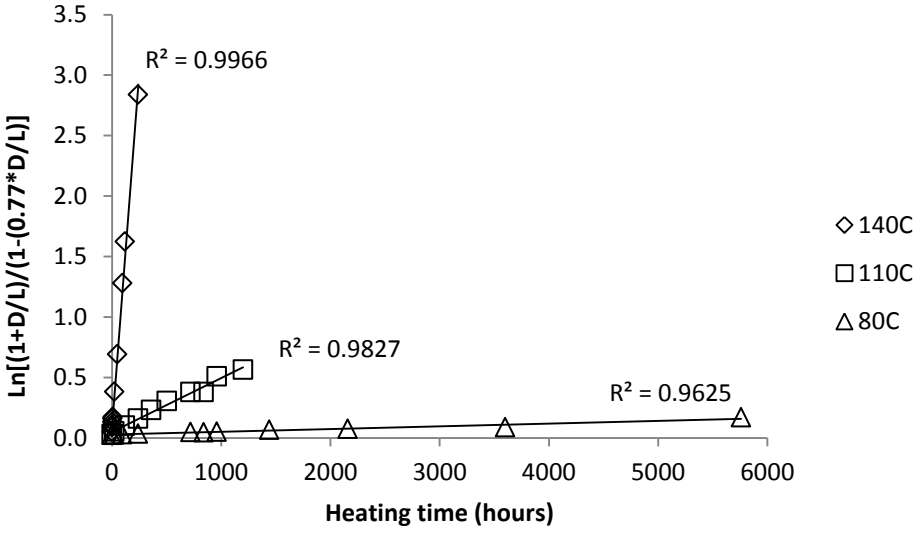

Figure 5.16: Plot of 1st-order log-transformed THAA isoleucine racemisation vs. heating time at three temperatures in chicken eggshell. The good linear fit allows a simple first-order equation to be used in calculating Arrhenius parameters.

The linear fit between the log-transformed data and heating time was good ($R^2 \geq 0.95$) across most of the range of D/L values observed in two amino acids (valine and isoleucine) in all three species; the fit was also reasonably good ($R_2 \geq 0.9$) for leucine in all three species, for Glx in chicken eggshell, and for phenylalanine in duck and chicken eggshell. In most cases, the 140^{o}C data were truncated and do not incorporate the observed reduction in racemisation rate during late diagenesis. Regression equations (Table 5.8) and Arrhenius parameters (Table 5.9) for racemisation were subsequently derived for five amino acids (valine, phenylalanine, leucine, isoleucine and Glx) using a reversible first-order kinetic (RFOK) model.

Chicken			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -13895.1*(1/T) + 28.5	0.9999	0.03 – 0.17
Leucine	*ln(k) = -15259.0*(1/T) + 24.4*	*0.9993*	*0.05 – 0.21*
Phenylalanine	*ln(k) = -15006.2*(1/T) + 24.4*	*0.9976*	*0.05 – 0.33*
Valine	ln(k) = -13585.4*(1/T) + 19.6	0.9999	0.02 – 0.17
Glx	*ln(k) = -12349.9*(1/T) + 16.5*	*0.9935*	*0.07 – 0.31*
Turkey			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -13695.4*(1/T) + 28.2	0.9999	0.02 – 0.12
Leucine	ln(k) = -14376.2*(1/T) + 30.6	0.9989	0.03 – 0.17
Phenylalanine	n.d.	n.d.	n.d.
Valine	ln(k) = -13906.3*(1/T) + 28.8	0.9999	0.03 – 0.12
Glx	n.d.	n.d.	n.d.
Duck			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -13267.0*(1/T) + 26.9	0.9994	0.00 – 0.17
Leucine	*ln(k) = -15286.7*(1/T) + 33.0*	*0.9944*	*0.05 – 0.27*
Phenylalanine	*ln(k) = -15516.0*(1/T) + 35.0*	*0.9997*	*0.04 – 0.36*
Valine	ln(k) = -13625.3*(1/T) + 28.0	0.9982	0.03 – 0.18
Glx	n.d.	n.d.	n.d.

Table 5.8: Regression equations derived for racemisation in six amino acids derived using the RFOK model. n.d. = not determined due to poor linearization (R2 < 0.9) of log-transformed data. Italicised entries represent cases where the linear fit between the log-transformed data and heating time was sub-optimal (0.9 < R2 < 0.95) at one or more temperatures. D/L ranges are calculated as the maximum D/L value for which data are available at all three temperatures (i.e. at the end-point of the 80°C experiment).

Chicken		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Isoleucine	122.7	20.3
Leucine	134.8	24.4
Phenylalanine	132.5	24.4
Valine	120.0	19.6
Glx	109.1	16.5
Turkey		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Isoleucine	121.0	28.2
Leucine	127.0	30.6
Valine	122.8	28.8
Duck		
	Ea (kJ/mol)	Ln(A) (s^{-1})
Isoleucine	120.3	27.9
Leucine	137.0	35.0
Phenylalanine	135.0	33.0
Valine	117.2	26.9

Table 5.9: Arrhenius parameters for THAA racemisation in the intra-crystalline fraction of the eggshell of four avian species derived using RFOK model. Ea = activation energy; ln(A) = frequency factor.

In isoleucine, valine and phenylalanine, activation energies fall within a relatively tight range in all three species (± ≤ 5.5 kJ/mol) (Table 5.9). The activation energy for leucine is higher in duck and turkey eggshell than in chicken eggshell (Table 5.9). In general, the activation energies for leucine and phenylalanine seem unrealistically high; this may result from the use of truncated data forced by DL ratios over 1. This may partly relate to poor resolution of L-phenylalanine and D-leucine using the reverse-phase HPLC method employed here, which may have led to overestimation of DL ratios in some cases. Slight differences in alanine

activation energy between species may relate to differences between the molecular environments, which are likely to incorporate a substantial component representing decomposed serine.

For valine and isoleucine, the activation energies derived are lower than those observed in ostrich eggshell (126 and 128 kJ/mol respectively), while the values derived for phenylalanine are higher than in ostrich eggshell (129 kJ/mol) (Crisp et al., 2013). This difference may be partly explained by the fact that these were derived over different D/L ranges (Crisp et al., 2013).

5.3.3.4.4: Modelling racemisation using a 3rd order scaling method

As discussed in a previous section (Section 5.3.2.3), the main drawback of FOK models is the forcing of a linear fit to the log-transformed data. 3rd-order scaling (Section 5.3.2.4) works from the raw data in order to avoid this particular bias. This method is purely mathematical, and has no underlying chemical rationale (Demarchi et al., 2013). Phenylalanine is not included, as the data for this amino acid provided a very poor fit to the 3rd order polynomial model. Regression equations (Table 5.10) and Arrhenius parameters (Table 5.11) were calculated for three amino acids (valine, leucine and isoleucine).

Chicken			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -14597.1*(1/T) + 38.1	0.9999	0.03 – 0.17
Leucine	ln(k) = -14399.0*(1/T) + 37.4	0.9961	0.05 – 0.21
Valine	ln(k) = -14161.6*(1/T) + 37.0	0.9999	0.02 – 0.17
Turkey			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -15208.2*(1/T) + 39.7	0.9999	0.02 – 0.12
Leucine	ln(k) = -14688.1*(1/T) + 38.4	0.9996	0.03 – 0.18
Valine	ln(k) = -15273.1*(1/T) + 39.9	0.9999	0.03 – 0.12
Duck			
	Regression equation	R^2	D/L range
Isoleucine	ln(k) = -14662.0*(1/T) + 38.4	0.9991	0.00 – 0.17
Leucine	ln(k) = -14775.3*(1/T) + 38.6	0.9999	0.05 – 0.27
Valine	ln(k) = -14553.3*(1/T) + 38.1	0.9994	0.03 – 0.18

Table 5.10: Regression equations for racemisation in three amino acids derived using a 3rd order scaling model. D/L ranges are calculated as the maximum D/L value for which data are available at all three temperatures (i.e. at the end-point of the 80^{o}C experiment).

icken		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Isoleucine	121.4	38.1
Leucine	119.8	37.6
Valine	117.7	37.0
Turkey		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Isoleucine	126.4	39.7
Leucine	122.1	38.4
Valine	127	39.9
Duck		
	Ea (kJ/mol)	Ln (A) (s^{-1})
Isoleucine	121.9	38.4
Leucine	122.8	38.6
Valine	121	38.1

Table 5.11: Arrhenius parameters for racemisation in three amino acids derived using a 3rd-order scaling method; the D/L range across which these data apply is also given.

Within the three amino acids for which Arrhenius parameters were derived using this method, the activation energies are fairly tight (maximum difference = 9.3 kJ/mol). The values obtained for valine and isoleucine are generally similar to those obtained for ostrich eggshell in chicken and duck eggshell (118 kJ/mol for valine, 120 kJ/mol for isoleucine) (N.B. Arrhenius parameters for leucine have not been derived in ostrich eggshell) (Crisp et al., 2013).

5.3.3.4.5: Comparing the models

A direct comparison between the models is only possible for three amino acids (valine, leucine and isoleucine). Unlike hydrolysis (Figure 5.14), there is no systematic difference between the models, and in most cases the difference between them is slight relative to that observed for hydrolysis (Figure 5.17). The only exception is in leucine racemisation, for which the RFOK model gives very high activation energies in chicken and duck eggshell (Figure 5.17). This probably relates to the inability of the RFOK data transformation to accommodate DL ratios above 1; the fact that leucine racemisation in turkey eggshell, in which D/L does not reach 1, falls close to the same line as the other amino acids supports this, and suggests that the 3rd order model is more accurate for this amino acid in chicken and duck eggshell (Figure 5.17).

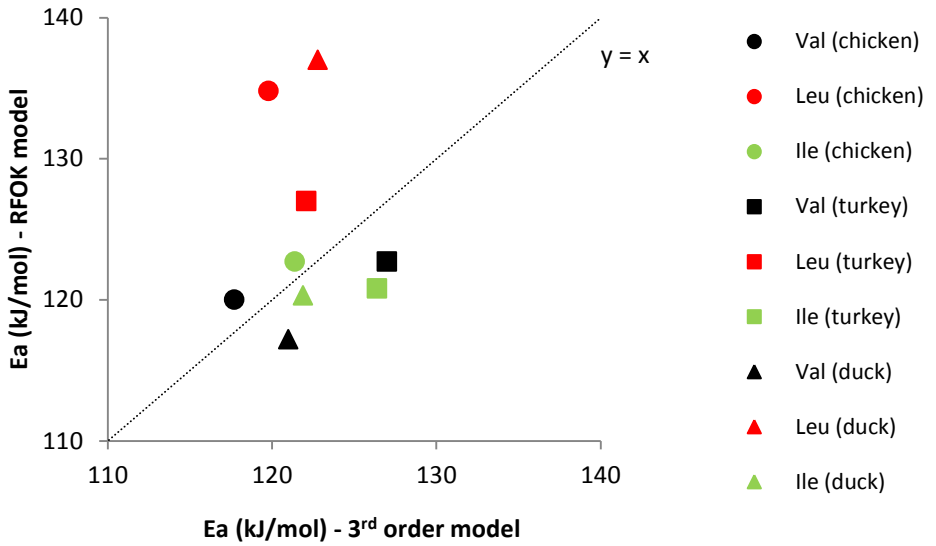

Figure 5.17: Comparison of activation energies for racemisation derived using RFOK and 3rd order models. According to these data the two models do not produce systematically different activation energies.

5.3.3.4.6: Racemisation vs. hydrolysis

One of the major assumptions of using artificially accelerated diagenesis to model pathways of diagenesis at lower temperatures is that the temperature dependence of different diagenetic processes is the same (or at least very similar) across a wide temperature range (e.g. Collins & Riley, 2000; Miller et al., 2000; Demarchi, 2009). This is particularly the case for racemisation, as reaction rate is ultimately dependent upon the rate of hydrolysis (e.g. Collins & Riley, 2000). Where possible, comparing the activation energies of hydrolysis and racemisation (Figures 5.18 & 5.19) derived here allows this to be tested, despite the potentially large error margins attached to these data (e.g. McCoy, 1987).

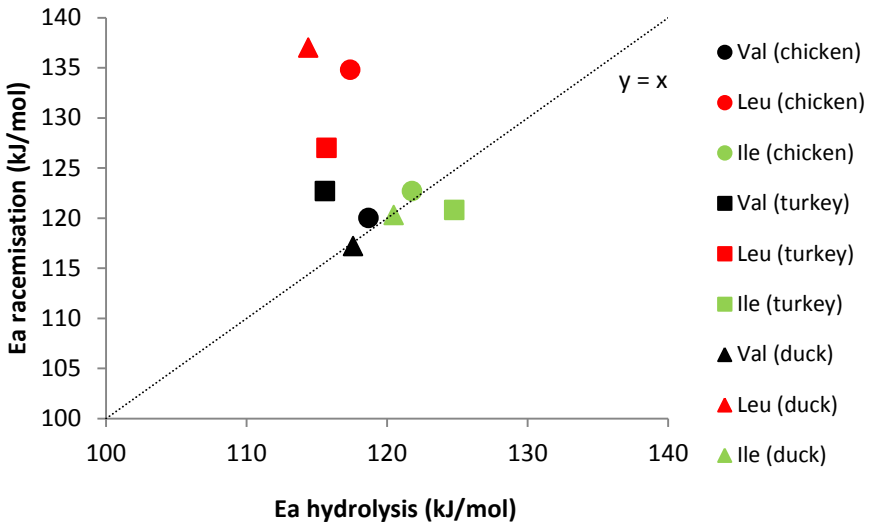

Figure 5.18: Comparison of activation energies for hydrolysis and racemisation derived using FOK models. The correspondence observed between isoleucine and valine in chicken and turkey is very strong, suggesting that these might provide a basis for estimating the temperature dependence of these reactions (Section 5.3.3.6).

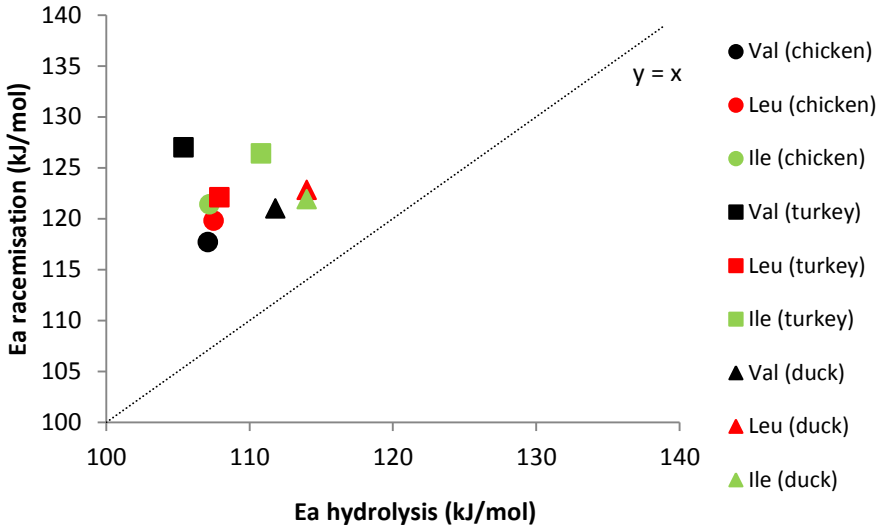

Figure 5.19: Comparison of activation energies for hydrolysis and racemisation derived using 3rd order model. Using this model, there are no examples where

the correspondence between the reactions is close enough to justify extrapolation to a wider temperature range.

When the first-order models are considered, the correspondence between the activation energies of racemisation and hydrolysis appears to be close for valine and isoleucine in chicken and duck eggshell (Figure 5.18). The same amino acids in turkey are a bit less closely matched, but can still be considered similar given the error margins of these techniques (difference = 7.1 kJ/mol for valine, 4 kJ/mol for isoleucine) (Figure 5.18). Based on these data, valine and isoleucine racemisation and hydrolysis can be considered potentially useful in modelling diagenesis at ambient temperatures, although this is complicated by the lack of data for decomposition. Based on these data, leucine is not useful for this purpose (although, as shown above, the activation energies derived for leucine racemisation are probably misleadingly high).

The 3rd order scaling model provides no example of amino acids in which the activation energies of hydrolysis and racemisation are closely matched (Figure 5.19). Based on these data, it could be concluded that the use of high temperature experiments to model diagenesis at lower temperatures is unfounded in the systems examined here. Much depends on the cause of the systematic difference in hydrolysis activation energies observed between the two models; if overestimation by the IFOK model is the cause, then these data cannot be used to model expected racemisation rates at lower temperature; if underestimation by the 3rd order model is the main cause, then there are some

examples of amino acids which can potentially be used in this context (Figure 5.18). This problem is not unique to avian eggshell, and the answer is not clear. For example, the same pattern has been observed in *Patella* shell; the activation energies of these reactions in isoleucine and valine are closely matched when derived using FOK models, but not when derived using the 3rd order scaling method (Demarchi, 2009; Demarchi et al., 2013). Answering this problem definitively is beyond the scope of this project. It is, however, possible to test whether the patterns of diagenesis observed here are also observed in samples of known age and deposition temperature (Section 5.3.3.6).

5.3.3.5: Decomposition

Of all of the known pathways of amino acid decomposition (Section 5.2.4), only decomposition of serine into alanine (measured as [Ser]/[Ala]) is directly measurable by the technique used in this study. As observed in some biominerals (Bada et al., 1978; Bada & Man, 1980; Penkman et al., 2008), but not in others (Demarchi, 2009), the decrease in [Ser]/[Ala] during diagenesis conforms to an exponential model in the eggshell of all three species (Figure 5.20). Although developing serine decomposition to alanine as a chronological tool (e.g. Westaway, 2009) is beyond the scope of this study, these data suggest that this might be viable in avian eggshell.

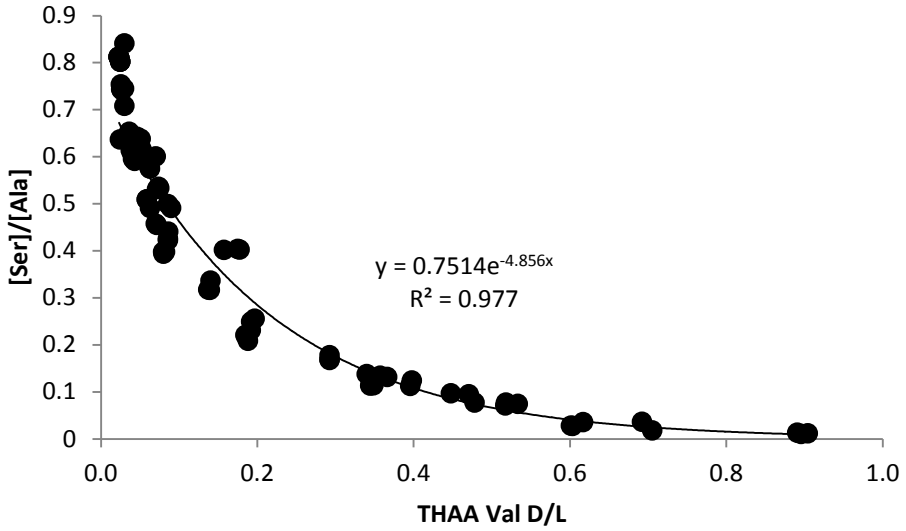

Figure 5.20: Exponential decrease in [Ser]/[Ala] during diagenesis in chicken eggshell heated at 80, 110 and 140°C, normalised against THAA valine racemisation. Similar exponential decreases (R^2 > 0.96) were observed in duck and turkey eggshell (data not shown). These data suggest that this facet of decomposition might provide a basis for relative dating (e.g. Westaway, 2009).

5.3.3.6.1: Archaeological AAR kinetics: overview

The Arrhenius parameters obtained in previous subsections are applicable only to the temperature range from which they were derived; correspondence with low-temperature diagenesis cannot be assumed. THAA D/L and %FAA values were obtained from archaeological fragments of chicken eggshell (context 305, Bornais, South Uist – Chapter 4) in order to test whether the temperature sensitivity of racemisation and hydrolysis is consistent across a wide temperature range. These samples date from between 860 and 1040 years before present (Sharples, pers comm). An effective deposition temperature of 7°C was estimated based on the geographical location, sediment type, sample

type and approximate sample depth (www.thermal-age.eu). There have been no examples of duck (or turkey) eggshell of sufficient quantity to allow this type of analysis from any site studied to date.

5.3.3.6.2: Variability

Ten separate fragments from the same context were prepared in duplicate (using the methods described in Section 5.3.2.1) in order to provide an indication of the level of variability in the samples (Table 5.12). Of the ten replicates, one reported far higher D/L and %FAA values than the other nine; this probably suggests that this fragment had experienced higher temperatures (possibly through exposure to fire) (e.g. Crisp et al., 2013). This fragment was not used in subsequent calculations of average DL ratios and %FAA. Among the other nine replicates, the values obtained are reasonably consistent (Table 5.12). In general, FAA concentrations were very low, and in some cases D-enantiomers were badly resolved (particularly in leucine).

	% FAA	THAA D/L
Asx	2.8 ± 0.9	0.21 ± 0.02
Glx	1.1 ± 0.4	0.07 ± 0.03
Ser	5.0 ± 0.4	0.20 ± 0.01
Thr	1.1 ± 0.2	
His	0.9 ± 0.1	
Gly	3.9 ± 1.7	
Arg	1.3 ± 0.3	
Ala	2.7 ± 3.1	0.06 ± 0.02
Val	1.3 ± 1.5	0.03 ± 0.02
Phe	1.7 ± 1.0	0.04 ± 0.01
Leu	0.8 ± 0.6	0.05 ± 0.01
Ile	0.9 ± 0.7	0.02 ± 0.01
Total	2.3 ± 1.0	

Table 5.12: Variability in %FAA and THAA D/L in chicken eggshell fragments from context 305, Bornais. Error margins represent one standard deviation from the mean. %FAA indicates the percentage of the total pool of each amino acid present as FAA.

5.3.3.6.3: Hydrolysis

Despite the very limited extent of %FAA observed (Table 5.12), the rate of hydrolysis in the archaeological samples could be calculated for four amino acids (valine, phenylalanine, leucine and isoleucine) using the IFOK model. The observed %FAA ranges were too low to allow meaningful comparison with the high temperature data using the 3rd order model, and have no discernible effect on the Arrhenius parameters derived (data not shown). Other amino acids were not considered due to poor linearization by the IFOK data transformation (Asx, Glx, serine) and/or due to inability to isolate the original pool from that resulting from diagenesis of other residues (alanine, glycine). Of those which were considered, the observed rate was consistent with the projected rate in three (valine, phenylalanine and leucine) across the observed %FAA range. The activation energies derived across the different temperature ranges are particularly close in valine (Figure 5.21) and isoleucine (Figure 5.22); reasonably close in phenylalanine; and somewhat divergent in leucine (Table 5.13). Although these data seem promising, they must be treated with extreme caution; the correspondence between high and low temperature data remains to be demonstrated across all but a minuscule diagenetic range.

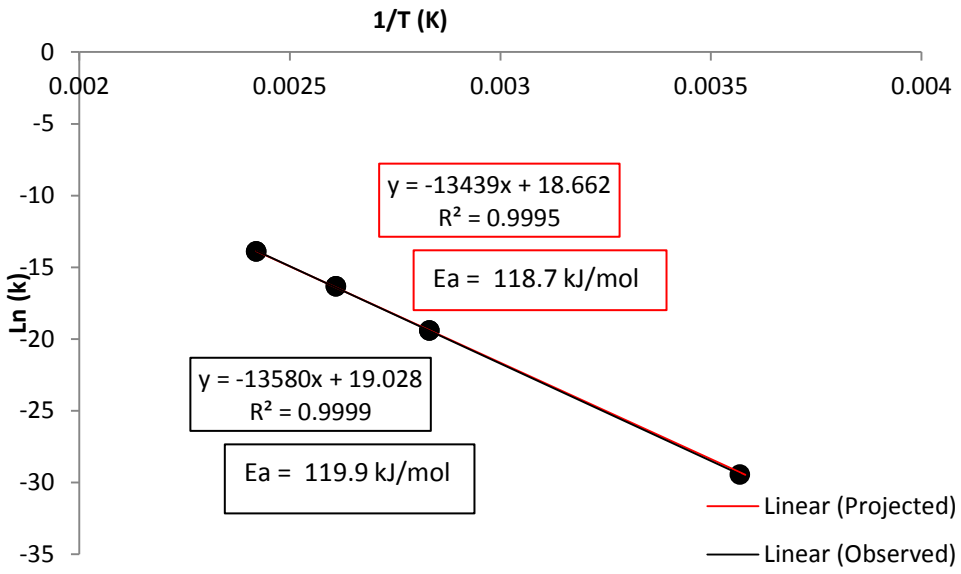

Figure 5.21: Very close correspondence of activation energies for valine hydrolysis in experimental and archaeological samples. The age of the archaeological samples is derived as the average of the earliest and latest possible dates; when the end points are used, $R^2 \geq 0.9999$ in all cases and activation energies differ by less than 1.5 kJ/mol between the earliest and latest date.

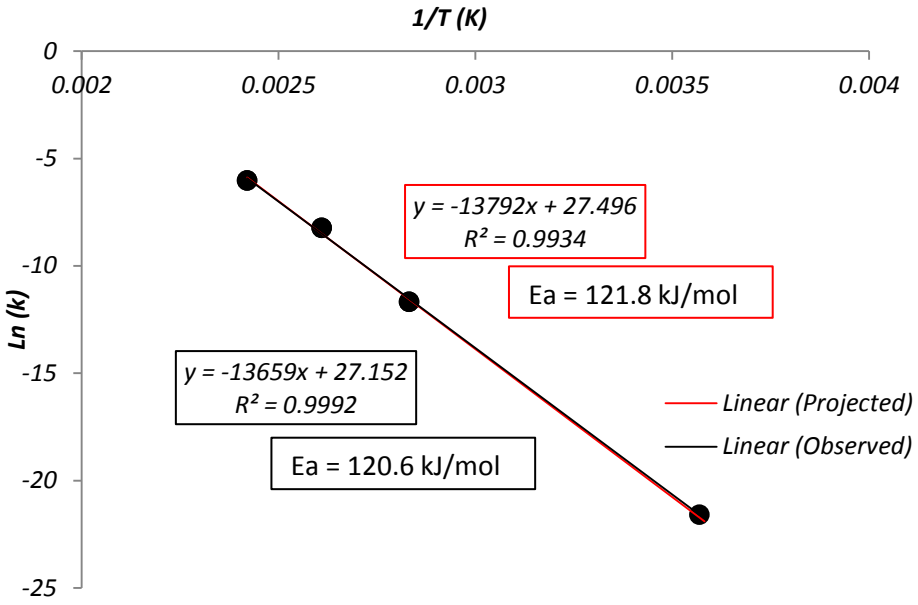

Figure 5.22: Very good correspondence between projected and observed hydrolysis rates in isoleucine. The age of the archaeological samples is derived as the average of the earliest and latest possible dates; when the end points are used, $R^2 \geq 0.9999$ in all cases and activation energies differ by less than 1.5 kJ/mol between the earliest and latest date.

Range (°C)	Val	Phe	Leu	Ile
80-140	118.7	115.6	117.4	121.8
7-140	119.9	121.5	128.3	120.6

Table 5.13: Activation energies (kJ/mol) for hydrolysis in three amino acids derived using different temperature ranges.

5.3.3.6.4: Racemisation

Using the first-order model, racemisation rates in the archaeological samples could be calculated for five amino acids (valine, phenylalanine, leucine, isoleucine and Glx). Others were not considered for the same reasons as given

above for hydrolysis. In general, the correspondence between the data from the two temperature ranges is not as tight as for hydrolysis (Table 5.14). As for hydrolysis, the archaeological data covered an insufficient D/L range to noticeably affect the Arrhenius parameters derived using the 3rd order model (data not shown).

Range (°C)	Glx	Val	Phe	Leu	Ile
80-140	109.1	120.0	132.6	135.8	122.7
7-140	121.0	128.4	137.6	127.4	129.9

Table 5.14: Activation energies for racemisation in five amino acids derived using different temperature ranges. Glx is included despite relatively poor conformation to FOK models.

The fact that the activation energies derived using the first-order model differ between the high and low temperature data across such a small D/L range (Figure 5.23) suggests that racemisation pathways are likely to be highly temperature-dependent. This has negative implications for the use of amino acid racemisation for absolute dating of non-ratite eggshell, although relative dating and calibration using independently dated samples may still be viable.

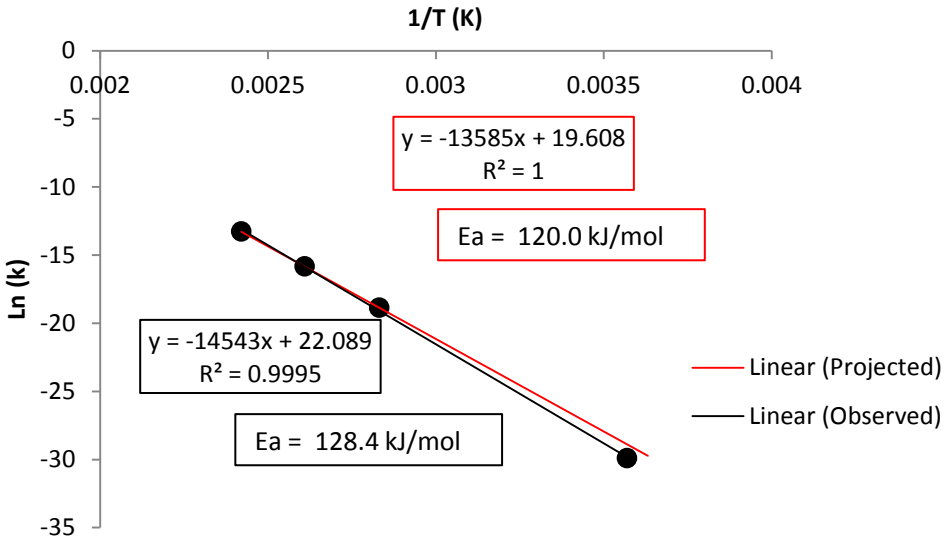

Figure 5.23: Correspondence of activation energy for valine racemisation derived using RFOK modelling of high temperature data with that derived from archaeological material. The age of the archaeological samples is derived as the average of the earliest and latest possible dates; when the end points are used, $R^2 \geq 0.9994$ in all cases and activation energies differ by less than 1.6 kJ/mol between the earliest and latest possible dates.

5.3.3.6.5: Conclusions on archaeological AAR kinetics

These results must be considered preliminary. The archaeological data available represent a very limited diagenetic range. Before extrapolation of the high temperature activation energies to low temperatures can be considered, the correspondence between high and low-temperature data observed for hydrolysis would need to be shown to be maintained across a wider range at low temperature. The successful development of a tool for identification of archaeological eggshell (Chapters 2-4) will aid the process of testing this. It cannot yet be claimed that these data provide a reliable means of estimating

reaction rates at low temperatures; far more analyses of eggshell assemblages from a range of locations and deposition temperatures is required. Future analyses should attempt to source older chicken eggshell in order to extend the temporal range of this part of the analysis, seek material from a wider range of deposition environments, and also extend the analysis to other avian taxa.

Section 4: Further investigations of intra-crystalline diagenesis

5.4.1: Overview

The previous section described the derivation of Arrhenius parameters for diagenesis in eggshell proteins and amino acids. This section will begin to explore the relationship between diagenesis and various levels of protein structure; it will also describe the sequencing of eggshell peptides; and the application of kinetic experiments on these to an attempt to compare patterns of diagenesis at high and low temperature. This will determine whether it is possible to model the expected survival time of eggshell peptides in archaeological samples from high temperature data.

5.4.2: Methods

5.4.2.1.1: Peptide sequencing: overview

Samples of modern chicken eggshell were prepared and analysed according to the optimised protocol described in a previous chapter (Section 2.2.3.4). Following initial analysis, tandem MS (MS/MS) (in which a high-intensity laser beam is used to induce peptide fragmentation) was used to derive sequence information for specific peptides (e.g. Steen & Mann, 2004). This study uses MALDI-TOF/TOF: target ions are isolated in the first time-of-flight detector using a velocity filter and fragmented by high-intensity exposure to the laser beam, before the specific trajectory of resulting 'product' ions is measured in the second TOF detector (Medzihradszky et al., 2000; Bienvenut et al., 2002; Yergey et al., 2002; Steen & Mann, 2004; de Hoffmann & Stroobart, 2007).

5.4.2.1.2: How MS/MS works

In positive ion MALDI mass spectrometry, charge is attached to target peptides by ionisation in the presence of excess protons. The addition of energy by exposure to the laser beam allows this charge to be transferred intra-molecularly from basic R groups (arginine, lysine, or histidine) to peptide backbone heteroatoms (Wysocki et al., 2000). In most peptides, the point of the peptide chain at which cleavage occurs during MS/MS depends upon the position this extra charge occupies (Wysocki et al., 2000). Cleavage is usually induced by the added proton becoming attached to the nitrogen atoms of peptide bonds, which are then weakened and dissociate (e.g. Papayannopolous 1995; Polce et al., 2000). Some post-translational modifications can affect the charge distribution within a peptide and thus affect cleavage patterns: for example, oxidation of methionine can cause a reduction in charge at the main protonation sites in some peptides, and facilitate migration of protons to amide nitrogens, leading to more efficient cleavage (Sun et al., 2010; Zong et al., 2010). A range of effects are also observed when multiple charges are attached to peptides (Wysocki et al., 2000): as this study targets only singly charged peptides, these are not discussed here.

These fragmentation pathways produce ion series in which each ion differs by the mass of one amino acid from its predecessor, as any one of the bonds within a peptide may potentially be disrupted (Wysocki et al., 2000). In reality, bonds between different amino acids vary widely in their susceptibility to fragmentation, and the factors governing this are incompletely understood. For example, enhanced cleavage is very common at the N-terminal side of proline,

aspartic acid, and at histidine when its side chain is protonated, while non-protonated histidine is particularly resistant to fragmentation (Wysocki et al., 2000). This can affect the fragmentation series which are actually observed. The most common ion series produced result from cleavage of the peptide bond; these are referred to as y-ions if the charge is retained at the C-terminus of the novel peptide (as it usually is in MALDI MS/MS), and b-ions if the charge is retained at the N-terminus (Papayannopoulos, 1995; Yergey et al., 2002; Steen and Mann, 2004). The y- and b-ion series are used to reconstruct the sequence of the peptide under analysis (see below). A range of other ion series can be produced by loss of particular atoms from y- and b-ions, by fragmentation within the side chains of some amino acids, and by loss of water and/or ammonia (Papayannopoulos, 1995). These can be useful as a cross-checking mechanism when reconstructing ion series by *de novo* sequencing (see below). Where fragmentation occurs on both sides of an amino acid, immonium ions with characteristic masses are produced. These often appear at the low-mass end of the mass spectrum, and provide some compositional information (although they provide no sequence information) (Papayannopoulos, 1995).

5.4.2.1.3: Peptide identification

Sequence data were derived from fragmentation spectra using a combination of MASCOT screening (Perkins et al., 1999) against the NCBInr and SwissProt protein databases (Table 5.15) and *de novo* sequencing (Table 5.16). Where Mascot and *de novo* sequences did not match, candidate sequences were screened against the SwissProt database using BLAST (Basic Local Alignment Search Tool) (Gasteiger et al., 2003). Most disagreements between Mascot and

de novo sequencing arose from differences in sequence (not in composition) at the N-terminal end of the peptide. This may reflect errors in transcription of the genomic data.

Parameter	Setting
Database	NCBInr/Swissprot
Enzyme	Trypsin
Global Modifications	None
Variable Modifications	Oxidation (H, W, M); Phosphorylation (S,T,Y); deamidation (N, Q)
Mass Tolerance (MS)	250 ppm
Mass Tolerance (MS/MS)	0.5 Da
Charge State	1+
Mass	Monoisotopic

Table 5.15: MASCOT search settings used in this study.

De novo sequencing, in which peptide sequence is reconstructed manually from the y- and b-ion series, can provide sequences for peptides which are not in current databases (e.g. Standing, 2003). As tryptic peptides terminate either in lysine or arginine (Olsen et al., 2004), the mass of the first y-ion (147 or 175) is diagnostic of the C-terminal residue. The y-ion series ascends from the low-mass range of the spectrum, while the b-ion series descends from the high-mass range (Table 5.16). *De novo* sequencing can provide a list of candidate sequences which can be screened against databases in order to cross-check the reliability of database mining, or can provide sequence information for previously unreported peptides.

y-ion *m/z*	Mass (Da)	Amino Acid	b-ion *m/z*	
175.1	175.1	R	n/a	C-terminal
246.1	71	A	752.5	
343.2	97.1	P	655.4	
414.2	71	A	584.4	
513.2	99	V	485.4	
642.3	129.1	E	356.3	
770.4	128.1	Q	228.2	
898.5	128.1	Q	n/a	
997.6	99.1	V	n/a	N-terminal

Table 5.16: Reconstruction of chicken eggshell peptide (m/z = 997.55) by de novo sequencing. In this case, it has been possible to construct the whole sequence using only the y-ion series; often both ion series are required, particularly if the high-mass end of the spectrum is poor. Although the b-ion series is incomplete (as it almost always is using these methods), the ions which are present are given.

5.4.2.2: Kinetic experiments

The kinetic experiments exploited the chicken eggshell powders heated at 140°C for derivation of Arrhenius parameters for diagenetic processes in eggshell proteins and amino acids (Section 5.3.2.1). Subsamples of these were prepared for mass spectrometry and analysed using the optimised protocol (Section 2.2.3.4).

5.4.3: Results

5.4.3.1: Sequenced eggshell peptides

15 peptides representing four different proteins were isolated and identified from the intra-crystalline fraction of chicken eggshell (Table 5.17). These

peptides represent a low proportion of the total sequence of these proteins (Table 5.18). A combined total of 17 peptides were sequenced from duck, goose, turkey and ostrich eggshell; as these are not used in subsequent analysis, they are contained in an appendix (Appendix B).

Peak m/z	Sequence	Protein
997.5	VQQEVAPAR	Ovocleidin-116
1042.5	EAFVPPVQR	Clusterin
1150.6	WGPGSHLAAVR	Ovocleidin-17
1225.6	ARVQQEVAPAR	Ovocleidin-116
1293.6	GVVGGMVVPEGHR	Ovocleidin-116
1309.6	GVVGG**M**VVPEGHR	Ovocleidin-116
1345.7	HIATNAVLFFGR	Ovalbumin
1346.7	HIAT**D**AVLFFGR	Ovalbumin
1348.7	VWIGLHRPAGSR	Ovocleidin-17
1687.8	GGLEPINFQTAADQAR	Ovalbumin
1688.8	GGLEPI**D**FQTAADQAR	Ovalbumin
1734.9	LGQAARPEVAPAPSTGGR	Ovocleidin-116
1774.9	ISQAVHAAHAEI**D**EAGR	Ovalbumin
1850.9	ELHPFLQHPVHGFHR	Clusterin
1858.9	ELINSWVESQTNGIIR	Ovalbumin
1859.9	ELINSWVESQT**D**GIIR	Ovalbumin
1872.8	DPWVWGSAHPQAQHTR	Ovocleidin-116
2137.0	AQQEVAPVPSMVVETVAPER	Ovocleidin-116

*Table 5.17: Peptides sequenced from the intra-crystalline fraction of chicken eggshell, given in single letter amino acid code. Residues which have undergone deamidation or oxidation are underlined and in bold text (e.g. **N**).*

Protein	Family	% sequenced
Clusterin	Clusterin	5.4
Ovocleidin-17	CTL	16.2
Ovocleidin-116	SCPP	11.2
Ovalbumin	Serpin	11.4

Table 5.18: Coverage of intra-crystalline proteins sequenced. CTL = C-type lectin; SCPP = secretory calcium-binding phosphoprotein.

A number of eggshell proteins of known sequence were not recovered from even the unheated samples. In most cases, it is difficult to explain this by recourse to their primary structure. Although the distribution of lysine and arginine governs the number of tryptic peptides within the mass range used (e.g. Olsen et al., 2004), virtual tryptic digestion of the unrepresented proteins (PeptideMass, www.expasy.org) does not predict lower representation of unrecovered proteins than of recovered proteins (independent samples T-test: $p = 0.75$) (Table 5.19). In some cases (ovotransferrin and serum albumin) the predicted number of tryptic peptides in the mass range used is very high (Table 5.19); their non-recovery is sufficient evidence to claim that these proteins are probably not present in the intra-crystalline pool. In others (lysozyme, ovocalyxin-32/36 and osteopontin), there are too few predicted tryptic peptides (Table 5.19) to rule out effects of primary structure on successful peptide recovery and sequencing. Based on these data, it is possible that these proteins are present in the intra-crystalline fraction but remain undetected by the techniques used here.

Protein	Length (AAs)	Predicted peptides	Recovered
Ovotransferrin	705	49	0
Serum Albumin	615	49	0
Ovalbumin	386	25	4
OC-116	743	24	7
Clusterin	448	24	2
OCX-36	459	19	0
OCX-32	275	8	0
Osteopontin	264	8	0
Lysozyme	147	7	0
OC-17	142	6	2

Table 5.19: Predicted number of tryptic peptides for eggshell proteins of known sequence. The predicted number represents a conservative minimum value (no missed cleavages allowed). In cases where multiple variants of a single peptide were recovered, these were grouped as one peptide.

5.4.3.2.1: A model for quantifying relative protein contributions

A mixing model was constructed to attempt to model the relative contribution of these proteins to the total pool. As mentioned above, this did not prove successful. Despite this, the model used might prove useful in future if more accurate input data can be obtained (see below); for this reason, it is presented despite being unsuccessful here.

This model uses the relative contribution of different amino acids to the total intra-crystalline pool in fresh chicken eggshell (Section 5.3.3.1) as the starting data input. This provides the first of two major sources of error; the RP-HPLC procedure used (Section 2.1.2.5) resolves only 16 of the 20 amino acids present

in proteins, and some of these poorly and/or only in one chiral form. Development of techniques which can accurately resolve a higher number of amino acids may be a pre-requisite for the successful use of a model such as this. The relative contribution of each of these amino acids to the primary sequence of the four identified proteins was then quantified. In order to minimise the error resulting from the missing and poorly resolved amino acids, these were deducted from the total number of residues in the sequence prior to calculation of relative contributions (inclusion of the poorly resolved residues had no significant effect on the observed patterns). The model was then constructed to apply the multiplier to the data for each protein which gives the best linear fit between the observed contribution of each amino acid to the whole system and the calculated contribution from each protein. Maximising the correlation score between these parameters should indicate the relative contribution of each protein to the total pool. The linear fit is not good; the maximum correlation score achieved ($c = 0.91$, $R^2 = 0.83$) is inadequate, and the model wrongly suggests that ovalbumin is not present (Tables 5.18 & 5.19). The fit is slightly improved if other eggshell proteins (OCX-32, OCX-36 and osteopontin) are included ($c = 0.93$, $R^2 = 0.86$), but this does not resolve the problem of ovalbumin being wrongly listed as absent (Table 5.20).

Protein	Relative Contribution (%)
OC-17	3.9
OC-116	76.3
Ovalbumin	0.0
Clusterin	6.2
OCX-32	0.0
OCX-36	13.5
Osteopontin	0.0

Table 5.20: Relative contributions of intra-crystalline proteins estimated using mixing model. Although these data are not accurate, the technique may be useful given more accurate data to work from.

The model suggests that OC-116 dominates the intra-crystalline pool; this is supported to some extent by observed protein concentrations in eggshell, where this is by far the most abundant protein (Hincke et al., 1999). The limited data available for concentrations of other proteins suggest that OC-17 should be more highly represented, as the concentration of this protein is roughly half that of OC-116 in the combined inter- and intra-crystalline pools (Mann, 1999). Given the catalytic role played by OC-17 (Freeman et al., 2010), it might not be found at high concentrations in the intra-crystalline fraction. Intra-crystalline peptides have been recovered from ovalbumin, but the model lists this protein as absent; this shows that the estimation of relative contribution is inaccurate.

Given more accurate input data to work from, there is hope for this method to provide a better estimate of the relative contribution of different proteins. The failure of this model probably relates to two major factors; inability to resolve all

amino acids in the system, and the likely presence of unidentified proteins in the intra-crystalline pool. This latter problem is likely to be the main cause of inaccuracy in the model here, and more sensitive characterisation of the intra-crystalline protein fraction should be an emergent focus of future research. Although this model may provide a useful template for future research if these data can be made available, this also highlights the dangers of applying mathematical models to incomplete datasets.

If these data were available, it would be possible to conduct a detailed investigation of the effects of protein primary structure on patterns of diagenesis. Unfortunately, conducting such an analysis in the absence of accurate starting data would require the unfounded assumption that the proteins in the intra-crystalline fraction are evenly distributed; this could only lead to inaccurate results. The characteristics of adjacent amino acids can affect patterns of hydrolysis, and subsequently racemisation and decomposition. Mechanisms by which this can occur include particularly hydrophobic residues forming peptide bonds which are relatively hydrolysis-resistant (the converse is true of hydrophilic residues); proximity to residues with large, bulky side chains which can act as a physical barrier to hydrolysis; and proximity to favoured points of chain scission (e.g. aspartic acid and serine) leading to more rapid diagenesis. Exploration of these factors would provide a more nuanced understanding of the diagenetic environment in eggshell proteins and its effect on observed patterns of hydrolysis and racemisation; while accurate data on the composition of the intra-crystalline fraction are lacking, this must remain a black

box. If such data become available in the future, this concept should be revisited as a matter of priority.

5.4.3.2.2: The effects of tertiary protein structure on diagenesis

Given the limitations described above, investigation of potential effects of protein 3D structure on the diagenesis of individual amino acids is not possible here. The effects of 3D structure on the relative survival of different peptides can be explored based on the available data. For example, if a particular part of a protein is buried in the core, it may be physically protected from hydrolysis. In the largely anhydrous intra-crystalline environment, water should become limiting; this may partially explain the retention of a hydrolysis-resistant pool of amino acids even after prolonged diagenesis (Section 5.3.3.3.1). It is also possible that some peptides may be protected from hydrolysis by direct interactions with the mineral phase (e.g. Freeman et al., 2011). 3D structure is available for two of the four proteins identified in the intra-crystalline fraction of chicken eggshell (OC-17 and ovalbumin) (Stein et al., 1991; Reyes-Grajeda et al., 2004). The structure of the other proteins remains unknown at present. For OC-17 and ovalbumin, it is possible to map the sections of identified sequence onto the entire protein (Figure 5.24). The potential effects of 3D conformation on diagenesis can then be considered (Section 5.4.3.4).

Figure 5.24: Sections of sequence identified in this study (shown in yellow) in ovocleidin-17 (a) and ovalbumin (b) from chicken eggshell. Image generated using CN3D.

5.4.3.3: Patterns of peptide survival during accelerated diagenesis

In the previous section, it was shown that the correspondence between high and low temperature data is reasonably good for valine and isoleucine hydrolysis (and to a lesser extent racemisation) in chicken eggshell (albeit across a minuscule diagenetic range) (Section 5.3.3.6). In order to relate peptide survival to the diagenetic processes explored in the previous section, the extent of valine and isoleucine racemisation and hydrolysis at the latest time point at which each peptide was observed at 140°C was recorded (Table 5.21).

Time (hours)	Val D/L	% Val FAA	Peptides (m/z)
0	0.02	0.23	1345.7, 1688.8, 1858.9, 1859,9, 1860.9
1	0.04	0.92	1687.8, 1775.8, 1851.0
2	0.04	1.49	997.6, 1225.6, 1872.8, 2137.0
4	0.06	2.43	1309.7, 1348.7, 1735.9
6	0.06	3.01	1346.7
8*	0.08	5.43	1042.6, 1150.6, 1293.7

*Table 5.21: The extent of valine racemisation and hydrolysis at latest survival of chicken eggshell peptides heated at 140°C. * indicates that these peptides were still observed at the end of the experiment; no peptides were recovered from 24 hour heated eggshell.*

5.4.3.4: Factors affecting peptide survival

The survival of particularly hydrolysis-resistant peptides after a significant amount of diagenesis (≥ 11% FAA) lends some support to the notion of a residual peptide-bound fraction of amino acids in chicken eggshell; this might explain the observed decrease in hydrolysis and racemisation rates during late diagenesis (Section 5.3.3.4.2) (e.g. Collins & Riley, 2000; Demarchi et al., 2013). The most robust peptides (i.e. those most resistant to hydrolysis at high temperature) (Table 5.21) share common characteristics; they are relatively short (9-13 residues), contain mostly hydrophobic residues (average = 75.7% ± 9.1%); few hydrophilic residues (6.7% ± 5.9%); and no consecutive hydrophilic residues in their primary sequence (Tables 5.16 & 5.20).

There is no clear relationship between peptide length and survival ($R^2 < 0.3$: data not shown). This is because a range of factors, including composition, sequence, location within the protein and interaction with the mineral phase may also affect peptide survival. Although there is an expected trend towards more hydrophobic peptides being relatively resistant to hydrolysis (Figure 5.25), this does not conform to a simple linear model, and nor does the converse calculation for hydrophilic (Figure 5.25) or charged residues.

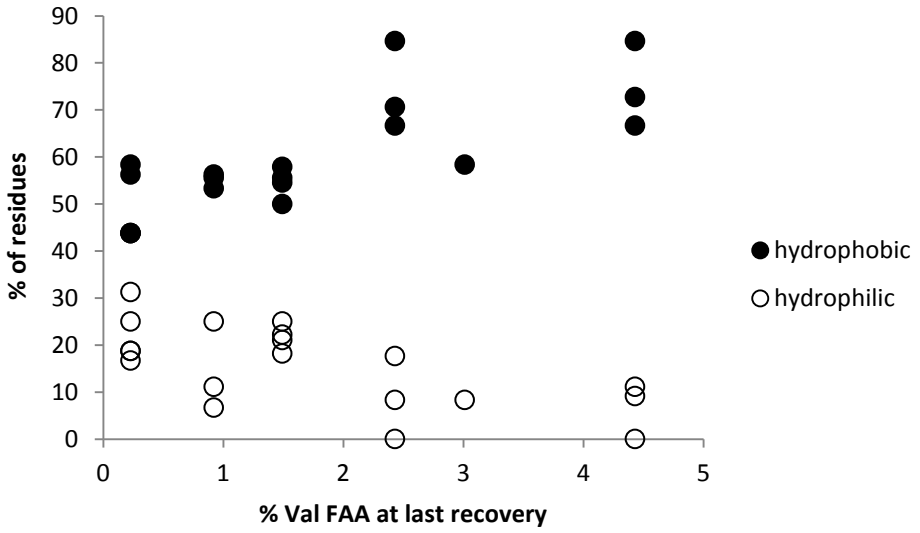

Figure 5.25: Relationship between peptide survival and hydrophobicity and hydrophilicity. Although general trends are visible, there is no clear relationship ($R^2 < 0.6$ in all cases). This agrees with previous analyses which showed that preserved amino acids in fossil materials tend to be dominated by hydrophobic residues (e.g. Goodfriend et al., 2000 and references therein).

Peptide sequence is also likely to contribute to the observed patterns of peptide survival. None of the most robust peptides exhibit consecutive hydrophilic residues, and those which do tend to be lost earlier (Figure 5.26). There is no

discernible trend in peptide survival relating to the grand average of hydropathicity (GRAVY), isoelectric point (pI) or aliphatic index (data not shown).

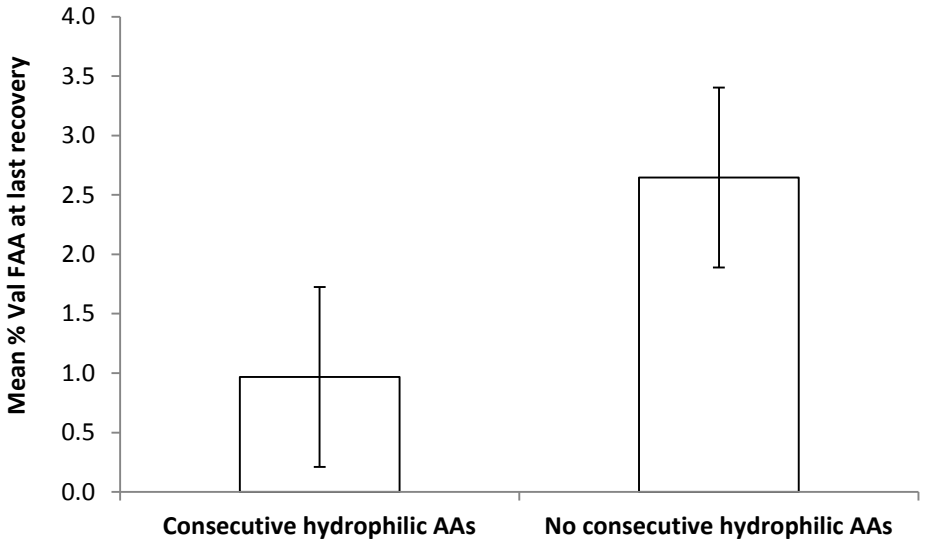

Figure 5.26: Survival of chicken eggshell peptides heated at 140°C. A difference is observed between those exhibiting (n = 10) and not exhibiting (n = 9) consecutive hydrophilic residues. Error margins represent one standard deviation from the mean.

The location of a peptide within the intact protein may also be important. For example, in the intra-crystalline fraction, water for hydrolysis may rapidly become limiting while the protein core remains largely un-hydrolysed (e.g. Walton, 1998). In intra-crystalline proteins, conditions should become anhydrous in the absence of water entering from the external environment, although water can be produced by a range of diagenetic processes (e.g. Collins et al., 1992). Peptides which interact directly with the mineral phase (e.g. Freeman et al., 2011) may also be physically protected from hydrolysis. Unfortunately, the 3D structure of ovocleidin-116 and clusterin is not available

at present, so this cannot be tested for peptides derived from these proteins. The structure of ovocleidin-17 and ovalbumin is available, and in these the most robust peptide(s) derived can be placed within the protein (Figures 5.27 & 5.28).

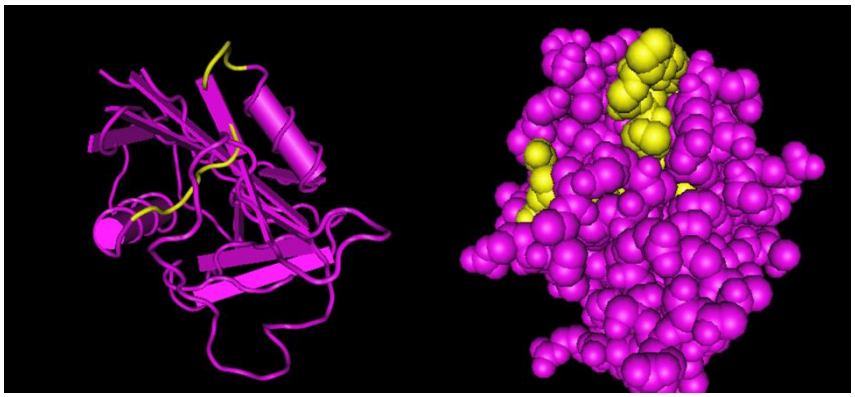

Figure 5.27: Two views of location of the most robust peptide (m/z = 1150.6) in OC-17. Image generated using CN3D.

In OC-17, the location of parts of the most robust peptide (*m/z* = 1150.6) at the outer surface of the protein (Figure 5.27) argues against physical protection within the protein being a major determinant of survival, especially as the exposed sections of the peptide contain the least hydrophobic components (one serine and two glycines) (Reyes-Grajeda et al., 2004). Only one of the six positions adjacent to these residues comprises an amino acid (tryptophan) with an aromatic side chain (these may act as a physical barrier to hydrolysis) (Reyes-Grajeda et al., 2004). Given the strong evidence that OC-17 plays a catalytic role in promoting crystal nucleation (Section 1.4), it seems unlikely that interaction with the mineral phase physically protects particular parts of the protein from hydrolysis. According to current models, OC-17 binds weakly to the mineral surface and is dislodged once a particular configuration is reached (Freeman et al., 2011). Of the three possible binding configurations observed for OC-17,

there are two in which the C-terminal arginine of the most robust peptide interacts with the mineral phase; this is the only residue in the peptide to do so (Freeman et al., 2011). This might seem to suggest that direct interaction plays a role in determining survival, but when compared to the less robust peptide sequenced from this protein (*m/z* = 1348.7, Table 5.16), this seems unlikely; the C-terminal arginine of this peptide is involved in all three identified binding configurations, and an additional serine is involved in one (Freeman et al., 2011). Despite this, and despite only being one residue longer, this peptide is lost earlier. It is likely that the hydrophobic composition and lack of consecutive hydrophilic residues in this peptide, combined with a rapidly diminishing pool of water for hydrolysis, is more important in determining peptide survival than its location in the protein or interaction with the mineral phase.

In ovalbumin, the most robust peptide (*m/z* = 1346.7 in the more robust deamidated form, 1345.7 in the non-deamidated form) is only moderately hydrophobic (58.3% of residues), and despite low representation of hydrophilic residues it is lost earlier than the peptide described above (Table 5.21).

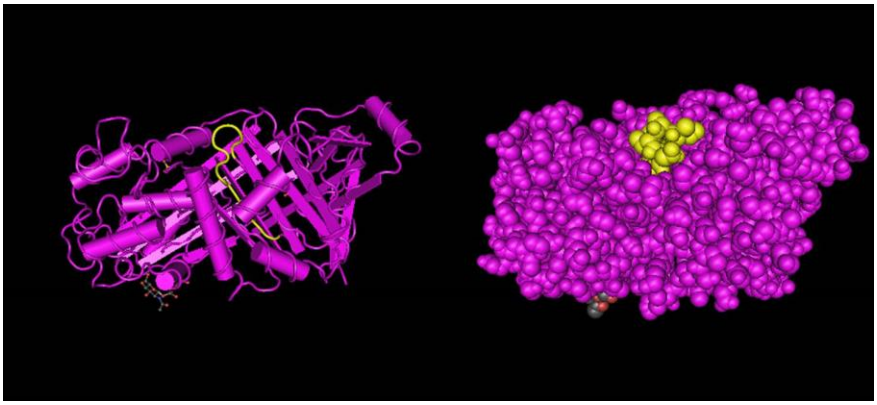

Figure 5.28: Two views of the most robust peptide (m/z =1346.7) in chicken ovalbumin. Image generated using CN3D.

There are two factors which could explain this. In the deamidated form of this peptide (which survives longer, probably due to the removal of a hydrophilic-to-hydrophilic (threonine-asparagine) peptide bond), a hydrophilic residue (threonine) and a negatively charged residue (aspartic acid) are exposed at the protein surface (Stein et al., 1991). In the non-deamidated form, both threonine and asparagine are exposed at the surface, and occur consecutively (Stein et al., 1991). This can lead to hydrolysis either by presenting an amenable site with a particularly labile peptide bond, and/or by breaking of the peptide chain at aspartic acid to form succinimides even when the residue is peptide-bound (Geiger & Clarke, 1987). The less robust ovalbumin peptides (*m/z* = 1687.8/1688.8 and 1858.9/1859.9/1860.9) form a contiguous sequence which is entirely exposed at the external surface, and which exhibits a relatively high proportion of hydrophilic residues (some of which are consecutive). All of this points to a combination of peptide composition and location being important in determining peptide survival; the interaction between these factors is unclear, and is sure to differ between proteins and between different sections of a protein. While interaction with the mineral phase is theoretically important in some cases, there is little evidence to support that scenario, at least for the peptides recovered here.

5.4.3.5: Testing the correspondence between high and low temperature peptide survival

A reasonably tight correspondence was observed between the activation energies of valine and isoleucine hydrolysis in the high and low-temperature experiments (Section 5.3.3.6.3). Although this has been demonstrated for only a very limited diagenetic range, it allows the expected peptide survival time to be roughly modelled from the kinetic data across the range observed. This is calculated using the regression equations derived for valine and isoleucine hydrolysis in a previous section (Figures 5.21 & 5.22). The value thus derived for reaction rate (k) can then be inserted into the IFOK equation (equation 1) in order to provide an estimate of the expected reaction rate at a given temperature (Table 5.22). Given the limited diagenetic range across which correspondence between high and low temperature data has been demonstrated, the error margins attached to these models will be very large, and it is not claimed that these represent even moderately accurate estimates of the temporal span of the eggshell identification technique. Despite this, they provide a means of testing whether the patterns of recovery observed at high temperature are replicated at low temperature; this can be used to provide a proxy indication of whether the pathways of hydrolysis are consistent at different temperatures.

Deposition temperature (^{o}C)	Projected survival time (years)	
	Valine	Isoleucine
7	8828.2	7755.3
10	5282.3	4626.5
15	2298.3	2003.2
20	1028.7	892.5
25	473.1	408.6

Table 5.22: Expected survival time of most robust chicken eggshell peptides estimated using the regression equations derived for valine and isoleucine hydrolysis. The error margins attached to these estimates are likely to be high, but cannot be quantified at present.

Direct archaeological evidence to test the predictions of these models is generally lacking at present, as the technique for taxonomic identification of eggshell was only developed as part of this study. Some preliminary tests are possible, and even these suggest that the observed correspondence between the relationship between heating temperature and reaction rate at different temperatures is misleading. The results obtained from the analysis of chicken eggshell from Pompeii (≥ 1934 years at an estimated deposition temperature of 16°C – www.thermal-age.eu) suggest that these models do not accurately predict rates (or even relative rates) of peptide survival at low temperatures (Table 5.23). A range of peptides were recovered from the material which the model predicts should be absent; this suggests that patterns of hydrolysis are different at low temperature, and that attempts to model expected survival time of peptides based on high temperature data are inaccurate. This has negative implications for the potential use of high temperature data to estimate peptides survival at lower temperatures and of intra-crystalline protein diagenesis in eggshell as an absolute dating tool.

Peak (m/z)	Predicted	Observed	% of samples	Mean S/N
997	Absent	Present	4.5	31.5
1042.6	Present	Present	100	90.4
1150	Present	Present	69.5	24.4
1224	Absent	Present	0.9	17.7
1293/1309	Present	Present	23.0	34.8
1345	Absent	Present	43.8	26.8
1346.7	Absent	Present	68.6	36.0
1348.7	Absent	Present	6.0	14.7
1687	Absent	Present	24.5	16.8
1688	Absent	Present	43.5	36.6
1735	Absent	Present	14.2	17.7
1774	Absent	Present	55.0	18.0
1851	Absent	Absent	0	0.0
1872	Absent	Present	0.3	7.6
2137	Absent	Present	2.7	13.9

Table 5.23: Correspondence between predicted recovery of chicken eggshell peptides estimated from high temperature data and observed recovery from archaeological samples. The archaeological samples used were from Pompeii. These data suggest that patterns of hydrolysis are temperature dependent.

5.4.4: Conclusions on this section and chapter

These data (Table 5.23) suggest that despite the apparent correspondence derived using FOK models, patterns of hydrolysis are in fact temperature dependent. If the intra-crystalline fraction is largely anhydrous, differences observed between high and low temperature diagenesis may relate (at least in part) to the temperature dependence of diagenetic reactions which can generate water for hydrolysis (e.g. decomposition or condensation reactions between sugars and proteins), and of alternative mechanisms of chain scission (Steinberg & Bada, 1981; Geiger & Clarke, 1987; Collins et al., 1992; Walton, 1998). Although quantifying these is beyond the scope of this project, it has

been made clear here that a full understanding of all contributing diagenetic reactions is required in order to accurately model low temperature diagenesis from high temperature data, even in cases where the correspondence appears to be good.

These data also demonstrate that the high and low temperature data are not likely to correspond closely when archaeological material representing a more extensive diagenetic range becomes available. This implies that, at least using FOK models, avian eggshell does not provide a suitable material for absolute AAR dating. Although it has not been possible to positively use the data presented in this chapter to relate observed patterns of diagenesis to the archaeological data, there are some promising observations. For one, the recovery of a range of peptides from the Pompeii material bodes well for the temporal span of the eggshell identification technique which forms the core aim of this project (although attempts to model this directly have been thwarted). In theory, if peptides are recovered from nearly 2000 year old material at Pompeii's deposition temperature, the technique should be viable for a number of millennia at British and northern European deposition temperatures. Although not (strictly speaking) a positive result, the question of whether avian eggshell can provide a suitable material for amino acid dating and/or palaeothermometry (at least in an absolute sense) has begun to be answered. This fulfils one of the stated aims of this project; to determine the usefulness of eggshell proteins in archaeological research. The data presented in this study suggest that these are very useful for taxonomic identification, and may become

useful for relative dating or externally calibrated dating, but not for absolute dating and palaeothermometry.

Chapter 6: Summary & Future Directions

Section 1: Summary of results

6.1.1: Overview

This section will present a brief summary of the objectives of this project, and will describe the extent to which these were met. Where these were not met, priorities for future research are described in the following section.

6.1.2.1: Production of an identification system for archaeological eggshell

The primary objective of this study was to produce a taxonomic identification system for archaeological eggshell fragments. This has been conducted successfully; the technique developed in this study (Chapters 2-3) is fast, accurate, and (unlike previously existing techniques) can cope with the large volume of material often recovered from archaeological contexts. This allows whole assemblages to be characterised; this removal of the need for sub-sampling removes subjectivity from the process. It is no longer necessary to assume that a sub-sample is representative of the entire assemblage; this has facilitated more accurate quantification of eggshell data. In addition to the development of PMF as an identification system, novel software has been developed in order to screen the reference database. This facilitated the application of the new technique to archaeological material (described below). The level of resolution achieved by the new technique is variable, but all taxa studied to date have been identified to at least the level of family (Chapter 2).

6.1.2.2: Applying the new technique to the archaeological record

Two case studies were conducted in order to test the viability of the new technique. The first analysed Anglo-Scandinavian eggshell assemblages from York, and highlighted the lack of wild taxa present; this was in marked contrast to the avian bone assemblages, which contained a wide range of wild taxa (Chapter 4). While the lack of wild eggshell may not be particularly surprising in an urban setting such as York, it is more surprising that it was absent from the majority of the island sites considered in this study (Chapter 4). A wide range of factors may explain this observation, but it seems likely that one major factor is the Norse system of land ownership and legal access to different types of resource. Although a wide range of questions remain open following these analyses (Chapter 4), the production of the new technique has allowed these questions to be asked, and in the future it may allow them to be answered.

6.1.2.3: Characterising intra-crystalline diagenesis in avian eggshell

Another aim of this project was to establish how useful eggshell proteins can be in archaeology. Intra-crystalline protein diagenesis (IcPD) has been studied in a range of biominerals in order to establish their usefulness for dating and/or palaeothermometry (Chapter 5). Arrhenius parameters were obtained for two major diagenetic processes (hydrolysis and racemisation) in the eggshell of three avian species, and potential effects of protein structure and conformation on observed patterns of diagenesis were explored (Chapter 5). Although the potential applicability of IcPD in avian eggshell to relative or calibrated dating remains to be fully explored, it has been shown here that it is not useful as an absolute dating method in the taxa analysed here (Chapter 5). This contributes

to a better understanding of the usefulness of eggshell in archaeology and geochronology.

6.1.2.4: Using intra-crystalline diagenesis to model the expected temporal span of the technique

The main application of IcPD in this study was an attempt to model the expected temporal span of the identification technique, and to establish the relationship between this and deposition temperature. This was not successful, as it proved impossible to accurately model low-temperature diagenesis using high-temperature data (Chapter 5). This led to a major discrepancy between expected and observed rates of peptide survival (Chapter 5). Given the failure of this approach (and the data are unequivocal; this approach does not work), it is not yet clear what the expected temporal range of the technique is; nor is it clear how this relates to deposition temperature. At present, the temporal range of the technique can only be estimated on a trial-and-error basis (Section 6.2.2.4).

Section 2: Emergent research priorities

6.2.1: Overview

This study has demonstrated the development and application of a novel technique for identification of archaeological eggshell fragments. Although this technique has been shown to be viable across a significant temporal span, there are still a range of improvements which can be made in future. This section will highlight the major research foci which emerge from this project.

6.2.2.1: Improving the reference collection

One of the main areas for potential future refinement is in the reference collection which forms the basis of the identification technique. In this study, the reference collection was drawn mostly from the YMT egg collections, with small contributions from other museums and individuals (Chapter 3). Due to the novel nature of this project, construction of this collection had to begin from scratch (Chapter 3); this was a lengthy and time-consuming process. Now that the efficacy of the technique has been demonstrated, and the research potential of museum egg collections has been demonstrated, it should become easier to source material to expand the reference collection beyond its current state (58 species in 13 orders). As the reference collection is largely geared towards British and northern European taxa, improvement of the coverage of certain taxa may be a pre-requisite for application to other regions. This will also facilitate better understanding of the level of taxonomic variability in eggshell proteins. All of the data pertaining to the reference collection, as well as the software developed for interrogation of it and detailed instructions for

installation and use, will be made publicly available online upon completion of this project.

6.2.2.2: Testing the correspondence of different identification techniques

One major area which remains to be explored is the correspondence between the pre-existing SEM identification technique (Sidell, 1993) and the technique developed in this study. It seems probable that the former has overestimated the diversity of the eggshell assemblage at Freswick Links (Sidell, 1995); an unsuccessful attempt was made to source this material in order to compare the methods. Now that the efficacy of the new ZooMS technique has been demonstrated, it might become possible to secure this material for analysis; this would allow testing of the correspondence between the techniques. Given the recovery of various taxa from Iron Age and Norse contexts at Bornais (Chapter 4), it seems unlikely that the technique developed in this study underestimates taxonomic diversity in eggshell assemblages. Of course, as described in a previous chapter, it is entirely possible that the Freswick Links assemblage is genuinely different from those considered in this study (Chapter 4). It would be beneficial for the study of archaeological eggshell if the two techniques correspond well, as both have relative advantages; the SEM technique is potentially capable of a higher degree of resolution in some cases (Sidell, 1993), while the ZooMS technique can handle a far greater volume of material (Stewart et al., 2013).

6.2.2.3: Expanding the geographical range of the analysis

In this project, the new identification technique was applied predominantly to Anglo-Scandinavian and Norse-era sites in York and in the Scottish islands (Chapter 4). Another area of priority is expansion of the application of the new technique across a wider geographical area. Considering other Norse and Norse-influenced areas will allow variability in Norse subsistence strategies across their geographical sphere of influence to be tested, and may shed light on the nature of changes in these over time. This may also begin to answer some of the questions regarding Norse-era use of wild eggs arising from this study (Chapter 4).

A range of other assemblages have been successfully tested, but are either not considered here or are not considered in detail. These include a relatively large assemblage from Pompeii (J. Thompson, pers comm), and smaller assemblages from Roman age sites in Austria (E. Gal, pers comm), classical Greece (D. Mylona, pers comm), Iron Age (5^{th} – 9^{th} century A.D.) sites in Hungary and Austria (B. Tugya, pers comm), Iron Age grave sites in Germany (C. Schuh, pers comm), and single fragments of shell from Bronze/Iron Age sites in Saudi Arabia and Ethiopia (A. Prust, pers comm). The successful analysis of all of this material begins to show the potential of using this method on archaeological investigation to provide information about past subsistence across a wider geographical (and temporal) range.

6.2.2.4: Establishing the temporal range of the new technique

The exact temporal limits of the technique remain to be established, but it has been shown that use of high temperature experiments to mimic natural diagenesis is not an accurate means of establishing the relationship between peptide survival, deposition temperature and time (Chapter 5). The failure of the application of the high temperature data to model peptide recovery at lower deposition temperatures with any degree of accuracy means that, for now at least, establishment of temporal limits has been conducted on a trial-and-error basis. The oldest material successfully identified so far is a fragment of ostrich shell from Taywa, Saudi Arabia (8^{th} – 3^{rd} century B.C.; A. Prust, pers comm); peptides identified by this technique have been observed to survive in ostrich shell for up to 15000 years at slightly lower deposition temperatures in southern Africa (B. Demarchi, pers comm). The oldest non-ratite shell yet analysed was chicken eggshell from Kalaeuria in Greece (4^{th} century B.C.; D. Mylona, pers comm). Although the temporal limits of the technique have not yet been established, unsuccessful analysis of 50000 year old eggshell from southern France suggests that truly ancient timescales may not be accessible using this technique (V. Laroulandie, pers comm).

Another angle which can be more fully explored in future is establishment of the relationship between high and low temperature diagenesis. More substantial assemblages of known species will be required to facilitate this; these can now be targeted. In particular, older chicken eggshell assemblages are a priority, as are those from a wider range of deposition environments. The only potentially useful assemblage analysed so far was from Pompeii, but there was not enough

material from any single context to allow testing of diagenesis (in many cases, a context at this site would be represented by one or two minuscule fragments of eggshell).

6.2.2.5: Integration of bone and eggshell evidence

In order to provide a better understanding of the manner in which people in the past utilised birds, closer investigation of the relationship between the zooarchaeological bone data and the eggshell data is required, particularly for the major domestic species and seabirds. In the past, it had been cautiously presumed in many cases that patterns of egg use would at least partially track patterns in use of the adult birds (as represented in the bone remains) (e.g. Graham-Campbell & Batey, 1998; Sharples, 2005; Harland et al., 2012). Close investigation of the Norse-era eggshell assemblages targeted in this study has shown that this is not the case, and a direct link between use of adult birds and their eggs cannot be presumed (Chapter 4). In the future, the new eggshell technique can facilitate better integration of the bone and eggshell data; this will allow the main functions of avian taxa at different points in history to be investigated in more detail.

6.2.2.6: More accurate characterisation of the intra-crystalline protein fraction

In this study, incomplete characterisation of the intra-crystalline protein fraction is one of the major barriers to investigation of the effects of protein structure on observed patterns of diagenesis. As a result, the effects of primary structure on diagenesis largely remain a subject of speculation. If more sensitive techniques

can characterise the intra-crystalline protein pool, it should become much easier to accurately model diagenesis in this system (or systems if the pool is heterogeneous). The 3D structure of the proteins involved also helps in this process. For chicken, some data is available, but characterisation of the intra-crystalline pool is at a very early stage; in other (non-ratite) species, this work has barely begun. More detailed and more widespread characterisation of eggshell proteins can only improve understanding of diagenesis in future.

6.2.2.7: Establishment of Arrhenius parameters for other diagenetic reactions

One of the major limitations of the study of diagenesis conducted here is the focus on only two reactions (hydrolysis and racemisation). As detailed in a previous section, a range of other processes occur during diagenesis, including decomposition, DKP and succinimide formation, and various minor processes which may release water into the intra-crystalline environment (Chapter 5). These latter, which include processes such as Maillard reactions between proteins and carbohydrates (Collins et al., 1992), may be particularly crucial to explaining the observed difference between high and low temperature peptide recovery across similar diagenetic ranges (Chapter 5). A far closer examination of the whole diagenetic process in eggshell is required before the data for any of the reactions studied can be accurately used. Application of this line of research to a greater diversity of avian species will also highlight differential patterns of diagenesis in different species. These are certain to occur due to differences between local molecular environments, and will dictate that the temporal limits of peptide recovery vary between taxa.

Appendix A – Species represented in the reference collection

Order	Species	Common Name	Species	Common Name
Accipitriformes	*P. unicinctus*	Harris's Hawk		
Anseriformes	*A. platyrhynchos*	Mallard/Domestic Duck	*B. canadensis*	Canada Goose
	S. mollissima	Eider	*C. olor*	Mute Swan
	M. merganser	Goosander	*B. bernicla*	Brant Goose
	B. leucopsis	Barnacle Goose	*A. anser*	Greylag/Domestic Goose
	B. clangula	Goldeneye	*A. ferina*	Pochard
	A. acuta	Pintail		
Caprimulgiformes	*C. europaeus*	Eurasian Nightjar		
Charadriiformes	*F. arctica*	Puffin	*S. hirundo*	Common Tern
	L. canus	Common Gull	*V. vanellus*	Lapwing
	R. tridactyla	Kittiwake	*G. gallinago*	Snipe
	S. paradisaea	Arctic Tern	*N. arquata*	Curlew
	U. aalge	Guillemot	*L. hyperboreus*	Glaucous Gull
	L. argentatus	Herring Gull	*L. fuscus*	Lesser Black-backed Gull
	H. ostralegus	Oystercatcher	*L. marinus*	Great Black-backed Gull
	A. torda	Razorbill	*L. minutus*	Little Gull
	C. ridibundus	Black-headed Gull	*S. rusticola*	Woodcock
	C. grille	Black Guillemot		
Ciconiiformes	*A.cinerea*	Grey Heron		
Columbiformes	*C. palumbus*	Wood Pigeon		
	C. livia	Rock Pigeon		
Falconiformes	*F. tinnunculus*	Kestrel		
Galliformes	*M. gallopavo*	Domestic Turkey	*P. colchicus*	Pheasant
	G. gallus	Domestic Chicken	*L. l. scotica*	Red Grouse
	C. coturnix	Common Quail	*P. perdix*	Grey Partridge
	P. cristatus	Peafowl		
Gruiformes	*F. atra*	Coot		
	G. chloropus	Moorhen		
Passeriformes	*T. merula*	Blackbird	*C. monedula*	Jackdaw
	C. frugilegus	Rook	*P. pica*	Magpie
	H. rustica	Swallow	*C. corone*	Carrion Crow
	T. guttata	Zebra Finch		
Pelecaniformes	*M. bassanus*	Gannet		
	P. carbo	Cormorant		
	P. aristotelis	Shag		
Procelariiformes	*P. puffinus*	Manx Shearwater		
	F. glacialis	Fulmar Petrel		
Struthioniformes	*S. camelus*	Ostrich		

All taxonomic distinctions follow Hackett et al (2008).

Appendix B – Peptides sequenced from other avian species

Species	Peak *m/z*	Sequence	Protein
Turkey	1418.8	HNLTNSILFFGR	Ovalbumin
Turkey	1677.8	GGLESINFQTAADQAR	Ovalbumin
Turkey	1848.9	ISQAVHAAYAEIYEAGR	Ovalbumin
Duck	1366.6	YSAWDDDELPR	Ansocalcin
Goose	1372.6	YSAWDDDE_	Ansocalcin
Duck	1382.6	YSA**W**DDDELPR	Ansocalcin
Goose	1722.7	_YGYFGQQLTWR	Ansocalcin
Duck	1723.7	_ YGYFGQELTWR	Ansocalcin
Duck	2132.98	EEEDNVWIGLHHWNQAR	Ansocalcin
Duck	2148.97	EEEDNV**W**IGLHHWNAQR	Ansocalcin
Duck	2164.96	EEEDNV**W**IGLHH**W**NAQR	Ansocalcin
Ostrich	962.5	G**W**IPFDGR	Struthiocalcin-2
Ostrich	991.5	YELPWKR	Struthiocalcin-1
Ostrich	1007.5	YELPWKR	Struthiocalcin-1
Ostrich	1059.9	LDYGSWYR	Struthiocalcin-2
Ostrich	1615.8	AGAHLASIHTSEEHR	Struthiocalcin-1
Ostrich	2491.2	FISQYHHGEEEEDVWIGLFR	Struthiocalcin-1

Peptides sequenced from other species (ostrich, turkey, domestic duck and domestic/greylag goose). _ indicates unidentified sections of sequence. Amino acids which are underlined and bold text have undergone post-translational modification

References

Abelson, P., 1957. Organic constituents of fossils. *Yearbook of the Carnegie Institute of Washington*, 53, pp.97-101.

Addadi, L., Raz, S. & Weiner, S., 2003. Taking advantage of disorder: amorphous calcium carbonate and its roles in biomineralization. *Advanced Materials*, 15(12), pp.959 - 970.

Aizenberg, J. et al., 1996. Stabilization of amorphous calcium carbonate by specialized macromolecules in biological and synthetic precipitates . *Advanced Materials*, 8(3), pp.222-226.

Aizenberg, J. et al., 2002. Factors involved in the formation of amorphous and crystalline calcium carbonate: a study of an ascidian skeleton. *Journal of the American Chemistry Society*, 124(1), pp.32 - 39.

Albarella, U. & Thomas, R., 2002. They Dined on Crane: Bird Consumption, Wild Fowling and Status in Medieval England. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.23-38.

Albeck, S. et al., 1993. Interactions of various skeletal intracrystalline components with calcite crystals. *Journal of the American Chemical Society*, 115, pp.11691-11697.

Allison, E.P., 1995. Bird Bone. In C. D. Morris, C. E. Batey, & D. J. Rackham, eds. *Freswick Links, Caithness: Excavation and survey of a Norse settlement*. Inverness: Highland Libraries, pp. 209-211

Allwood, A.C. et al., 2009. Controls on development and diversity of Early Archean stromatolites. *Proceedings of the National Academy of Sciences,* 106(24), pp.9548-9555.

Arias, J.L. et al., 1991. Collagens of the chicken eggshell membranes. *Connective Tissue Research*, 26, pp.37 - 45.

Arias, J.L. et al., 1992. Partial biochemical and immunochemical characterization of avian eggshell extracellular matrices. *Archives of Biochemistry and Biophysics*, 298(1), pp.293 - 302.

Arias, J.L. et al., 1993. Biomineralization and egg-shells - cell-mediated acellular compartments of mineralized extracellular matrix. *International Review of Cytology*, 145, pp.217 - 250.

Bada, J.L., 1971. Kinetics of the non-biological decomposition and racemization amino acids in natural waters. *Advances in Chemistry Series*, 106, pp.309-331.

Bada, J.L., 1972. The dating of fossil bones using the racemization of isoleucine. *Earth and Planetary Science*, 15, p.223.

Bada, J.L., 1985. Racemization of amino acids. In G. Barrett, ed. *Chemistry and Biochemistry of the Amino Acids*. London: Chapman & Hall, pp. 399-414.

Bada, J.L., 1991. Amino Acid Cosmogeochemistry. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 333, pp.349 - 358.

Bada, J.L. et al., 1984. Accelerator mass spectrometry radiocarbon ages of amino acid extracts from Californian Palaeoindian skeletons. *Nature*, 312, pp.442 - 444.

Bada, J.L. et al., 1978. Decomposition of hydroxy amino acids in foraminiferal tests: kinetics, mechanism and geochronological implications. *Earth Planetary Science Letters*, 41, pp.67 - 76.

Bada, J.L. et al., 1986. Isoleucine stereoisomers on the Earth. *Nature*, 319, pp.314 - 316.

Bada, J.L. et al., 1994. Amino acid racemization in amber-entombed insects: implications for DNA preservation. *Geochimica et Cosmochimica Acta*, 58(14), pp.3131 - 3135.

Bada, J. & Man, E., 1980. Amino acid analysis in Deep Sea drilling project cores: kinetics and mechanisms of some reactions and their applications in geochronology and in paleotemperature and heat flow determinations. *Earth Science Reviews*, 16, pp.21-55.

Bada, J. & Protsch, R., 1973. Racemization reaction of aspartic acid and its use in dating fossil bones. *Proceedings of the National Academy of Science*, 70, pp.1331-1334.

Bada, J. & Schroeder, R., 1972. Racemization of isoleucine in calcareous marine sediments: kinetics and mechanism. *Earth and Planetary Science Letters*, 15, pp.1-11.

Bada, J.L., Schroeder, J.A. & Carter, G.F., 1974. New evidence for the antiquity of man in North America deduced from aspartic acid racemization. *Science*, 184, pp.791 - 793.

Bada, J.L. & Schroeder, R.A., 1975. Amino acid racemization reactions and their geochemical implications. *Naturwissenschaften*, 62, pp.71 - 79.

Bada, J.L. & Shou, M.-Y., 1980. Kinetics and mechanism of amino acid racemization in aqueous solution and in bones. In P. E. Hare, T. C. Hoering, & K. King Jr., eds. *Amino Acid Biogeochemistry*. Wiley, pp. 235 - 255.

Bada, J.L., Wang, X.S. & Hamilton, H., 1999. Preservation of key biomolecules in the fossil record: current knowledge and future challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354, pp.77 - 87.

Baker, P., 2008. Procurement, Presentation and Consumption of Domestic and Wildfowl at Windsor Castle, England in the 12th to 14th c. In W. Prummel, J. T. Zeiler, & D. C. Brinkhuizen, eds. *Birds in Archaeology: Proceedings of the 6th Meeting of the Bird Working Group in Groningen (23.8-27.8.2008)*. Eelde: Barkhuis, pp. 57-70.

Baldwin, J., 2009. Harvesting seabirds and their eggs on the Irish Sea islands (Part 1: the Welsh islands, Lundy and Scilly). *Folk Life: Journal of Ethnological Studies*, 47, pp.76-96.

Baldwin, J., 2010. Harvesting seabirds and their eggs on the Irish Sea islands (Part 2: Mainland of Man). *Folk Life: Journal of Ethnological Studies*, 48(2), pp.144-169.

Baldwin, J.R., 2011. Harvesting Seabirds and their eggs on the Irish Sea islands (Part 3: the Calf of Man). *Folk Life: Journal of Ethnological Studies*, 49(1), pp.30-74.

Barrett, J.H., 2012. *Being an Islander: Production and identity at Quoygrew, Orkney, AD 900-1600*, Cambridge: McDonald Institute for Archaeological Research.

Barrett, J.H., Beukens, R.P. & Nicholson, R.A., 2001. Diet and ethnicity during the Viking colonization of northern Scotland: evidence from fish bones and stable carbon isotopes. *Antiquity*, 75, pp.145-54.

Barrett, J.H. & Gerrard, J.F., 2012. The Quoygrew Sequence. In J. H. Barrett, ed. *Being an Islander: Production and identity at Quoygrew AD 900 -1600*. Cambridge: McDonald Institute for Archaeological Research, pp. 47-96.

Barrett, J., Hall, A., Johnstone, C., Kenward, H., O'Connor, T., & Ashby, S. (2007). Interpreting the plant and animal remains from Viking-age Kaupang. In D. Skre (Ed.), *Kaupang in Skiringssal* (pp. 283-319). Aarhus: Aarhus University Press.

Beacham, E.B. & Durand, S.R., 2007. Eggshell and the archaeological record: new insights into turkey husbandry in the American Southwest. *Journal of Archaeological Science*, 34, pp.1610-1621.

Becker, C. & Grupe, G., 2012. Archaeometry meets archaeozoology: Viking Haithabu and medieval Schleswig revisited. *Archaeological and Anthropological Sciences*, 4, pp.241-262.

Becking, J., 1975. The ultrastructure of the avian eggshell. *Ibis*, 117(2), pp.143-151.

Bejenaru, L., Stanc, S.M. & Oleniuc, F.C., 2008. Domestic Birds in the Medieval Settlements on the Territory of Romania. In W. Prummel, J. T. Zeiler, & D. C. Brinkhuizen, eds. *Birds in Archaeology: Proceeding of the 6th Meeting of the ICAZ Bird Working Group in Grongingen (23.8-27.8.2008)*. Eelde: Barkhuis, pp. 29-36.

Belcher, A.M. et al., 1996. Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature*, 381, pp.56 - 58.

Bentov, S. et al., 2010. Stabilization of amorphous calcium carbonate by phosphate rich organic matrix proteins and by single phosphoamino acids. *Journal of Structural Biology*, 171, pp.207 - 215.

Berman, A., Addadi, L. & Weiner, S., 1988. Interactions of sea-urchin skeleton macromolecules with growing calcite crystals - a study of intracrystalline proteins. *Nature*, 331, pp.546-548.

Berman, A. et al., 1990. Intercalation of Sea Urchin Proteins in Calcite: Study of a Crystalline Composite Material. *Science*, 250(4981), pp.664-667.

Bienvenut, W. et al., 2002. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics*, 2, pp.868-876.

Bigelow, G.F., 1985. Sandwick, Unst and Late Norse Shetland Economy. In B. Smith, ed. *Shetland Archaeology*. Lerwick: The Shetland Times Ltd., pp. 95-127.

Bingle, C.D. & Craven, C.J., 2004. Meet the relatives: a family of BPI and LBP-related proteins. *Trends in Immunology*, 25, pp.53 - 55.

Blake, C.C.F. et al., 1965. Structure of hen egg-white lysozyme. *Nature*, 206(4986), pp.757 - 761.

Blench, R. & MacDonald, K.C., 2012. Chickens. In K. F. Kiple & K. C. Ornelas, eds. *The Cambridge World History of Food*. Cambridge: Cambridge University Press, pp. 496-499.

Bond, G., Board, R. & Scott, V., 1988. A comparative study of changes in the structure of avian eggshells during incubation. *Zoological Journal of the Linnean Society*, 92, pp.105-113.

Bond, J.M., 2007. The mammal bone. In J. Hunter, J. M. Bond, & N. Smith, eds. Investigations in Sanday, Orkney, vol. 1: Excavations at Pool, Sanday Multi-period Settlement from Neolithic to Late Norse Times. Kirkwall: The Orcadian/Historic Scotland, pp. 207-262.

Brant, A., 1998. A brief history of the turkey. *World's Poultry Science Journal*, 54(4), pp.365-373.

Brinton, K. & Bada, J., 1995. Comment on Aspartic acid racemization and protein diagenesis in corals over the last 350 years'' by G.A. Goodfriend, P.E. Hare and E.R.M. Druffel. *Geochimica et Cosmochimica Acta*, 59, pp.415-416.

Brooks, A.S. et al., 1990. Dating Pleistocene archaeological sites by protein diagenesis in ostrich eggshell. *Science*, 248, pp.60 - 64.

Buckley, M. et al., 2008. Comparing the survival of osteocalcin and mtDNA in archaeological bone from four European sites. *Journal of Archaeological Science*, 35, pp.1756 - 1764.

Buckley, M. et al., 2009. Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 23, pp.3843 - 3854.

Buckley, M. et al., 2010. Distinguishing between archaeological sheep and goat bones using a single collagen peptide. *Journal of Archaeological Science*, 37, pp.13 - 20.

Cann, J.H., De Deckker, P. & Murray-Wallace, C.V., 1991. Coastal aboriginal shell middens and their palaeoenvironmental significance, Robe Range, South Australia. *Transactions of the Royal Society of South Australia*, 115, pp.161 - 175.

Chien, Y.-C. et al., 2008. Ultrastructural matrix-mineral relationships in avian eggshell, and effects of osteopontin on calcite growth in vitro. *Journal of Structural Biology*, 163, pp.84 - 99.

Chien, Y.-C., Hincke, M.T. & McKee, M.D., 2009. Ultrastructure of avian eggshell during resorption following egg fertilization. *Journal of Structural Biology*, 168, pp.527 - 538.

Clarke, S.J. et al., 2007. The geochronological potential of isoleucine epimerisation in cassowary and megapode eggshells from archaeological sites. *Journal of Archaeological Science*, 34, pp.1051 - 1063.

Clarke, S.J. & Murray-Wallace, C., 2006. Mathematical expressions used in amino acid racemisation geochronology - A review. *Quaternary Geochronology*, 1, pp.261-278.

Coghlan, M. et al., 2012. Egg forensics: an appraisal of DNA sequencing to assist in species identification of illegally smuggled eggs. *Forensic Science International: Genetics*, 6, pp.268-273.

Colfen, H. & Antonietti, M., 2005. Mesocrystals: inorganic superstructures made by highly parallel crystallization and controlled alignment. *Angewandte Chemie International Edition* 44, pp.5576-5591.

Colfen, H. & Mann, S., 2003. Higher-order organization by mesoscale self-assembly and transformation of hybrid nanostructures. *Angewandte Chemie Int. Ed.*, 42, pp.2350-2365.

Collins, M. et al., 1991. Macromolecules in brachiopod shells: characterization and diagenesis. *Lethaia*, 24(4), pp.387-397.

Collins, M.J. et al., 1992. Experimental evidence for condensation reactions between sugars and proteins in carbonate skeletons. *Geochimica et Cosmochimica Acta*, 56, pp.1539 - 1544.

Collins, M.J. et al., 2009. Is amino acid racemization a useful tool for screening for ancient DNA in bone? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 276, pp.2971 - 2977.

Collins, M.J. & Riley, M.S., 2000. Amino acid racemization in biominerals: the impact of protein degradation and loss. In G. A. Goodfriend et al., eds. *Perspectives in Amino Acid and Protein Geochemistry*. Oxford University Press, pp. 120 -142.

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

Coltherd, J.B., 1966. The Domestic Fowl in Ancient Egypt. *Ibis*, 108(2), pp.217-223.

Crawford, B.E., 1987. *Scandinavian Scotland*, Leicester: University of Leicester Press.

Crenshaw, M., 1972. The soluble matrix from Mercenaria mercenaria shell. *Biomineralization*, 6, pp.6-11.

Crisp. M. et al., 2013. Isolation of the intracrystalline proteins and kinetic studies in Struthio camelus (ostrich) eggshell for amino acid geochronology. *Quaternary Geochronology*, 16, pp.110-128.

De Reggi, M. & Gharib, B., 2001. Protein-X, pancreatic stone-, pancreatic thread-,reg-protein, P19, lithostathine, and now what? *Current Protein and Peptide Science*, 2, pp.19 - 42.

Demarchi, B., 2009. *Geochronology of coastal prehistoric environments: a new closed system approach using amino acid racemisation*. University of York.

Demarchi, B. et al., 2011. Amino acid racemization dating of marine shells: a mound of possibilities. *Quaternary International*, 239(1-2), pp.114-124.

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. *Quaternary Geochronology*, *16*, 158-172.

Demarchi, B., Rogers, K., Fa, D. A., Finlayson, C. J., Milner, N., & Penkman, K. E. H. (2013). Intra-crystalline protein diagenesis (IcPD) in Patella vulgata. Part 1: Isolation and testing of the closed system. *Quaternary Geochronology*, *16*, 144-157.

Dennis, J.E. et al., 1996. Microstructure of matrix and mineral components of eggshells from white leghorn chickens (*Gallus gallus*). *Journal of Morphology*, 228, pp.287 - 306.

Dobney, K. & Jaques, D., 2002. Avian signatures for identity and status in Anglo-Saxon England. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.7-21.

Dominguez-Vera, J.M. et al., 2000. The effect of avian uterine fluid on the growth behavior of calcite crystals. *Poultry Science*, 79, pp.901 - 907.

Dunn, I.C. et al., 2008. Polymorphisms in eggshell organic matrix genes are associated with eggshell quality measurements in pedigree Rhode Island Red hens. *Animal Genetics*, 40, pp.110-114.

Eastham, A. & Gwynn, I.A., 1997. Archaeology and the electron microscope: eggshell and neural network analysis of images in the Neolithic. *Anthropozoologica*, 25, pp.85-94.

Egloff, C. et al., 2009. A nondestructive method for obtaining maternal DNA from avian eggshells and its application to embryonic viability determination in herring gulls (Larus argentatus). *Molecular Ecology Resources*, 9(1), pp.19-27.

Eisenberg, D. et al., 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of Molecular Biology*, 179(2), pp.125-142.

Falini, G. et al., 1996. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science*, 271, pp.67 - 69.

Feng, Q.L. et al., 2000. Polymorph and morphology of calcium carbonate crystals induced by proteins extracted from mollusk shell. *Journal of Crystal Growth*, 216, pp.459 - 465.

Fenn, J. et al., 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926), pp.64-71.

Fernandez, M.S., Araya, M. & Arias, J.L., 1997. Eggshells are shaped by a precise spatio-temporal arrangement of sequentially deposited macromolecules. *Matrix Biology*, 16, pp.13 - 20.

Fernandez, M.S. et al., 2003. Localization of osteopontin in oviduct tissue and eggshell during different stages of the avian egg laying cycle. *Journal of Structural Biology*, 143, pp.171 - 180.

Freeman, C.L. et al., 2010. Structural control of crystal nuclei by an eggshell protein. *Angewandte Chemie International*, 49, pp.5135 - 5137.

Freeman, C. et al., 2011. Simulations of Ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation. *Journal of Physical Chemistry C*, 115, pp.8175-8183.

Fu, G. et al., 2005. CaCO3 biomineralization: acidic 8-kDa proteins isolated from aragonitic abalone shell nacre can specifically modify calcite crystal morphology. *Biomacromolecules*, (6), pp.1289 - 1298.

Fujii, N. & Saito, T., 2004. Homochirality and life. *Chemical Record*, 4(5), pp.267 - 278.

Fumihito, A. et al., 1996. One subspecies of the red junglefowl (Gallus gallus gallus) suffices as the matriarchic ancestor of all domestic breeds. *Proceedings of the National Academy of Sciences*, 91, pp.12505-12509.

Gaines, S. & Bada, J., 1988. Aspartame decomposition and epimerization in the diketopiperazine and dipeptide products as a function of pH and temperature. *Journal of Organic Chemistry*, 53, pp.2757-2764.

Gasteiger, E et al., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31, pp.3784-3788.

Gautron, J., Hincke, M.T. & Nys, Y., 1997. Precursor matrix proteins in the uterine fluid change with stages of eggshell formation in hens. *Connective Tissue Research*, 363(3), pp.195 - 210.

Gautron, J. et al., 2001. Ovocalyxin-32, a novel chicken eggshell matrix protein. *Journal of Biological Chemistry*, 276(42), pp.39243 - 39252.

Gautron, J. et al., 2007. Cloning of ovocalyxin-36, a novel chicken eggshell protein related to lipopolysaccharide-binding proteins (LPB) bactericidal permeability-increasing proteins (BPI) and Plunc family proteins. *Journal of Biological Chemistry*, 282, pp.5273 - 5286.

Gebauer, D. et al., 2009. The multiple roles of additives in CaCO3 crystallization: a quantitative case study. *Advanced Materials*, 21, pp.435 - 439.

Gebauer, D. et al., 2010. Proto-calcite and proto-vaterite in amorphous calcium carbonates. *Angewandte Chemie International*, 49, pp.8889 - 8991.

Gebauer, D., Volkel, A. & Colfen, H., 2008. Stable prenucleation calcium carbonate clusters. *Science*, 322, pp.1819 - 1822.

Geiger, T. & Clarke, S., 1987. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. *Journal of Biological Chemistry*, 262, pp.785 - 794.

Gill, B.J., 2000. Morphometrics of Moa eggshell fragments (Aves: Dinornithiformes) from Late Holocene dune sands of the Karikari Peninsula, New Zealand. *Journal of the Royal Society of New Zealand*, 30(2), pp.131-145.

Goodfriend, G.A., 1987. Chronostratigraphic studies of sediments in the Negev Desert, using amino acid epimerization analysis of land snail shells. *Quaternary Research*, 28, pp.374 - 392.

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

Goodfriend, G.A., 1992. Rapid racemization of aspartic acid in mollusc shells and potential for dating over recent centuries. *Nature*, 357, pp.399 - 401.

Goodfriend, G., Brigham-Grette, J. & Miller, G., 1996. Enhanced age resolution of the marine Quaternary record in the Arctic using aspartic acid racemization dating of bivalve shells. *Quaternary Research*, 45, pp.176-187.

Goodfriend, G., Kashgarian, M. & Harasewych, M., 1995. Use of aspartic acid racemization and post-bomb 14C to reconstruct growth rate and longevity of the deep-water slit shell Entemnotrochus adansonianus . *Geochimica et Cosmochimica Acta*, 59, pp.1125-1129.

Goodfriend, G.A. & Meyer, V.R., 1991. A comparative study of the kinetics of amino acid racemization/epimerization in fossil and modern mollusk shells. *Geochimica et Cosmochimica Acta*, 55, pp.3355 - 3367.

Goodfriend, G.A. & Rollins, H.B., 1998. Recent barrier beach retreat in Georgia: dating exhumed salt marshes by aspartic acid racemization and post-bomb radiocarbon. *Journal of Coastal Research*, 14(3), pp.960 - 969.

Gotliv, B.-A., Addadi, L. & Weiner, S., 2003. Mollusk shell acidic proteins: in search of individual functions. *ChemBioChem*, 4, pp.522 - 529.

Gower, L.B. & Odom, D.J., 2000. Deposition of calcium carbonate films by a polymer-induced liquid-precursor (PILP) process. *Journal of Crystal Growth*, 210, pp.719 - 734.

Graham-Campbell, J. & Batey, C.E., 1998. *Vikings in Scotland: an Archaeological Survey*, Edinburgh: Edinburgh University Press.

Griffiths, D. & Harrison, J., 2011. Orkney: a Viking-Age settlement under wind-blown sand. *Medieval Archaeology*, 55, pp.321-326.

Grotzinger, J.P., Watters, W.A. & Knoll, A.H., 2000. Calcified metazoans in thrombolite-stromatolite reefs of the terminal Proterozoic Nama group, Namibia. *Paleobiology*, 26(3), pp.334 - 359.

Gueta, R. et al., 2007. Local atomic order and infrared spectra of biogenic calcite. *Angewandte Chemie Internation Edition*, 46, pp.291-294.

Hackett, S.J. et al, 2008. A phylogenomic study of birds reveals their evolutionary history. *Science*, 320, pp.1763-1768.

Hall, R., 1989. Archaeological Introduction. In T. O'Connor, *Bones from Anglo-Scandinavian Levels at 16-22 Coppergate*. London: Council for British Archaeology, pp. 141-142.

Hamilton-Dyer, S., 2002. The Bird Resources of Medieval Novgorod, Russia. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.99-107.

Hare, P., 1971. Effect oh hydrolysis on the racemisation rate of amino acids. *Carnegie Institute of Washington Yearbook*, 70, pp.256-258.

Hare, P.E. et al., 1993. Chemical clocks and thermometers: diagenetic reactions of amino acids in fossils. *Carnegie Institute Washington Year Book*, 92, pp.80 - 85.

Hare, P.E. & Abelson, P.H., 1967. Racemization of amino acids in fossil shells. *Carnegie Institute Washington Year Book*, 65, pp.526 - 528.

Hare, P. & Gil-Av, E., 1979. Separation of D and L amino acids by liquid chromatography: use of chiral eluants. *Science*, 204(15), pp.1226-1228.

Hare, P.E., Hoering, T.C. & King Jr., K. eds., 1980. *Biogeochemistry of Amino Acids*, New York: John Wiley & Sons.

Hare, P. & Mitterer, R., 1967. Non-protein amino acids in fossil shells. *Yearbook of the Carnegie Institute of Washington*, 73, pp.576-581.

Hare, P.E. & Mitterer, R.M., 1969. Laboratory simulation of amino acid diagenesis in fossils. *Carnegie Institute Washington Year Book*, 67, pp.205 - 208.

Harland, J., 2012. Animal husbandry; the mammal bone. In J. Barrett, ed. *Being an Islander: Production and Identity at Quoygrew, Orkney A.D. 900-1600*. Cambridge: McDonald Institute for Archaeological Research, pp. 139-154.

Harland, J.F. & Barrett, J.H., 2012. The Maritime Economy: Fish bone. In J. H. Barrett, ed. *Being an Islander: Production and identity at Quoygrew AD 900 -1600*. Cambridge: McDonald Institute for Archaeological Research, pp. 115-138.

Harland, J.F. et al., 2012. Fowling: the Bird Bone. In J. H. Barrett, ed. *Being an Islander: Production and identity at Quoygrew AD 900 -1600*`. Cambridge: McDonald Institute for Archaeological Research, pp. 155-160.

Harper, J., 1972. The tardy domestication of the duck. *Agricultural History*, 46(3), pp.385-389.

Harrison, C.J.O., 1980. A re-examination of British Devensian and earlier Holocene bird bones in the British Museum (Natural History). *Journal of Archaeological Science*, 7, pp.53-68.

Henzel, W.J. et al., 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proceedings of the National Academy of Sciences*, 90, pp.5011 - 5015.

Hernandez Carrasquilla, F., 1992. Some comments on the introduction of domestic fowl in Iberia. *Archaeofauna*, 1, pp.45-53.

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

Hincke, M.T. et al., 1995. Purification and immunochemistry of a soluble matrix protein of the chicken eggshell (ovocleidin-17). *Calcified Tissue International*, 56, pp.578 - 583.

Hincke, M.T. et al., 1999. Molecular cloning and ultrastructural localization of the core protein of an eggshell matrix proteoglycan, Ovocleidin-116. *Journal of Biological Chemistry*, 274, pp.32915 - 32923.

Hincke, M.T. et al., 2000. Identification and localization of lysozyme as a component of eggshell membranes and eggshell matrix. *Connective Tissue Research*, 44, pp.16 - 19.

Hincke, M.T., Nys, Y. & Gautron, J., 2010. The role of matrix proteins in eggshell formation. *Journal of Poultry Science*, 47, pp.208 - 219.

Hoering, T., 1980. The organic constituent of fossil mollusc shells. In P. Hare, T. Hoering, & K. King Jr, eds. *Biogeochemistry of Amino Acids*. New York: Wiley, pp. 193-201.

de Hoffmann, E. & Stroobart, V., 2007. *Mass Spectrometry: Principles & Applications* 3rd ed., Chichester, UK: John Wiley & Sons.

Hollemeyer, K., Altmeyer, W. & Heinzle, E., 2007. Identification of furs of domestic dog, raccoon dog, rabbit and domestic cat by hair analysis using MALDI-ToF mass spectrometry. *Spectroscopy Europe*, 19(2), pp.8 - 15.

Hollemeyer, K. et al., 2008. Species identification of Oetzi's clothing with matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry based on peptide pattern similarities of hair digests. *Rapid Communications in Mass Spectrometry*, 22, pp.2751 - 2767.

Horvat-Gordon, M. et al., 2008. Ovocleidin-116 is present in avian skeletal tissues. *Poultry Science*, (87), pp.1618 - 1623.

Hunn, E. et al., 2003. *A study of traditional use of birds' eggs by the Huna Tlingit*, Seattle.

Huntington, J.A. & Stein, P.E., 2001. Structure and properties of ovalbumin. *Journal of Chromatography B*, 756, pp.189 - 198.

James, P. et al., 1993. Protein identification by mass profile fingerprinting. *Biochemical and Biophysical Research Communications*, 195(1), pp.58 - 64.

Johnson, B.J. et al. 1997, The determination of late Quaternary paleoenvironments at Equus Cave, South Africa, using stable isotopes and amino acid racemization in ostrich eggshell. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 136, pp.121-137.

Johnson, B.J. & Miller, G.H., 1997. Archaeological implications of amino acid racemization. *Archaeometry*, 39(2), pp.265 - 287.

Kanginakudru, S. et al., 2008. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern chicken. *BMC Evolutionary Biology*, 8, pp.174-188.

Kaufman, D., 2006. Temperature sensitivity of aspartic acid and glutamic acid racemization in the foraminifera Pulleniatina. *Quaternary Geochronology*, 1, pp.188-207.

Kaufman, D. & Manley, W., 1998. A new procedure for determining DL amino acid ratios in fossils using reverse phase liquid chromatography. *Quaternary Geochronology*, 17, pp.987-1000.

Kaufman, D. & Miller, G., 1995. Isoleucine epimerization and amino acid composition in molecular-weight separations of Pleistocene Genyornis eggshell. *Geochimica et Cosmochimica Acta*, 59, pp.2757-2765.

Kaufman, D. & Sejrup, H., 1995. Isoleucine epimerization in the high-molecular-weight fraction of Pleistocene Arctica. *Quaternary Science Reviews*, 14, pp.337-350.

Kawasaki, K. & Weiss, K., 2003. Mineralized tissue and vertebrate evolution: the secretory calcium-binding phosphoprotein gene cluster. *PNAS*, 100, pp.4060-4065.

Keepax, C.A., 1981. Avian egg-shell from archaeological sites. *Journal of Archaeological Science*, 8, pp.315-335.

Keung, W.M. & Azari, P., 1982. Structure and function of ovotransferrin. II. Iron-transferring activity of iron-binding fragments of ovotransferrin with chicken embryo red cells. *Journal of Biological Chemistry*, 257, pp.1184 - 1188.

Kightly, C. (1984). *Country voices: life and lore in farm and village*. London: Thames & Hudson.

Kimber, R.W.L., Griffin, C.V. & Milnes, A.R., 1986. Amino acid racemization dating: evidence of apparent reversal in aspartic acid racemization with time in shells of Ostrea. *Geochimica et Cosmochimica Acta*, 50, pp.1159 - 1161.

Kimber, R.W.L. & Griffin, C.V, 1987. Further evidence of the complexity of the racemization process in fossil shells with implications from amino acid racemization dating. *Geochimica et Cosmochimica Acta*, 56, pp.839 - 846.

Koga, N., Nakagoe, Y. & Tanaka, H., 1998. Crystallization of amorphous calcium carbonate. *Thermochimica Acta*, 318, pp.239 - 244.

Krattiger, P. et al., 2010. Investigation of active crystal morphogenesis peptide sequences from peptide libraries by crystallization on peptide functionalized beads. *Colloids and Surfaces A: Physiochemical and Engineering Aspects*, 354(1-3), pp.218 - 225.

Kriausakul, N. & Mitterer, R.M., 1978. Isoleucine epimerisation in peptides and proteins: kinetic factors and application to fossil proteins. *Science*, 201, pp.1011 - 1014.

Kriausakul, N. & Mitterer, R.M., 1980. Some factors affecting the epimerization of isoleucine in peptides and proteins. In P. E. Hare, T. C. Hoering, & K. King Jr., eds. *Biogeochemistry of Amino Acids*. Wiley, p. 283 296.

Kriausakul, N. & Mitterer, R.M., 1989. Calculation of amino acid racemization ages based on apparent parabolic kinetics. *Quaternary Science Reviews*, 8, pp.353 - 357.

Kysely, R., 2010. Review of the oldest evidence of domestic fowl Gallus gallus f. domestica from the Czech Republic in its European context. *Actazoologica Cracoviensia*, 53A(1-2), pp.9-34

Lakshminarayanan, R. et al., 2005. Structure-function relationship of avian eggshell matrix proteins: a comparative study of two major eggshell matrix proteins, ansocalcin and OC-17. *Biomacromolecules*, 6, pp.741 - 751.

Lakshminarayanan, R. et al., l, 2006. Formation of transient amorphous calcium carbonate precursor in quail eggshell mineralization: an in vitro study. *Biomacromolecules*, 7, pp.3202 - 3209.

Lakshminarayanan, R., Kini, R.M. & Valiyaveettil, S., 2002. Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix. *Proceedings of the National Academy of Science*, 99, pp.5155 - 5159.

Lam, R.S.K. et al., 2007. Synthesis-dependant structural variations in amorphous calcium carbonate. *CrystEngComm*, 9, pp.1226 - 1236.

Lammie, D. et al., 2006. Scanning microfocus small angle X-ray scattering study of the avian eggshell. *Journal of Bionic Engineering*, 3, pp.11 - 18.

Lavelin, I., Meiri, N. & Pines, M., 2000. New insight in eggshell formation. *Poultry Science*, 79, pp.1014 - 1017.

Leach, R.M., 1982. Biochemistry of the organic matrix of the eggshell. *Poultry Science*, 61, pp.2040 - 2047.

Lee, P. & Prys-Jones, R., 2008. Extracting DNA from museum bird eggs, and whole genome amplification of archive DNA. *Molecular Ecology Resources*, 8(3), pp.551-560.

Lepot, K., Benzenara, K., Brown, G.E., Philippot, P., 2008. Microbially influenced formation of 2.7 billion year old stromatolites. *Nature Geoscience,* 1(2), pp.118-121.

Li, H. et al., 2009. Visualizing the 3D internal structure of calcite single crystals grown in agarose hydrogels. *Science*, 326, pp.1244 - 1247.

Lightman, B. (2000). The story of nature: Victorian popularizers and scientific narrative. *Victorian Review*, 1-29.

Liska, A.J. & Shevchenko, A., 2003. Expanding the organismal scope of proteomics: Cross-species protein identification by mass spectrometry and its implications. *Proteomics*, 3, pp.19 - 28.

Liu, Y.-P. et al., 2006. Multiple maternal origins of chickens: Out of the Asian jungles. *Molecular Phylogenetics and Evolution*, 38, pp.12-19.

Luff, R. L. (2012). Ducks. In K. F. Kiple & K. C. Ornelas (Eds.), *The Cambridge World History of Food* (pp. 517-524). Cambridge: Cambridge University Press.

Mac Con Iomaire, M. & Cully, A., 2007. The History of eggs in Irish cuisine and culture. In R. Hosking, ed. *Eggs in Cookery: Proceedings of the Oxford Symposium on Food and Cookery*. Totnes: Prospect, pp. 137-149.

MacDonald, K. C., & Blench, R. (2012). Geese. In K. F. Kiple & K. C. Ornelas (Eds.), *The Cambridge World History of Food* (pp. 529-531). Cambridge: Cambridge University Press.

Magee, J.W. et al., 2009. Evaluating Quaternary dating methods: radiocarbon, U-series, luminescence and amino acid racemization dates of a late Pleistocene emu egg. *Quaternary Geochronology*, 4, pp.84 - 92.

Makowiecki, D. & Gotfredsen, A.B., 2002. Bird Remains of Medieval and Post-Medieval Coastal Sites at the Southern Baltic Sea, Poland. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.65-84.

Manaseryan, N. & Balyan, L., 2002. The Birds of Ancient Armenia. *Acta Zoologica Cracoviensia*, 45 (Special Issue: Proceedings of the 4th meeting of the ICAZ bird working group), pp.405-414.

Manley, W., Miller, G. & Czywczynski, J., 2000. Kinetics of aspartic acid racemization in Mya and Hiatella: modelling age and paleotemperatyre of high-latitude Quaternary mollusks. In G. Goodfriend et al., eds. *Perspectives in Amino Acid and Protein Geochemistry*. Oxford University Press, pp. 202-218.

Mann, K., 1999. Isolation of a glycosylated form of the chicken eggshell protein ovocleidin and determination of the glycosylation site. Alternative glycosylation/phosphorylation at an N-glycosylation sequon. *FEBS Letters*, 463, pp.12 - 14.

Mann, K., 2004. Identification of the major proteins of the organic matrix of emu (Dromaius novaehollandiae) and rhea (Rhea americana) eggshell calcified layer. *British Poultry Science*, 45(4), pp.483 - 490.

Mann, K., 2008. Proteomic analysis of the chicken egg vitelline membrane. *Proteomics*, 8, pp.2322 - 2332.

Mann, K. et al., 2000. The amino-acid sequence of the abalone (Haliotis laevigata) nacre protein perlucin. *European Journal of Biochemistry*, 267, pp.5257 - 5264.

Mann, K. et al., 2003. Disulfide-linked heterodimeric clusterin is a component of the chicken eggshell matrix and egg white. *Matrix Biology*, 22, pp.397 - 407.

Mann, K., Hincke, M.T. & Nys, Y., 2002. Isolation of ovocleidin-116 from chicken eggshells, correction of its amino acid sequence and identification of disulfide bonds and glycosylated Asn. *Matrix Biology*, 21, pp.383 - 387.

Mann, K., Macek, B. & Olsen, J.V., 2006. Proteomic analysis of the acid-soluble organic matrix of the chicken eggshell calcified layer. *Proteomics*, 6, pp.3801 - 3810.

Mann, K. & Siedler, F., 1999. The amino acid sequence of ovocleidin 17, a major protein of the avian eggshell calcified layer. *Biochemistry and Molecular Biology International*, 47(6), pp.997-1007.

Mann, K. & Siedler, F., 2004. Ostrich (Struthio camelus) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications. *Biochimica et Biophysica Acta*, 1696(1), pp.41-50.

Mann, K. & Siedler, F., 2006. Amino acid sequences and phosphorylation sites of emu and rhea eggshell C-type lectin-like proteins. *Comparative Biochemistry and Physiology*, 143, pp.160 - 170.

Manson-Bahr, P. (1959). Recollections of some famous British ornithologists. *Ibis*, *101*(1), 53-64

Marie, B. et al., 2009. Nacre evolution: a proteomic approach. *Material Research Society Symposium Proceedings*, 1187.

Marie, B. et al., 2011. Novel Proteins from the Calcifying Shell Matrix of the Pacific Oyster *Crassostrea gigas*. *Marine Biotechnology*, 13, pp.1159-1168.

Marin, F. & Luquet, G., 2005. Molluscan biomineralization: The proteinaceous shell constituents of *Pinna nobilis* L. *Materials Science and Engineering C*, 25, pp.105-111.

Marin, F. et al., 2005. Caspartin and calprismin, two proteins of the shell calcitic prisms of the Mediterranean fan mussel *Pinna nobilis*. *Journal of Biological Chemistry*, 280(40), pp.33895-33908.

Marin, F. et al., 2008. Molluscan shell proteins: primary structure, origin, and evolution. *Current Topics in Developmental Biology*, 80, pp.209-76.

Marin-Garcia, L. et al., 2008. Chemical recognition of carbonate anions by proteins involved in biomineralization processes and their influence on calcite crystal growth. *Crystal Growth and Design*, 8(4), pp.1340 - 1345.

McCoy, W.D., 1987. The precision of amino acid geochronology and paleothermometry. *Quaternary Science Reviews*, 6, pp.43 - 54.

McGovern, T.H. et al., 1983. A study of the floral and faunal remains from two Norse farms in the Western Settlement, Greenland. *Arctic Anthropology*, 20(2), pp.93-120.

McGovern, T. et al., 2006. Coastal connections, local fishing, and sustainable egg harvesting: patterns of Viking Age inland wild resource use in Myvatn district, Northern Iceland. *Environmental Archaeology*, 11(2), pp.187-205.

Medina, M. et al., 2011. Late Holocene subsistence and social integration in Sierras of Cordoba (Argentina): the South-American ostrich eggshells evidence. *Journal of Archaeological Science*, 38, pp.2071-2078.

Medzihradszky, K. et al., 2000. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Analytical Chemistry*, 72(3), pp.552-558.

Meldrum, F.C. & Sear, R.P., 2008. Now you see them. *Science*, 322, pp.1802 - 1803.

Miksik, I. et al., 2007. Proteins of insoluble matrix of avian (Gallus gallus) eggshell. *Connective Tissue Research*, 48, pp.1-8.

Miller, G.H. et al., 1992. Pleistocene geochronology and palaeothermometry from protein diagenesis in ostrich eggshells: implications for the evolution of modern humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 337, pp.149 - 157.

Miller, G.H. et al., 1999. Earliest modern humans in southern Africa dated by isoleucine epimerization in ostrich eggshell. *Quaternary Science Reviewa*, 18, pp.1537 - 1548.

Miller, G.H. et al., 1999. Pleistocene extinction of Genyornis newtoni: Human impact on Australian megafauna. *Science*, 283, pp.205 - 208.

Miller, G.H. et al., 2000. Isoleucine epimerization in eggshells of the flightless Australian birds Genyornis and Dromaius. In G. A. Goodfriend et al., eds. *Perspectives in Amino Acid and Protein Geochemistry*. Oxford University Press, pp. 161 - 181.

Miller, G. H., & Brigham-Grette, J. (1989). Amino acid geochronology: resolution and precision in carbonate fossils. *Quaternary International*, *1*, 111 - 128.

Milner, N., Barrett, J. & Welsh, J., 2007. Marine resource intensification in Viking Age Europe: the molluscan evidence from Quoygrew, Orkney. *Journal of Archaeological Science*, 34, pp.1461-1472.

Milner, N. & Barrett, J.H., 2012. The Maritime Economy: Mollusc Shell. In J. H. Barrett, ed. *Being an Islander: Production and identity at Quoygrew AD 900 -1600*. Cambridge: McDonald Institute for Archaeological Research, pp. 103-114.

Mine, Y., 1995. Recent advances in the understanding of egg white protein functionality. *Trends in Food Science & Technology*, 6, pp.225-232.

Morris, C.D. & Rackham, D.J., 1992. *Norse and Later Settlement and Subsistence in the North Atlantic*, Glasgow: Universtiy of Glasgow.

Morris, C.D., Batey, C.E. & Rackham, D.J., 1995 (eds). *Freswick Links, Caithness: Excavation and survey of a Norse settlement*. Inverness: Highland Libraries

Mitterer, R.M., 1993. The diagenesis of proteins and amino acids in fossil shells. In M. H. Engel & S. A. Macko, eds. *Organic Geochemistry*. Plenum, pp. 739 - 752.

Mitterer, R.M. & Kriausakul, N., 1984. Comparison of the rates and degrees of isoleucine epimerization in dipeptides and tripeptides. *Organic Geochemistry*, 7, pp.91 - 98.

Mulville, J., 2002. The role of cetacea in prehistoric and historic Atlantic Scotland. *International Journal of Osteoarchaeology* , 12, pp.34-48.

Mulville, J., 2010. Red deer on Scottish Islands. In T. O'Connor & N. Sykes, eds. *Extinctions and Invasions: a Social History of British Fauna*. Windgather Press, pp. 43-50.

Mulville, J., Bond, J. & Craig, O., 2005. The white stuff; milking in the outer Scottish Isles. In J. Mulville & A. Outram, eds. *The Zooarchaeology of Fats, Oils, Milk and Dairying*. Oxford: Oxbow Books, pp. 167-182.

Murray, K. et al., 1965. The hydrolysis of arginine. *The Journal of Biological Chemistry*, 240, pp.705-709.

Murray-Wallace, C. & Kimber, R., 1993. Further evidence for apparent "parabolic" racemization kinetics in Quaternary molluscs. *Australian Journal of Earth Sciences*, 40, pp.313-317.

Nakahara, H., Kakaei, M. & Bevelander, G., 1981. Studies on the formation of the crossed lamellar structure in the shell of Strombus gigas. *The Veliger*, 23, pp.207-211.

Neuberger, A., 1948. Stereochemistry of amino acids. *Advances in Protein Chemistry*, 4, pp.297-383.

Nickl, H. & Henisch, N., 1969. Growth of calcite crystals in gels. *Journal of the Electrochemical Society*, 116, pp.1258-1260.

Nys, Y. et al., 2004. Avian eggshell mineralization: biochemical and functional characterization of matrix proteins. *Comptes Rendus Palevol*, 3, pp.549 - 562.

O Corrain, D., 2001. The Vikings in Ireland. In A.-C. Larsen, ed. *The Vikings in Ireland*. Roskilde: The Viking Ship Museum, pp. 17-29.

O'Connor, T., 1989. Bones from Anglo-Scandinavian Levels at 16-22 Coppergate. Council for British Archaeology.

O'Connor, T., 2000. Bones as evidence of meat production and distribution in York. In E. White, ed. *Feeding a city: York : the provision of food from Roman times to the beginning of the twentieth century*. Devon: Prospect Books, pp. 43-60.

O'Connor, T., 2010. Livestock and deadstock in early medieval Europe from the North Sea to the Baltic. *Environmental Archaeology*, 15, pp.1-15.

Okumura, T. et al., 2010. Characteristics of biogenic calcite in the prismatic layer of a pearl oyster, *Pinctada fucata*. *Micron*, 41(7), pp.821-826.

Okumura, T. et al., 2012. Microstructural Variation of Biogenic Calcite with Intracrystalline Organic Macromolecules. *Crystal Growth and Design*, 12, pp.224-240.

Olsen, J., Ong, S.-E. & Mann, M., 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics*, 3.6, pp.608-614.

Olsen, S. J. (2012). Turkeys, Part 1. In K. F. Kiple & K. C. Ornelas (Eds.), *The Cambridge World History of Food* (pp. 578-583). Cambridge: Cambridge University Press.

Orme, C.A. et al., 2001. Formation of chiral morphologies through selective binding of amino acids to calcite surface steps. *Nature*, 411, pp.775 - 779.

Oskam, C. et al., 2010. Fossil avian eggshell preserves ancient DNA. *Proceedings of the Royal Society B: Biological Sciences*, 277(1690), pp.1991-2000.

Oskam, C. et al., 2011. Molecular and morphological analyses of avian eggshell excavated from a late thirteenth century oven. *Journal of Archaeological Science*, 38, pp.2589-2595.

Osterstrom, O. & Lilja, C., 2012. Evolution of avian eggshell structure. *Journal of Morphology*, 273, pp.241-247.

Oviatt, C. et al., 1999. Reinterpretation of the Burmester Core, Bonneville Basin, Utah. *Quaternary Research*, 52, pp.180-184.

Panheleux, M. et al., 1999. Organic matrix composition and ultrastructure of eggshell: a comparative study. *British Poultry Science*, 40, pp.240 - 252.

Papayannopoulos, I., 1995. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrometry Reviews*, 14, pp.49-73.

Pappin, D.J.C., Hojrup, P. & Bleasby, A.J., 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology*, 3, pp.327 - 332.

Pellegrini, A. et al., 2004. Proteolytic fragments of ovalbumin display antimicrobial activity. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1672(2), pp.76-85.

Penkman, K.E.H., 2005. *Amino acid geochronology: a closed system approach to test and refine the UK model*.

Penkman, K., 2010. Amino acid geochronology: its impact on our understanding of the Quaternary stratigraphy of the British Isles. *Journal of Quaternary Science*, 25(4), pp.501 - 514.

Penkman, K. et al., 2007. Testing the aminostratigraphy of fluvial archives: the evidence from intra-crystalline proteins within freshwater shells. *Quaternary Science Reviews*, 26(22-24), pp.2958-2969.

Penkman, K et al., 2008. *British Aggregates: an improved chronology using amino acid racemization and degradation of intracrystalline amino acids (IcPD)*,

Penkman, K.E.H. et al., 2008. Closed-system behaviour of the intra-crystalline fraction of amino acids in mollusc shells. *Quaternary Geochronology*, 3, pp.2 - 25.

Penkman, K.E.H. et al., 2010. Amino acid geochronology of the type Cromerian of West Runton, Norfolk, UK. *Quaternary International*, 228(1-2), pp.25-37.

Penkman, K. et al., 2011. A chronological framework for the British Quaternary based on Bithynia opercula. *Nature*, 476, pp.446-449.

Perkins, D. et al., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20(18), pp.3551-3567.

Phillips, D.C., 1967. The hen egg-white lysozyme molecule. *Proceedings of the National Academy of Sciences*, 57(3), pp.484 - 496.

Pines, M., Knopov, V. & Bar, A., 1995. Involvement of osteopontin in egg shell formation in the laying chicken. *Matrix Biology*, 14, pp.765 - 771.

Politi, Y. et al., 2008. Transformation mechanism of amorphous calcium carbonate into calcite in the sea urchin larval spicule. *Proceedings of the National Academy of Science*, 105(45), pp.17362-17366.

Portugal, S.J. et al., 2010. Can museum egg specimens be used for proteomic analyses? *Proteome Science*, 8, pp.40 - 45.

Pouget, E.M. et al., 2009. The initial stages of template-controlled CaCO3 formation revealed by Cryo-TEM. *Science*, 323, pp.1455 - 1458.

Prajanban, B., Shawsuan, L., Daduang, S., Kommanee, J., Roytrakul, S., Dhiravisit, A., & Thammasirirak, S. (2012). Identification of five reptile egg whites proteins using MALDI-TOF mass spectrometry and LC/MS-MS analysis. *Journal of Proteomics*, *75*(6), 1940-1959.

Preece, R. & Penkman, K., 2005. New faunal analyses and amino acid dating of the Lower Palaeolithic site at East Farm, Barnham, Suffolk. *Proceedings of the Geologists Association*, 116, pp.363-377.

Protection of Birds Act 1954 - Legislation.gov.uk. Retrieved March 16, 2013, from http://www.legislation.gov.uk/ukpga/1954/30/contents/enacted

Qian, Y. et al., 1993. Kinetics of peptide hydrolysis and amino-acid decomposition at high temperature. *Geochimica et Cosmochimica Acta*, 57, pp.3281-3293.

Radkiewicz, J. et al., 1996. Accelerated racemization of aspartic acid and asparagine residues via succinimide intermediates: an ab initio theoretical exploration of mechanism. *Journal of the American Chemical Society*, 118, pp.9148-9155.

Rafalska, J.K., Engel, M.H. & Lanier, W.P., 1991. Retardation of racemization rates of amino acids incorporated into melanoidin. *Geochimica et Cosmochimica Acta*, 55, pp.3669 - 3675.

Reyes-Grajeda, J.P., Moreno, A. & Romero, A., 2004. Crystal structure of ovocleidin-17, a major protein of the calcified Gallus gallus eggshell: implications in the calcite mineral growth pattern. *Journal of Biological Chemistry*, 279, pp.40876 - 40881.

Reynolds, S.J., Mand, R. & Tilgar, V., 2004. Calcium supplementation of breeding birds: directions for future research. *Ibis*, 146, pp.601 - 614.

Robinson, D.S. & King, N.R., 1963. Carbonic anhydrase and formation of the hen's egg shell. *Nature*, 4892(199), pp.497 - 498.

Rose, M.L.H. & Hincke, M.T., 2009. Protein constituents of the eggshell: eggshell-specific matrix proteins. *Cellular and Molecular Life Sciences*, 66, pp.2707-2719.

Rowe, M., 1986. Archaeological dating. *Journal of Chemical Education*, 63(1), pp.16-20.

Russell, D., White, J., Maurer, G., & Cassey, P. (2010). Data-poor egg collections: cracking an important research resource. *Journal of Afrotropical Zoology*, *6*, 77-82.

Sarikaya, M. et al., 2003. Molecular biomimetics: nanotechnology through biology. *Nature Materials*, 2, pp.577 - 585.

Sepetov, N. et al., 1991. Rearrangement, racemization and decomposition of peptides in aqueous solution. *Peptide Research*, 4(5), pp.308-313.

Serjeantson, D., 1988. Archaeological and ethnographic evidence for seabird exploitation in Scotland. *Archaeozoologia*, 2, pp.209-24.

Serjeantson, D., 2001. The great auk and the gannet: a prehistoric perspective on the extinction of the great auk. *International Journal of Osteoarchaeology* , 11, pp.43-55.

Serjeantson, D., 2002. Goose Husbandry in Medieval England, and the Problem of Ageing Goose Bones. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.39-54.

Serjeantson, D., 2009. *Cambridge Manuals in Archaeology: Birds*, Cambridge: Cambridge University Press.

Sharples, N., 2005. *A Norse Farmstead in the Outer Hebrides: Excavations at Mound 3, Bornais, South Uist*, Oxford & Oakville: Oxbow Books.

Sharples, N., 2012. *A Late Iron Age Farmstead in the Outer Hebrides: Excavations at Mound 1, Bornais, South Uist*, Oxford and Oakville: Oxbow Books.

Shevchenko, A. et al., 2001. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole Time-of-Flight mass sopectrometry and BLAST homology searching. *Analytical Chemistry*, 73, pp.1917-1926.

Sidell, E.J., 1993 (a). A methodology for the identification of avian eggshell from archaeological sites. *Archaeofauna. Archaeornithology: Birds and the Archaeological Record*, 2, pp.45-51.

Sidell, E.J., 1993 (b). *A methodology for the identification of archaeological eggshell*, Philadelphia: University of Pennsylvania.

Sidell, E. J. (1995). Eggshell. In C. D. Morris, C. E. Batey, & D. J. Rackham (Eds.), *Freswick Links: Excavation and survey of a Norse settlement* (pp. 211-213). Inverness: Highland Libraries.

Silyn-Roberts, H. & Sharp, R.M., 1986. Crystal growth and the role of the organic network in eggshell biomineralization. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 227(1248), pp.303-324.

Simkiss, K., 1968. The structure and formation of the shell and shell membranes. In T. C. Carter (Ed), *Egg Quality: A Study of the Hen's Egg* (pp. 3-25). Edinburgh: Oliver & Boyd.

Smith, G.G. & Baum, R., 1987. First-order rate constants for the racemization of each component in a mixture of isomeric dipeptides and the diketopiperazines. *Journal of Organic Chemistry*, 52(11), pp.2248 - 2255.

Smith, G.G. & De Sol, B., 1980. Racemization of amino acids in dipeptides shows COOH > NH2 for non-sterically hindered residues. *Science*, 207, pp.765-767.

Smith, G.G. & Evans, R.C., 1980. The effect of structure and conditions on the rate of racemization of free and bound amino acids. In P. E. Hare, T. C. Hoering, & K. King Jr., eds. *Biogeochemistry of Amino Acids*. Wiley, pp. 257 - 282.

Smith, G.G., Williams, K.M. & Wonacott, D.M., 1978. Factors affecting the rate of racemization of amino acids and their significance to geochronology. *Journal of Organic Chemistry*, 52, pp.2248 - 2255.

Song, R-Q. & Colfen, H., 2010. Mesocrystals – ordered nanopartical superstructures. *Advanced Materials* 22, pp.1301-1330.

Song, R.-Q. & Colfen, H., 2011. Additive controlled crystallisation. *CrystEngComm*, 13, pp.1249-1276.

Stadelman, W.J., 2012. Chicken Eggs. In K. F. Kiple & K. C. Ornelas, eds. *The Cambridge World History of Food*. Cambridge: Cambridge University Press, pp. 499-508

Standing, K., 2003. Peptide and protein de novo sequencing by mass spectrometry. *Current Opinion in Structural Biology*, 13, pp.595-601.

Stathoplos, L. & Hare, P., 1993. Bleach removes labile amino acids from deep sea foraminifera. *Journal of Foraminiferal Research*, 23, pp.102-107.

Steen, H. & Mann, M., 2004. The ABC's (and XYZ's) of peptide sequencing. *Nature Reviews: Molecular Cell Biology*, 5, pp.699 - 711.

Stein, L. & Badyaev, A., 2011. Evolution of eggshell structure during rapid range expansion in a passerine bird. *Functional Ecology*, 25, pp.1215-1222.

Stein, P.E. et al., 1991. Crystal structure of uncleaved ovalbumin at 1.95 A resolution. *Journal of Molecular Biology,* 221 (3), pp.941-959.

Steinberg, S.M. & Bada, J.L., 1981. Diketopiperazine formation during investigation of amino acid racemization in dipeptides. *Science*, 213, pp.544 - 545.

Steinberg, S.M. & Bada, J.L., 1983. Peptide decomposition in the neutral pH region via the formation of diketopiperazines. *Journal of Organic Chemistry*, 48, pp.2295 -2298.

Stephenson, R. & Clarke, S., 1989. Succinimide formation from Aspartyl and Asparingyl Peptides as a model for the spontaneous degradation of proteins. *Journal of Biological Chemistry*, 264(11), pp.6164-6170.

Stewart, J.R.M. et al., 2013. ZooMS: making eggshell visible in the archaeological record. *Journal of Archaeological Science*, 40, pp.1797-1804.

Stone, D.J., 2006. The Consumption and Supply of Birds in Late Medieval England. In C. M. Woolgar, D. Serjeantson, & T. Waldron, eds. *Food in Medieval England: Diet and Nutrition*. Oxford: Oxford University Press, pp. 148-161.

Stolarski, J. & Mazur, M., 2005. Nanostructure of biogenic versus abiogenic calcium carbonate crystals. *Acta Palaeontologica Pol.*, 50(4), pp.847 - 865.

Storey, A.A. et al., 2007. Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proceedings of the National Academy of Sciences*, 104, pp.10335-10339.

Storey, A.A. et al., 2010. Mitochondrial DNA from 3000 year old chickens at the Teouma site, Vanuatu. *Journal of Archaeological Science*, 37, pp.2459-2468.

Storey, A.A. et al., 2012. Investigating the Global Dispersal of Chickens in Prehistory Using Ancient Mitochondrial DNA Signatures. *PLOS One*, 7(7), pp.1-11.

Sun, F., Liu, R., Zong, W., Tian, Y., Wang, M., & Zhang, P. (2010). A Unique Approach to the Mobile Proton Model: Influence of Charge Distribution on Peptide Fragmentation. *Journal of Physical Chemistry B*, *114*, 6350-6353.

Sykes, G.A., Collins, M.J & Walton, D.I., 1995. The significance of a geochemically isolated intracrystalline organic fraction within biominerals. *Organic Geochemistry*, 23(11/12), pp.1059 - 1065.

Tait, C., 2012. *The Orkney Guide Book*, Kirkwall: Charles Tait Photographic.

Tanaka, K. et al., 1988. Protein and polymer analyses up to m/z 100000 by Laser Ionization Time-of-Flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2(8), pp.151-153.

Texier, P.-J. et al., 2010. A Howiesons Poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at Diepkloof Rock Shelter, South Africa. *Proceedings of the National Academy of Sciences*, 107(14), pp.6180-6185.

Thys, S. & van Neer, W., 2008. Bird Remains from Late Medieval and Postmedieval sites in Brussels, Belgium. In W. Prummel, J. T. Zeiler, & D. C. Brinkhuizen, eds. *Birds in Archaeology: Proceedings of the 6th Meeting of the ICAZ Bird Working Group in Grongingen (23.8-27.8.2008)*. Eelde: Barkhuis, pp. 71-86.

Tian, X. et al., 2010. What makes an egg unique? Clues from evolutionary scenarios of egg-specific genes. *Biology of Reproduction*, 83, pp.893-900.

Tixier-Boichard, M., Bed'hom, B. & Rognon, X., 2011. Chicken Domestication: From Archaeology to Genomics. *Comptes Rendus Biologies*, 334, pp.197-204.

Tomiak, P. et al., 2013. Testing the limitations of artificial protein degradation kinetics using known age massive Porites coral skeletons. *Quaternary Geochronology*, 16, pp.87-109.

Towe, K., 1980. Preserved organic ultrastructure: An unreliable indicator for Paleozoic amino acid biogeochemistry. In P. Hare, T. Hoering, & K. King Jr., eds. *The Biogeochemistry of Amino Acids*. New York: Wiley, pp. 65-74.

Towe, K. & Thompson, G., 1972. The structure of some bivalve shell carbonates prepared by ion beam thinning. *Calcified Tissue Research*, 10, pp.38-48.

Tullet, S.G., 1975. Regulation of avian eggshell porosity. *Journal of Zoology*, 177, pp.339 - 348.

Tyler, C. & Fowler, S., 1978. The distribution of organic cores, cones, cone junctions and pores in the eggshells of wild birds. *Journal of Zoology*, 186, pp.1-14.

Tyrberg, T., 2002. The Archaeological Record of Domesticated and Tamed Birds in Sweden. *Acta Zoologica Cracoviensia*, 45(Special Issue: Proceedings of the 4th Meeting of the ICAZ Bird Working Group), pp.215-231.

Vallentyne, J., 1964. Biogeochemistry of organic matter II: Thermal reaction kinetics and transformation products of amino compounds. *Geochimica et Cosmochimica Acta*, 32, pp.1353-1356.

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

van Duin, A.C.T. & Collins, M.J., 1998. The effects of conformational constraints on aspartic acid racemization. *Organic Geochemistry*, 29(5 - 7), pp.1227 - 1332.

Vocadlo, D.J. et al., 2001. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature*, 412(6849), pp.835 - 838.

Voinescu, A.E. et al., 2007. Mineralization of CaCO3 in the presence of egg white lysozyme. *Langmuir*, 23, pp.12269 - 12274.

Walton, D., 1998. Degradation of intracrystalline proteins and amino acids in fossil brachipods. *Organic Geochemistry*, 28(6), pp.389-410.

Wang, X. et al., 2009. Lysozyme mediated calcium carbonate mineralization. *Journal of Colloid and Interface Science*, 332, pp.96 - 103.

Wehmiller, J., 1977. Amino acid studies of the Del Mar, California, midden site: apparent rate constants, ground temperature models, and chronological implications. *Earth and Planetary Science Letters*, 37, pp.184-196.

Wehmiller, J.F., 1980. Intergeneric differences in apparent racemization kinetics in mollusks and foraminifera: implications for models of diagenetic racemization. In P. E. Hare, T. C. Hoering, & K. King Jr., eds. *Biogeochemistry of Amino Acids*. Wiley, pp. 341 - 355.

Wehmiller, J., 1981. Kinetic model options for interpretation of amino acid enatiomeric ratios in quaternary mollusks: comments on a paper by Kvenvolden et al (1979). *Geochimica et Cosmochimica Acta*, 45, pp.261-264.

Wehmiller, J.F., 1982. A review of amino acid racemization studies in Quaternary molluscs: stratigraphic and chronological applications in coastal and interglacial sites, Pacific and Atlantic coasts, US, UK, Baffin Island and Tropical Islands. *Quaternary Science Reviews*, 1, pp.83 - 120.

Wehmiller, J. & Bealknap, D., 1978. Alternative kinetic models for the interpretation of amino acid enantiomeric ratios in Pleistocene mollusks: examples from California, Washington, and Florida. *Quaternary Research*, 9, pp.330-348.

Wehmiller, J. & Hare, P., 1971. Racemization of amino acids in marine sediments. *Science*, 173, pp.907-911.

Weiner, S. & Addadi, L., 1997. Design strategies in mineralized biological materials. *Journal of Materials Chemistry*, 7(5), pp.689-702.

Weiner, S., Sagi, I. & Addadi, L., 2005. Choosing the crystallization path less travelled. *Science*, 309, pp.1027-1028.

Wellman-Labadie, O., Lakshminarayanan, R. & Hincke, M.T., 2008. Antimicrobial properties of avian eggshell-specific C-type lectin-like proteins. *FEBS Letters*, 582, pp.699-704.

West, B. & Zhou, B.X., 1988. Did chickens go north? New evidence for domestication. *Journal of Archaeological Science*, 15, pp.515-533.

West, D. (1975). Coal landholding in mediaeval Scandinavia. *Holdsworth Law Review*, *3*, 16-19.

Westaway, R., 2009. Calibration of decomposition of serine to alanine in Bityhnia opercula as a quantitative dating technique for Middle and Late Pleistocene sites in Britain. *Quaternary Geochronology*, 4, pp.251-259.

Wildlife and Countryside Act 1981 - Legislation.gov.uk. Retrieved March 16, 2013, from http://www.legislation.gov.uk/ukpga/1981/69

Wilson, H. & Cannan, R., 1937. The glutamic acid-pyrrolidone carboxylic acid system. *The Journal of Biological Chemistry*, 119, pp.309-331.

Wilt, F.H., 1999. Matrix and mineral in the sea urchin larval skeleton. *Journal of Structural Biology*, 126, pp.216 - 226.

Wilt, F.H., Killian, C.E. & Livingston, B.T., 2003. Development of calcareous skeletal elements in invertebrates. *Differentiation*, 71, pp.237-250.

Wong, M. et al., 1984. Collagen in the egg shell membranes of the hen. *Developmental Biology*, 104, pp.28 - 36.

Wysocki, V. et al., 2000. Mobile and localized protons: a framework for understanding peptide dissociation. *Journal of Mass Spectrometry*, 35, pp.1399-1406.

Xie, A.-J. et al., 2005. Crystal growth of calcium carbonate with various morphologies in different amino acid systems. *Journal of Crystal Growth*, 285, pp.436 - 443.

Xing, J., Wellman-Labadie, O., Gautron, J., & Hincke, M. (2007). Recombinant eggshell ovocalyxin-32; expression, purification and biological activity of the glutathione S-transferase fusion protein. *Comparative Biochemistry and Physiology part B: Biochemistry and Molecular Biology*, *147*(2), 172-177.

Xu, L.-Q. et al., 2011. A 700-year record of mercury in avian eggshells of Guangjin Island, South China Sea. *Environmental Pollution*, 159, pp.889-896.

Yamane, T., Miller, D.L. & Hopfield, J.J., 1981. Discrimination between D- and L-tyrosine transfer ribonucleic acids in peptide chain elongation. *Biochemistry*, 20(25), pp.7059 - 7064.

Yergey, A. et al., 2002. De Novo sequencing of peptides using MALDI/TOF-TOF. *Journal of the American Society for Mass Spectrometry*, 13, pp.784-791.

Zong, W. et al., 2010. Influence of charge distribution on the discrepant MS/MS fragmentation of the native and oxidized FMRF: evidence for the mobile proton model. *Journal of Peptide Science*, 16, pp.687-692.