1

Biominerals and the Fossil Record

1.1 Why Study Old Biominerals?

Biomineralized tissues have a good chance of surviving in the fossil record and of preserving, like a geochemical time capsule, a snapshot of the environmental conditions in which the organism (or tissue) itself was formed. This has been known since the first decades of the 20th century and is the basis for reconstructing past variations in sea temperature by measuring the stable isotopic composition of fossil biominerals (e.g. shells, foraminifera; Emiliani, 1955). Studying the way in which biominerals are formed can thus reveal important information about our planet's past. At the same time, it can teach us how to build new materials with superior characteristics (so-called biomimetic and bioinspired materials).

However, fossil biominerals can also reveal the evolutionary history of life itself. Heinz Lowenstam, an intellectual giant of biomineralization, presented his famous table of biomineralized organisms (redrawn here as Table 1.1) at the 'Biogeochemistry of Amino Acids' meeting at Airlie House, Warrenton, Virginia, in 1978 (Lowenstam, 1980). Observing that the evolution of biomineralized organisms has its roots in deep time (~650–550 Ma ago), and that the majority of biominerals 'appear' during the Cambrian 'explosion', he suggested that: 'extended to the fossil record of the Phanerozoic, […] the study of the organic and bioorganic fractions [in biominerals] holds promise to trace pathways of biochemical evolution.' Lowenstam's proposal was not without substantiation: in 1978 it had already been known for a couple of decades that organic matter could be recovered from fossils (Abelson, 1954). This is due to the fact that the mineral component can protect the biological fraction (*biomolecules*) from degradation: after the death of an organism, a range of biological and chemical factors immediately induce the breakdown

Amino acids and Proteins in Fossil Biominerals: An Introduction for Archaeologists and Palaeontologists, First Edition. Beatrice Demarchi.

Table 1.1 Types of biogenic minerals and their occurrence in extant phyla, redrawn from Lowenstam (1980).

		PHOSPHATES								CARBONATES					Fe-SULFIDES		Mn-OXIDES	Fe-OXIDES						SILICA	SULFATES			OXALATES		HALIDES	
	PHYLUM	Amorph. Hydr. Ferric Phosphate	Amorph. Whitlockite Precursor	Amorph. Brushite Precursor	Amorph. Dahllite Precursor	Brushite	$Ca_3Mg_3(PO_4)_4$	Francolite	Dahllite	Amorphous Ca Carbonate	Monohydrocalcite	Vaterite	Aragonite	Calcite	Hydrotroilite	Pyrite	“Todokorite”	Amorph. Ferrihydr.	Ferrihydrite	Lepidocrocite	Goethite	Maghemite	Magnetite	Opal	Barite	Celestite	Gypsum	Weddelite	Whewellite	Amorph. Fluorite Precursor	Fluorite
MONERA									*		*		*	*	*	*	*		*			?	*								
PROTOCTISTA	*DINOFLAGELLATA*													*																	
	HAPTOPHYTA												?	*																	
	BACILLARIOPHYTA																							*							
	PHAETOPHYTA												*																		
	RHODOPHYTA											*	*	*																	
	CHLOROPHYTA												*																?		
	ZYGNEMATOPHYTA																												*		
	RHYZOPODEA																							?	*						
	SYPHONOPHYTA												*	*															*		
	CHAROPHYTA													*											*						
	HELIOZOATA																							*							
	RADIOLARIATA																							*		*					
	FORAMINIFERA												*	*				*						*							
	MYXOMYCOTA									*				*																	
	CILIOPHORA								*																						
FUNGI	*BASIDIOMYCOTA*																		*									*	*		
	DEUTEROMYCOTA								*										*									?	?		
ANIMALIA	*PORIFERA*												*	*						*				*							
	COELENTERATA												*	*													*				
	PLATYHELMINTHES		*							*				*																	
	ECTOPROCTA				*								*	*																	
	BRACHIOPODA							*						*																	
	ANELLIDA	*	*				*						*	*				*	*					*							
	MOLLUSCA	*	*	*	*	*		*		*	*	*	*	*					*	*	*		*	*				*		*	*
	ARTHROPODA				*					*		*	*	*									*	*					*		*
	SIPUNCULA												*	*																	
	ECHINODERMATA	*												*										*				*			*
	CHORDATA				*			*	*		*	*	*	*				*	*				*					*		*	
PLANTAE	*BRYOPHYTA*													*					*								?		*		
	TRACHAEPHYTA								?			*	*	*					*					*			*		*		

of the organic matter (soft tissues), but if these organics are somehow protected by the mineral phase, then degradation will be slow (or slower). It follows that ancient molecules (nucleic acids, proteins, lipids, carbohydrates) will be found mainly in sub-fossil (hereafter "fossil") mineralized substrates, and that we can target these for molecular palaeontology/archaeology.

At the time of writing, the field of palaeobiogeochemistry has expanded so much that we now use 'fossil' biomolecules routinely in order to recover information on the age since death of an organism (e.g. in the case of protein diagenesis/amino acid racemization dating; see Chapter 4), as well as on the physiology, phylogeny, and biogeography of extinct and extant organisms (from Neanderthals to crops), the evolution of diet, agricultural and husbandry practices, and even diseases. Indeed, we now have a whole dedicated field of 'biomolecular archaeology', which can be split into different subdisciplines (palaeogenomics, palaeoproteomics, palaeolipidomics, stable isotope geochemistry). Many recent reviews of the literature deal with the potential of new technologies for studying ancient biomolecules (Cappellini et al., 2018; Hendy et al., 2018; Welker, 2018).

And yet, these technological developments have not so far been able to fulfil Lowenstam's dream: in order to reconstruct evolutionary patterns in deep time, it is necessary to retrieve sequence data (DNA or proteins, bearing phylogenetic information), from *really* old fossils. Since DNA degrades more rapidly than proteins, the latter would be the molecules of choice. Despite repeated claims of preservation of proteins in Cretaceous (dinosaur) samples (Abelson, 1956; Miller and Wyckoff, 1968; Schweitzer et al., 1997; Asara et al., 2007; Cleland et al., 2015; Schroeter et al., 2017), which have been repeatedly dismissed (Collins et al., 2000; Buckley et al., 2008, 2017; Saitta et al., 2019), the oldest endogenous peptide sequences have been reported from Plio-Pleistocene fossils (eggshell: Demarchi et al., 2016; tooth enamel: Cappellini et al., 2019). In fact, comparatively little effort has gone into systematically assessing the survival potential of organic matrices other than those for bone, eggshell and teeth. (Incidentally, bone has been taking the lion's share of the effort, despite its being well known as a poor repository of endogenous molecules due to its porous and 'leaky' nature.) As a result, many scientific papers reporting protein sequence data bear two striking features: (1) the missed opportunity of linking research in ancient biomolecules with the latest theories and discoveries in the field of biomineralization; (2) the limited engagement with issues of diagenesis, which were instead at the forefront of debate at the time Lowenstam published his work.

In order to fill this gap, this book will focus on the mechanisms of degradation and survival of proteins in fossil biominerals, and look at the extent to which this affects the use of ancient proteins as a means of understanding the past. Applications for dating purposes, as well as in the newer field of 'palaeoproteomics', will be considered. But, before we begin, we must clarify what biominerals actually *are*.

1.2 What are Biominerals?

Biominerals are hybrid nanocomposite biomaterials, in which an organic fraction produces mineralized structures (skeletons) from inorganic precursor ions (Weber and Pokroy, 2015). Or, as Weiner and Dove define them, ‘where the distinctions between the chemical, biological and earth science disciplines melt away’ (Weiner and Dove, 2003). Biominerals display highly sophisticated architectures, exceptional mechanical properties and serve a variety of purposes, from protection of the soft body to scattering of light (Addadi and Weiner, 2014). Among the many studies which review research on biominerals, a recent historical overview has been provided by Addadi and Weiner (2014). This, like many others, highlights the fact that, despite the natural variety of biogenic products, their *chemical composition* is generally very similar: two-thirds of known biominerals are made of calcium carbonate (in its various polymorphs, i.e. calcite, aragonite, vaterite and, in some cases, amorphous) or calcium phosphate (e.g. bone apatite). Silica-based skeletons are also formed by some organisms, while other minerals, e.g. magnetite, are rarely found in biominerals (Lowenstam, 1981; Carter, 1989).

The organic fraction of biominerals is highly complex and contains, among other things, water, which is mainly present as fluid inclusions in the carbonate (Hudson, 1967; Gaffey, 1988). In shells, the amount of water has been found to be consistent within the same species but extremely variable between different taxa, most likely because these inclusion waters are vestiges of metabolic fluids produced by the tissues (Lécuyer and O’Neil, 1994). Other than water, the organic compounds present in biominerals are lipids (n-alkanes, free and bound cholesterol, and bound fatty acids: Stern, 1996), carbohydrates and proteins. Proteins represent the biggest fraction (~65–90%), followed by lipids (0.8–3%) and carbohydrates (0.2–0.3%), according to the estimates of Goulletquer and Wolowicz (1989). However, their study only compared three species of mollusc shell and it is likely that these percentages vary considerably, especially for carbohydrates: for example, chitin, one of the main carbohydrates in many biominerals, can represent up to 5–10% of the organic matrix. It is noteworthy that water is integral to the structure of organic molecules (up to 20% by weight in proteins), and that, as such, this hydration layer plays an important role in both the functioning and the degradation of the molecules themselves (see Chapter 3). In general, there is a striking lack of data on the composition of the organic fraction in different biominerals. Despite this poor knowledge, and notwithstanding the fact that it is commonly accepted that the three classes of biomolecules must all play a role in biomineralization (Farre and Dauphin, 2009; Luquet et al., 2012; Rao et al., 2014), it is mainly the protein fraction that has been studied.

The way in which inorganic and organic matrices are organized spatially in different biominerals is also still far from clear, despite the fact that the organic matrix within biogenic minerals was first observed around 50 years ago, using transmission electron microscopy (TEM). These early studies of the bivalves *Mercenaria*, *Mytilus* (Towe and Thompson, 1972) and *Pinctada*

(Nakahara and Bevelander, 1970) showed the presence of both a 'frothy' structure *within* single crystals and a 'bright layer' *between* crystalline units. This is a crucial observation which, supported by experimental data on the effect of a strong bleaching treatment on the organic matrix in shells (Crenshaw, 1972), allowed researchers to put forward the hypothesis that there were two pools of organics within biominerals: an intracrystalline matrix and an intercrystalline matrix. This concept was explored with regard to biomineralization and the organic–inorganic interface (Albeck et al., 1996) but went largely undetected in the field of geochemistry and especially protein diagenesis dating, as we will see in Chapter 4.

1.3 How and Why are Biominerals Formed?

Early work on biomineralized organisms, therefore, showed that building a mineralized skeleton involves a sophisticated balance between organic and inorganic matrices. In general, a high level of order can be found within the mineral skeleton, from nanometre to macroscopic length scales (Vielzeuf et al., 2008; Arakaki et al., 2015) although amorphous exceptions are known, e.g. in earthworm granules (Hodson et al., 2015). This extraordinary architecture is formed under physiological conditions, i.e. at temperatures and pressures that are lower than those required to produce the same materials by chemical synthesis (Wang and Nilsen-Hamilton, 2013). Since this implies an energetic cost, possessing a hard skeleton must be sufficiently advantageous as to warrant this investment (Lowenstam and Weiner, 1989; Knoll, 2003; Murdock and Donoghue, 2011). Indeed, the fact that this strategy pays off is shown by the fact that possessing a biomineralized skeleton or tissue is not a prerogative of a single lineage of organisms: corals, molluscs, turtles and humans are all able to build biogenic minerals (Table 1.1, adapted from Lowenstam, 1980; Figure 1, adapted from Murdock and Donoghue, 2011).

This peculiarity has captured the interest not only of biomineralization scientists, but also of evolutionary biologists, who are concerned with a fundamental question (Murdock and Donoghue, 2011): was there a common ancestor which developed the ability to mineralize a skeleton, or are we looking at independent evolution pathways for each lineage? The fact that biomineralizing groups are dispersed among other groups that do *not* build hard skeletons seems to suggest that the evolution of this ability was independent (Figure 1.1). Indeed, most biomineralized skeletons appear during the 'Cambrian explosion' as the result of the diversification of an original biomineralization 'episode' that had occurred during the (highly alkaline) Ediacaran (635–541 Ma ago). In the Ediacaran, most taxa were soft bodied, but some weakly calcified taxa also began to appear (*Cloudina*, *Namacalathus*, *Namapoikia*), supporting the idea of the diversification of an original feature, likely a 'biomineralization toolkit' (Livingston et al., 2006; Ramos-Silva et al., 2013; Wood et al., 2017).

But how can mineralization actually occur within biological systems? Early views based on classical nucleation theory posited that the early stages of

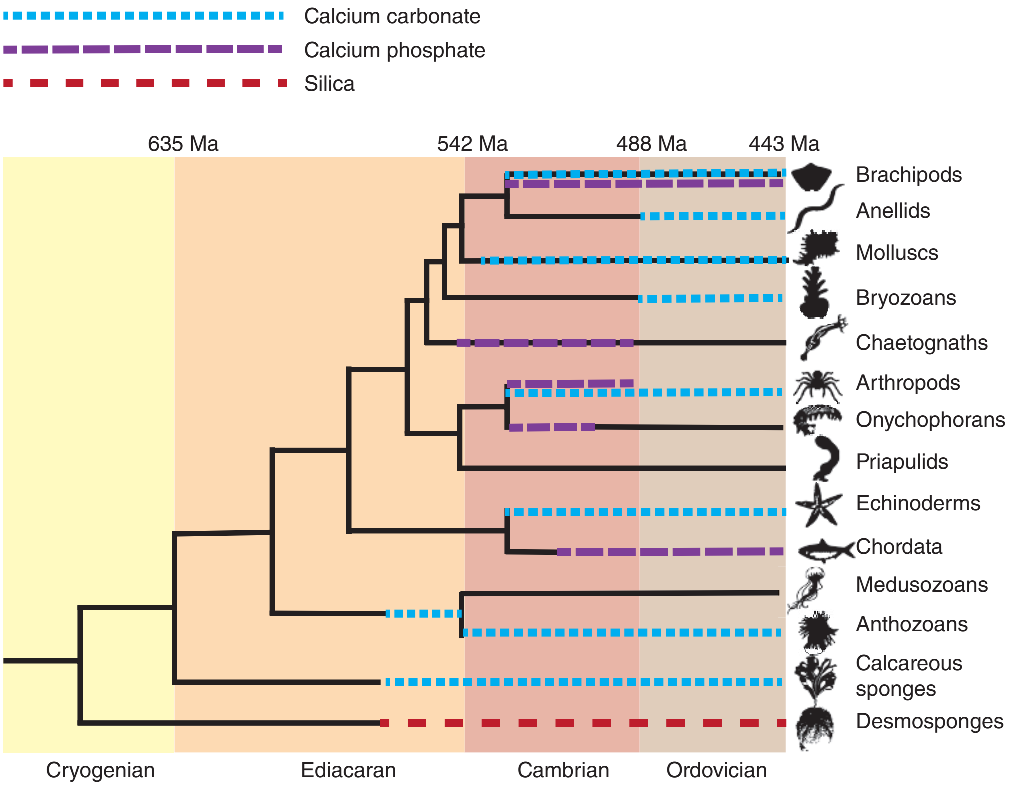

Figure 1.1 Evolutionary biology has had a longstanding interest in biomineralization. The phylogenetic tree shows that most biomineralized organisms appear in the Cambrian fossil record (i.e. the 'Cambrian explosion') and that mineralized skeletons are most frequently composed of calcium carbonate, calcium phosphate or silica. Adapted from Murdock and Donoghue (2011).

mineral formation are governed by random collisions between dissolved ions in supersaturated homogeneous solutions to form metastable clusters, which can then disintegrate further or grow (Gebauer et al., 2014). Recent mechanistic and experimental studies have shown that this is not possible; given the energetic cost involved in mineralization under physiological conditions, a biological catalyst is needed: the 'biomolecular toolkit'. This initiated the interaction between inorganic ions (calcium, carbonate, phosphate) and the extracellular matrix in the first biomineralized organisms ca. 550 Ma ago (Wood et al., 2017). It is likely that this toolkit used to be a relatively simple one but, after many million years of evolution, has morphed into an extremely complex 'molecular machinery', i.e. the organic matrix we find in biominerals today. This is so difficult to characterize fully that, as a result, many of the fundamental mechanisms controlling biogenic mineralization are still fairly obscure (De Yoreo et al., 2015). However, the exponential growth in the number of studies on biomineralization during the past three decades (Figure 1.2) has

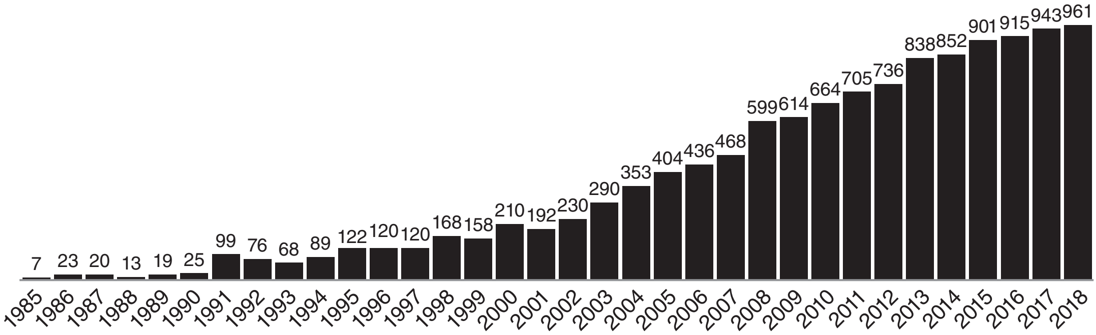

Figure 1.2 Number of articles with keyword 'biomin*' published between 1985 and 2018 (source: Web of Science).

so far allowed us to clarify three aspects of biologically controlled mineralization (Mann, 1983).

Biomineralization:

i is regulated by the genome;
ii results in the formation of mineral phases that are not in equilibrium with the environment;
iii is modulated by an organic matrix (extracellular) which is then entrapped in the mineral phase (Marin et al., 2012) and which can be recovered in biominerals of geological ages (Abelson, 1954).

In practice, this knowledge has changed our understanding of the way in which minerals are formed (Gebauer et al., 2008): new theories propose that, initially, ions assemble into stable amorphous clusters ('prenucleation clusters'), which can be conceived of as highly dynamic polymers of ion pairs. Subsequently, via a multi-step sequence, clusters are formed, followed by amorphous intermediates and, finally, crystals. Organic species, and especially proteins, can efficiently regulate many of these steps (Evans, 2013). In fact, the progress from ions (or atoms and molecules) to crystals can occur via a range of mechanisms, summarized recently by De Yoreo et al. (2015) (Figure 1.3).

In summary, biominerals are a diverse group of biological entities, which have a long evolutionary history, spanning at least 550 million years, and which display a fascinating set of coordinated biochemical mechanisms whereby macromolecules promote and regulate the growth of a mineralized skeleton. Proteins in particular play a major role in this process. This interplay not only confers exceptional material properties to the biogenic mineral – a feature which modern engineers strive to replicate in the creation of biomimetic and bioinspired materials – but also guarantees it a better chance of survival post mortem, in the fossil record.

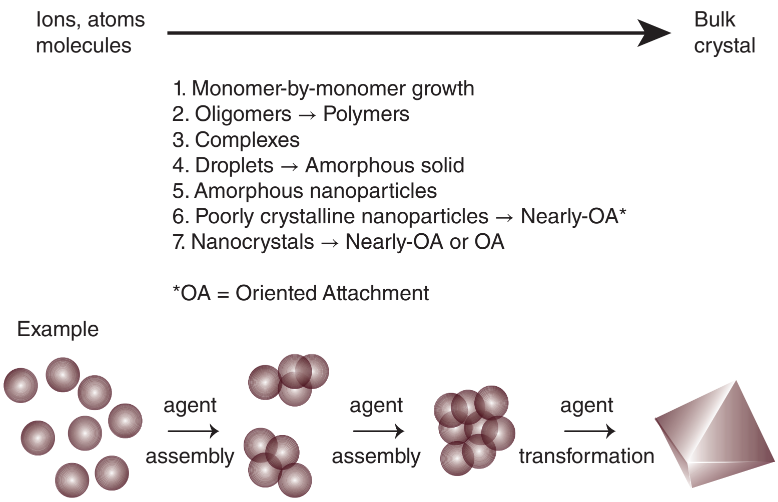

Figure 1.3 Schematic representation of nonclassical nucleation, adapted from Evans (2013) and De Yoreo et al. (2015).

1.4 'Biomineralization Toolkit': From Proteins to Proteomes

The fact that proteins are an important (but not the only) part of the biomineralization toolkit has been known for at least two decades (Sigel et al., 2008) and early biomolecular research efforts in the field were directed towards isolating, purifying and sequencing these molecules from a range of biominerals in order to clarify just how similar the 'biomineralization toolkit' is across various phyla. The first organisms and tissues to be investigated were mollusc shell (especially taxa producing pearls), avian eggshell and animal bone/tooth, due to their commercial and medical relevance. The study of these first sequences showed that, at least in mollusc shell, there were typically two subsets of organic matrices: the acid-soluble and the acid-insoluble (protein, proteoglycan and chitin) matrices (Miyamoto et al., 1996; Sarashina and Endo, 1998; Mann et al., 2000; Weiss et al., 2000; Marin et al., 2007; Marie et al., 2008). Furthermore, these early analyses showed that different proteins could be isolated from different microstructures and mineralogies (e.g. prisms and nacre/mother-of-pearl in mollusc shells) (Marie et al., 2012). Some similarities also began to emerge: C-lectin type proteins were isolated and purified from *Haliotis* shell (Weiss et al., 2000), avian eggshell (Mann and Siedler, 1999, 2004; Mann, 2004), sea urchins (Illies et al., 2002); highly acidic proteins were found commonly in mollusc shells (Gotliv et al., 2005; Marie et al., 2007) and a large percentage of them are intrinsically disordered proteins (IDPs), which are thought to be important for initiating biomineralization (Boskey and Villarreal-Ramirez, 2016).

Knowledge of molecular sequences is not sufficient for understanding the way in which these proteins (and their three-dimensional structures) interact with inorganic ions, clusters, amorphous precursors and crystals throughout the process of mineralization. Mathematical modelling can be used effectively to simulate the organic–inorganic interactions at each stage and is therefore able to elucidate the role of a certain biomolecule. An example is the study of the chicken eggshell C-lectin Ovocleidin-17 (OC-17), which showed that this molecule can bind to the carbonate ions of amorphous calcium carbonate via basic (positively charged) amino acid residues and is then released from the surface of the formed calcite crystal to start its catalytic cycle once more (Freeman et al., 2011). An important proviso of these types of computational studies is that not only the protein sequence, but also its structure, needs to be known.

An alternative strategy to computer simulation is the experimental approach, which often involves purifying a protein and then observing the behaviour of the mineral phase in vitro (Falini et al., 1996). It is also possible to create synthetic proteins (or peptides) with domains of known primary sequence and structural properties, e.g. rigid lectin-like domains versus intrinsically disordered. For example, a study of the three domains of the sea urchin spicule matrix protein SM50, using a small ubiquitin-like modifier (SUMO) fusion protein as a model, has shown that the three domains are necessary and sufficient to allow the formation and growth of the sea urchin spicule (Rao et al., 2013). In an earlier study, recombinant nacrein proteins with and without carbonic anhydrase and Gly–X–Asn domains were tested in order to elucidate nacrein's role as a negative regulator of calcification (Miyamoto et al., 2005).

Recent efforts in molecular biology and the sequencing of hundreds of new organisms, including biomineralizing taxa, have shown that the number and variety of proteins trapped in even apparently simple systems (e.g. avian eggshell) is such that a single-molecule approach is no longer viable. As such, it would be more correct to refer to the proteins pertaining to biominerals as 'biomineralized proteomes'. Chapter 3 briefly summarizes the main protein families found in biominerals. It is worth mentioning here, though, a recent review of the 'secretome' of marine mollusc shells by Kocot et al. (2016), which highlights the extraordinary diversity of biomolecules secreted by the molluscs' mantle cells in order to modulate the calcification of the mineral matrix, including mechanical and aesthetic characteristics (e.g. hardness, colour). Interestingly, it appears that most proteins involved in biomineralization (not only in shells but also in other animals) display repetition of low complexity domains (RLCDs), which are easily shuffled, lost or acquired in evolutionary processes (Kocot et al., 2016).

The quest for the universal biomineralization toolkit is at a crucial stage, with more and more proteins being identified and their function elucidated by homology. It is clear that there is an interplay between different proteins and/or different protein domains, and that the study of single molecules can only give a partial picture of this complex molecular machinery. It is also clear that only some of the proteins belonging to the proteome perform a mineralization-related task. There is need, however, for a novel strategy able to select protein sequences that are truly relevant for biomineralization and distinguish them from those which only 'happen' to be incorporated in biominerals. Such an approach might stem from the study of fossil proteomes, i.e. the surviving proteins within (sub)fossil biominerals.

Box 1.1 Terminology

Amino acids are aminoalkanoic acids, each containing a hydrogen atom –H, an amino group $-NH_2$, a carboxylic group –COOH and an –R group, or side-chain, bound to a central carbon atom –C. The amine and carboxylic acid functional groups mean that amino acids can act both as an acid and as a base. The –R group is specific to each of the 20 common, standard or coded amino acids. Amino acids are usually referred to by trivial names (e.g. alanine), but a three-letter abbreviation system is commonly used (e.g. Ala) for amino acids which make up peptides and proteins, as well as a one-letter system (e.g. A) for identifying peptides through a string of letters.

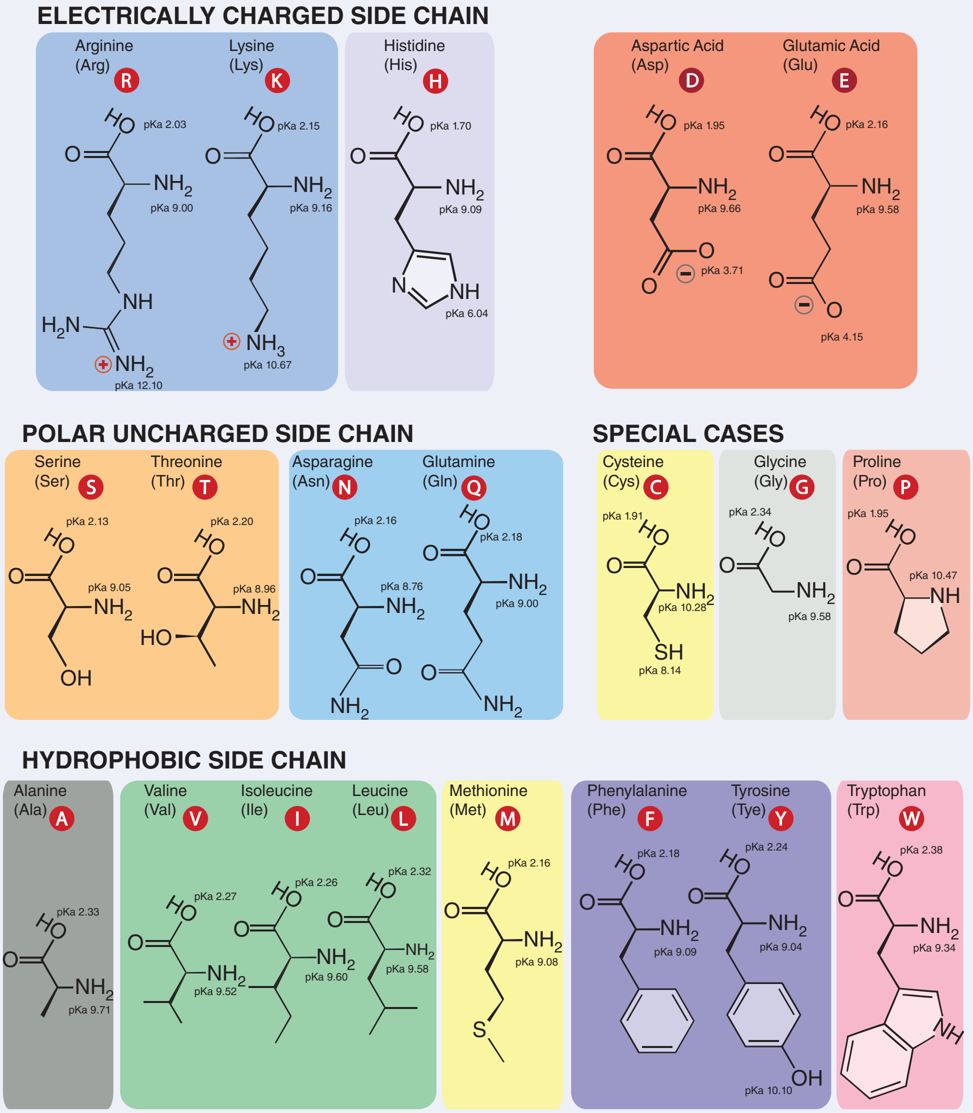

Image modified from the original at https://commons.wikimedia.org/wiki/File:Amino_Acids.svg and shared under a Creative Commons Attribution-Share Alike 3.0 Unported licence.

Peptides are formed by condensation of the respective amino and carboxy groups of two different amino acids, with the elimination of a water molecule and the formation of an amino linkage (–CO–NH–) (also known as a peptide bond). Generally, peptides are considered to be between two and twenty amino acids long (Barrett and Elmore, 1998). Proteins are polymers of amino acids with molecular weights in the 1–1000 kDa region, although the size boundary between peptide and protein is not absolute: there are some very small proteins (osteocalcin, 100 residues long; insulin, 51 residues) and also some very long peptides (amyloid beta, 36–43 residues). Proteins and peptides generally retain an unreacted amino group at one end (N terminus) and an unreacted carboxyl group at the other (C terminus) and, at neutral pH, each terminus carries an ionic charge. Conventionally, amino acid residues in peptide chains are numbered from the N terminus to the C terminus. They are also written from left to right, since this is the direction followed during biosynthesis.

The number, type and sequence of amino acids in the chain determines the *primary* structure of the protein, while the three-dimensional configuration of the peptide chains in space is called the *secondary* structure (e.g. alpha-helices, beta-strands and beta-sheets). Further interactions and coiling through ionic interactions, hydrogen bonds, van der Waals forces or sulfur bridges can bring distant portions of the primary and secondary structures close together, thus determining the formation of a *tertiary* structure. The *quaternary* structure involves the association of two or more polypeptide chains into a multi-subunit, or oligomeric, protein (Barrett and Elmore, 1998). Crucial for the diversification of the protein's functions are post-translational modifications (PTMs): functional groups can be added (as in phosphorylation, methylation, glycosylation, acetylation) or parts of the protein can be degraded (for example enzymatically), in order to achieve the normal functioning of the living cell.

The source of information on nucleotide and protein sequences is the NCBI (National Center for Biotechnology Information: https://www.ncbi.nlm.nih.gov/) repository, which is also linked to the protein databank (PDB: http://www.rcsb.org/pdb), in which any protein structure that has been solved is available. It is important to note that proteins enclosed in biominerals are often not well characterized, mainly due to missing or incomplete genomic information for the biomineralizing organism. In the example below, the cDNA sequence of OC-17, a major biomineralization protein (Lakshminarayanan et al., 2002; Freeman et al., 2011) which had been sequenced directly in 1999 and the structure of which has been known since 2004 (Mann and Siedler, 1999; Reyes-Grajeda et al., 2004), was missing from the chicken genome and had to be determined using transcriptome assembly and cloning in 2014 (Zhang et al., 2014).

cDNA sequence of Ovocleidin-17, a major biomineralization protein from chicken eggshell (Zhang et al., 2014)

1 GGTGCCGGGA CGCGGAACGG CCGCCCGAAC GGGAACGCAA TGGCACCGAC GTGGGCGCTG
61 CTGGGCTGCG TTCTGCTGCT CCCCTCCCTG CGAGGGGATC CGGACGGCTG CGGCCCGGGT
121 TGGGTGCCGA CCCCCGGCGG CTGCCTCGGC TTCTTCAGCC GGGAGCTCAG CTGGAGCCGC
181 GCCGAGTCGT TCTGCCGCCG TTGGGGTCCC GGTTCCCACC TGGCGGCGGT GCGCAGCGCG
241 GCGGAGCTGC GGCTCCTCGC GGAGCTCCTC AACGCGTCGC GGGGCGGCGA CGGCAGCGGG
301 GAGGGGGCGG ACGGCCGCGT CTGGATCGGC CTCCACCGCC CCGCCGGGAG CCGTTCGTGG
361 CGGTGGTCGG ACGGCACCGC GCCGCGCTTC GCTTCGTGGC ACCGAACGGC CAAAGCGCGG
421 CGGGGGGGGC GGTGCGCGGC GCTGCGGGAC GAGGAGGCCT TCACCTCGTG GGCCGCCCGG
481 CCGTGCACAG AGCGCAATGC CTTCGTCTGC AAAGCCGCCG CCTGAATGGA CAACAACACA
541 ACAACACAAC AACACAACAA CGCAACAACG CAACAACGCA ACAACGCAAC GACCCCCAAC
601 ACTGCAATAA ACGGACCCAC AGCAGC

Primary (amino acid) sequence of OC-17 (Mann and Siedler, 1999)

1 MAPTWALLGC VLLLPSLRGD PDGCGPGWVP TPGGCLGFFS RELSWSRAES FCRRWGPGSH
61 LAAVRSAAEL RLLAELLNAS RGGDGSGEGA DGRVWIGLHR PAGSRSWRWS DGTAPRFASW
121 HRTAKARRGG RCAALRDEEA FTSWAARPCT ERNAFVCKAA A

Secondary (alpha-helix, beta-sheets and random coil configurations represented in red, yellow and light grey, respectively) and tertiary structure of OC-17 (Reyes-Grajeda et al., 2004).

1.5 Fossil Biominerals, Fossil Proteomes

Biomineralized organisms have a greater chance of entering and being preserved in the fossil record than do soft tissues, to the point that the typical 'fossil' is almost shorthand for 'fossilized skeletal tissue' of a variety of organisms (the ones classified in Table 1.1). The importance of these fossils for archaeology and earth sciences is obvious: they allow us to reconstruct the past, from deep geological times (let us just think of the 'index' or 'guide' fossils used to define and identify a geological period) until more recently (when clade *Homo* enters the world scene in the Quaternary, see Chapter 4).

The survival of organic matter in 'fossil' organisms is a discovery that dates back to the 19th century (Schaffer, 1889). Identification of these compounds as individual amino acids (Abelson, 1954), and of their chiral properties, led to the consequent development of a relative dating technique (amino acid racemization, AAR). The field of amino acid biogeochemistry grew very rapidly, reaching a pinnacle with regard to the integration of geochronological applications, biogeochemistry and biomineralization at a symposium held in 1978 in Virginia. The papers are collected in the text 'Biogeochemistry of amino acids' (Hare et al., 1980), which (remarkably) is still today an exceptional source of knowledge (or indeed of open questions) for palaeobiogeochemists.

Subsequent advances focused specifically on the potential use of different biominerals as substrates for protein diagenesis and geochemistry. This potential derives not only from the state of preservation, but also from the specific characteristics of the biomineral. In particular, biominerals can be classified into three broad categories with regard to their (generalized) behaviour towards protein diagenesis: open-system biominerals, closed-system biominerals and biominerals which may contain a closed system.

1.5.1 Open and Closed Systems: Isolating the Intracrystalline Organic Fraction in the Laboratory

Most biominerals display a complex microstructure (crystals take the form of prisms, platelets, cones, lamellae or needles) within which the organic framework is interspersed in a variety of ways. Despite this array of microstructural arrangements, the organic component can be classified more simply in terms of its spatial relationship with the crystals, whatever shape these might be. In the 1960s and 1970s amino acid geochronologists (including Hare, Wehmiller, King and Ritter) discussed the idea that the organic framework had components of different mobility (Wehmiller, 1980). Sykes et al. (1995) put forward a model that theorized the existence of two pools of proteins: one '*intercrystalline*', which is located *between* mineral crystals and shows an open-system behaviour (i.e. leaching of amino acids out of the system, poor reproducibility of the data, contamination issues); the second '*intracrystalline*', which is somewhat trapped *within* the crystals, can be isolated via a chemical treatment (strong oxidation, usually by bleaching with sodium hypochlorite, NaOCl) and can behave as a closed system (Collins and Riley, 2000; Penkman et al., 2008; Bright and Kaufman, 2011; Demarchi et al., 2013). Figure 1.4 shows this theoretical model in which an 'operational intracrystalline fraction' can be isolated.

In reality, of course, things are more complex, but recent visualization studies performed using a variety of techniques based on electron microscopy and tomography (Robach et al., 2005; Gries et al., 2009; Gordon and Joester, 2011; Li et al., 2011; Younis et al., 2012; Hendley et al., 2015; Liu et al., 2015, 2017) show that Towe and Thompson's early work (Towe and Thompson, 1972) was accurate and that the model by Sykes and colleagues stands. The main characteristics of the organic intracrystalline matrix thus revealed are:

- it is embedded within *single crystals*, i.e. biogenic crystals fracture and diffract as single crystals (see a review of evidence in Weber and Pokroy, 2015);
- the occlusions have different shapes (voids, islets, chain-like, haze-like) at the nanometre scale (for a summary see Liu et al., 2017);
- the protein component of the intracrystalline matrix tends to be highly acidic and often glycosylated (Marin et al., 2007);
- the conformation of these proteins tends to be open, unfolded, with 'intrinsically disordered domains' (Boskey and Villarreal-Ramirez, 2016).

In fossil biominerals the intercrystalline organic matrix, which is not contained within occlusions, will be the first to succumb to the effects of time and of the burial environment, and even if its degradation is somewhat halted, this fraction is likely to contain exogenous compounds: if a biomineral contains mainly intercrystalline proteins, it will display an open-system behaviour. Open-system biominerals are prone to contamination and leaching of amino acids from the system (Brooks et al., 1990; Sykes et al., 1995), and the proteins

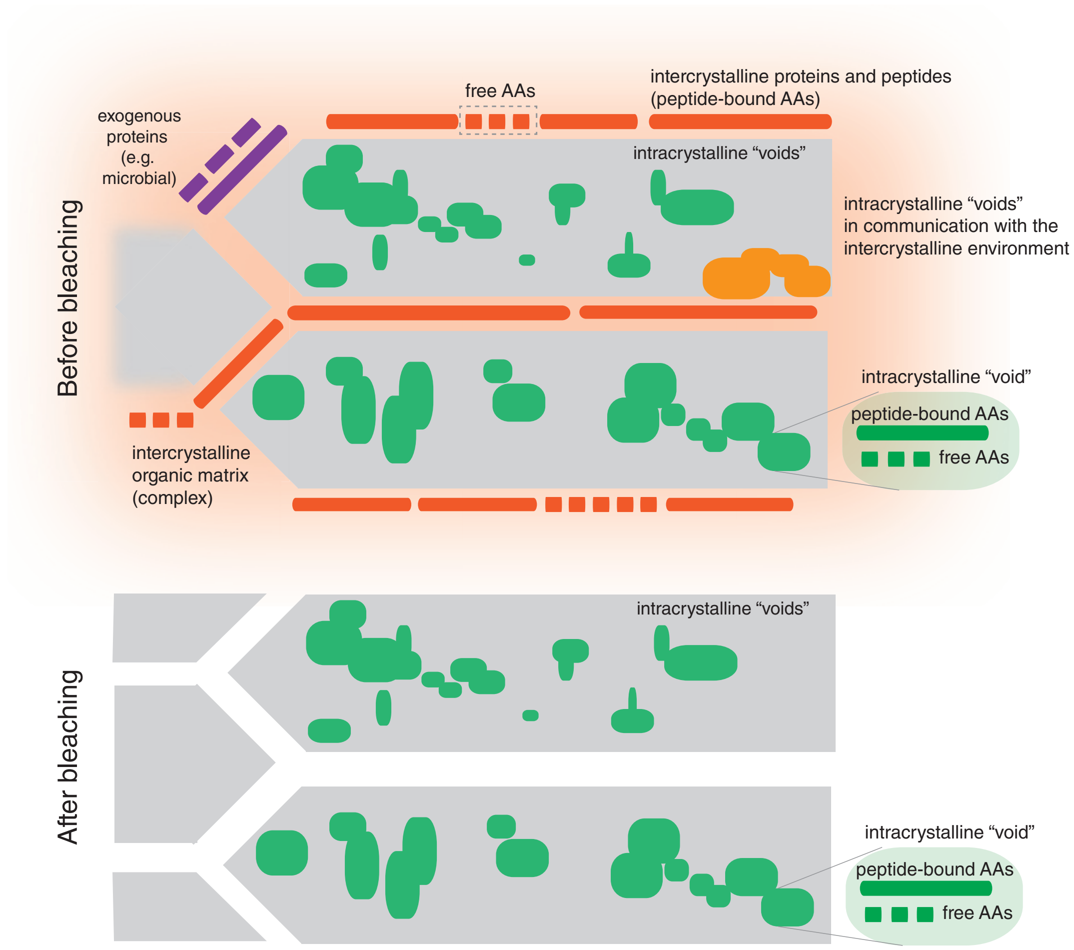

Figure 1.4 The intracrystalline fraction of proteins (peptide-bound and free amino acids) is occluded in nanometric 'voids' within the crystal. The morphology of the occlusions is variable: voids, islets, chain-like, haze-like shapes (Liu et al., 2015). The intercrystalline organic matrix also contains peptide-bound and free amino acids (as well as other compounds formed with the nonproteinaceous components of the matrix, not shown in the figure, and contamination from exogenous proteins) but these can be removed by a strong oxidation pretreatment (e.g. bleaching). This treatment leaves behind only the original proteins, degraded in situ. Modified after Sykes et al. (1995).

contained in an open system are not protected by the action of environmental factors such as pH, presence of water, effect of clay surfaces, of buffers, of metal ions (Smith and Evans, 1980). One example of an open-system biomineral is bone, since the individual mineral elements are too small to encapsulate proteins. In the case of bone, it may be that proteins become trapped in mineral aggregates (Salamon et al., 2005), though this has not yet been the focus of detailed studies. The fact that the organic matrix within bones behaves as an open system can be a real hindrance to archaeological and palaeontological applications since, without in-depth analyses, it is extremely difficult to assess whether we are looking at the original indigenous molecules, or at modern (or ancient) contamination. This is especially problematic when bulk amino acids are extracted from the bone matrix, typically for measuring the extent of racemization for geochronological purposes (e.g. Bada, 1985; see also Chapter 4). Later applications of ancient DNA and palaeoproteomics have been successful in extracting biomolecular *sequences*, which can be authenticated and distinguished from contamination (e.g. Ostrom et al., 2000; Höss et al., 1996), thus reducing the impact of the open-system behaviour of bone.

Conversely, early work on avian eggshells (especially ratite) postulated that they retain both original organic matter and degradation products during diagenesis (Brooks et al., 1990), and avian eggshell has been classed as a closed system. However, later research demonstrated that even avian eggshell is more akin to systems in which a *fraction* of proteins behaving as a closed system can be isolated from the complex organic matrix (Crisp et al., 2013). It is likely that the extremely rapid formation of the eggshell in utero (~20 hours for chicken eggshell, see Figure 3.3) means that a large proportion of proteins is trapped within the large eggshell crystals; this happens fairly indiscriminately, as indicated by the complexity of the eggshell proteome (Mann, 2015; Demarchi et al., 2016). Eggshell has been the traditional target for amino acid racemization dating from the 1990s onwards, and more recent studies in palaeoproteomics have demonstrated that eggshell offers an ideal substrate for exceptional preservation (Demarchi et al., 2016), but also for interesting archaeological insights into the relationship between humans and their environment (Stewart et al., 2014; Jonuks et al., 2018; Demarchi et al., 2019).

Systematic investigation of a range of biominerals has shown that a functionally defined intracrystalline *fraction* can be isolated from many taxa of mollusc shells, corals, brachiopods and ostracods (Walton, 1998; Penkman et al., 2008, 2011; Bright and Kaufman, 2011; Hendy et al., 2012; Demarchi et al., 2013; Bosch et al., 2015; Pierini et al., 2016). Unlike eggshell, these tissues and organisms are formed over longer time spans, even tens of hundreds of years: for example, the oldest animal in the world was an *Arctica islandica* bivalve mollusc, which was 'fished' from the seafloor when it was 507 years old, thus making it the world's oldest animal (Butler et al., 2013). These organisms, therefore, are much more complex, because they will include 'old' proteins due to lack of tissue turnover, as well as 'young'.

The isolation of the intracrystalline fraction in the laboratory, despite taking many decades to become standard procedure, is in practice a very simple step (see Chapter 4): the biomineral is cleaned (whole or in fragments, depending on its size) by ultrasonication in water or by briefly soaking in a weak acid (e.g. acetic acid), base (e.g. sodium hydroxide), or calcium-chelating agent (ethylenediaminetetraacetic acid). For shells, it is common to separate different microstructural layers by abrasion (Marie et al., 2012), after which the selected mineral fraction is ground and sieved until a homogeneous particle size (usually <500 μm) is obtained. The dried powders are submerged in an excess of oxidizing agent (e.g. NaOCl, 12% wt/vol), often helping the bleaching process by gently agitating the suspension.

Powders are rinsed in water and methanol (e.g. Penkman et al., 2008) until all the bleach has been removed. The bleached powders should theoretically only contain the intracrystalline proteins embedded in the mineral. The next step is therefore to extract and characterize these proteins and use them for a variety of applications in archaeology and earth sciences. However, the results vary slightly according to the experimental settings used (e.g. the concentration of the oxidant, the duration of the treatment, and the optimal particle size need to be tested for each biominerals) and this indicates that the intracrystalline fraction is indeed not a perfect closed system (nor is always the same in all biological systems, even for two shells of the same species). Therefore, the operational definition usually refers to crushed biomineral (around 500 µm particle size) exposed to 12% wt/vol NaOCl for 48 hours (Penkman et al., 2008).

1.5.2 'Diagenetically isolated' intracrystalline proteins

The concept of intracrystalline proteins has proven to be extremely powerful both for biomineralization studies (Drake et al., 2013; Ramos-Silva et al., 2013) and for geochronology/palaeoproteomics (Chapters 4 and 5). In fossils, however, the situation is complicated due to the matter of degradation: entombed proteins cannot be expected to survive the effects of time untouched (even permafrost-buried remains degrade, albeit slowly). This complication is, at the same time, a positive factor: diagenesis acts as an effective agent for isolating the intracrystalline fraction, by removing the more 'mobile' components and leaving behind a stable system (see e.g. Wehmiller and Miller, 2000). Therefore, the many studies that used geological and archaeological samples for dating purposes from the 1960s onwards without a bleaching step were successful *because the analyses effectively targeted a closed system* (Chapter 4). Re-evaluating those early datasets, preferably those obtained from well-understood stratigraphic sequences, by analysing bleached and unbleached samples, would be an excellent way to gain a better understanding of the behaviour of different fractions of proteins.

As a final remark to this chapter, it is important to highlight that, in general, the information we retrieve from fossil biominerals will be biased towards the most degradation-resistant molecules. Therefore, it is crucial that the mechanisms of diagenesis and preservation are understood before embarking on a discussion of the potential applications.

References

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

Abelson, P.H. (1956). Paleobiochemistry. *Sci. Am.* 195: 83–96.

Addadi, L. and Weiner, S. (2014). Biomineralization: Mineral formation by organisms. *Phys. Scr.* 89: 098003.

Albeck, S., Addadi, I., and Weiner, S. (1996). Regulation of calcite crystal morphology by intracrystalline acidic proteins and glycoproteins. *Connect. Tissue Res.* 35: 365–370.

Arakaki, A., Shimizu, K., Oda, M. et al. (2015). Biomineralization-inspired synthesis of functional organic/inorganic hybrid materials: Organic molecular control of self-organization of hybrids. *Org. Biomol. Chem.* 13: 974–989.

Asara, J.M., Schweitzer, M.H., Freimark, L.M. et al. (2007). Protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry. *Science* 316: 280–285.

Bada, J.L. (1985). Aspartic acid racemization ages of California Paleoindian Skeletons. *Am. Antiq.* 50: 645–647.

Barrett, G.C. and Elmore, D.T. (1998). *Amino Acids and Peptides*. Cambridge University Press.

Bosch, M.D., Mannino, M.A., Prendergast, A.L. et al. (2015). New chronology for Ksâr 'Akil (Lebanon) supports Levantine route of modern human dispersal into Europe. *Proc. Natl. Acad. Sci. U.S.A.* 112: 7683–7688.

Boskey, A.L. and Villarreal-Ramirez, E. (2016). Intrinsically disordered proteins and biomineralization. *Matrix Biol.* 52–54: 43–59.

Bright, J. and Kaufman, D.S. (2011). Amino acid racemization in lacustrine ostracodes, part I: Effect of oxidizing pre-treatments on amino acid composition. *Quat. Geochronol.* 6: 154–173.

Brooks, A.S., Hare, P.E., Kokis, J.E. et al. (1990). Dating Pleistocene archeological sites by protein diagenesis in ostrich eggshell. *Science* 248: 60–64.

Buckley, M., Walker, A., Ho, S.Y.W. et al. (2008). Comment on 'Protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry.' *Science* 319: 33.

Buckley, M., Warwood, S., van Dongen, B. et al. (2017). A fossil protein chimera; difficulties in discriminating dinosaur peptide sequences from modern cross-contamination. *Proc. R. Soc. B* 284: 20170544.

Butler, P.G., Wanamaker, A.D., Scourse, J.D. et al. (2013). Variability of marine climate on the North Icelandic Shelf in a 1357-year proxy archive based on growth increments in the bivalve *Arctica islandica*. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 373: 141–151.

Cappellini, E., Prohaska, A., Racimo, F. et al. (2018). Ancient biomolecules and evolutionary inference. *Annu. Rev. Biochem.* 87: 1029–1060.

Cappellini, E., Welker, F., Pandolfi, L., and Madrigal, J.R. (2019). Early Pleistocene enamel proteome from Dmanisi resolves Stephanorhinus phylogeny. *Nature* 574: 103–107.

Carter, J.G. (ed.) (1989). *Skeletal Biomineralization: Patterns, Processes and Evolutionary Trends*, vol. 5. American Geophysical Union.

Cleland, T.P., Schroeter, E.R., Zamdborg, L. et al. (2015). Mass spectrometry and antibody-based characterization of blood vessels from *Brachylophosaurus canadensis*. *J. Proteome Res.* 14: 5252–5262.

Collins, M.J. and Riley, M.S. (2000). Amino acid racemization in biominerals: The impact of protein degradation and loss. In: *Perspectives in Amino Acid and Protein Geochemistry* (eds. G. Goodfriend, M.J. Collins, M. Fogel, et al.), 120–141. Oxford University Press.

Collins, M.J., Gernaey, A.M., Nielsen-Marsh, C.M. et al. (2000). Slow rates of degradation of osteocalcin: Green light for fossil bone protein? *Geology* 28: 1139–1142.

Crenshaw, M.A. (1972). The soluble matrix from *Mercenaria mercenaria* shell. *Biomineralization* 6: 6–11.

Crisp, M., Demarchi, B., Collins, M. et al. (2013). Isolation of the intra-crystalline proteins and kinetic studies in *Struthio camelus* (ostrich) eggshell for amino acid geochronology. *Quat. Geochronol.* 16: 110–128.

Demarchi, B., Rogers, K., Fa, D.A. et al. (2013). Intra-crystalline protein diagenesis (IcPD) in *Patella vulgata*. Part I: Isolation and testing of the closed system. *Quat. Geochronol.* 16: 144–157.

Demarchi, B., Hall, S., Roncal-Herrero, T. et al. (2016). Protein sequences bound to mineral surfaces persist into deep time. *eLife,* 5: e17092.

Demarchi, B., Presslee, S., Gutiérrez-Zugasti, I. et al. (2019). Birds of prey and humans in prehistoric Europe: A view from El Mirón Cave, Cantabria (Spain). *J. Archaeol. Sci. Rep.* 24: 244–252.

De Yoreo, J.J., Gilbert, P.U.P.A., Sommerdijk, N.A.J.M. et al. (2015). Crystal growth. Crystallization by particle attachment in synthetic, biogenic, and geologic environments. *Science* 349: aaa6760.

Drake, J.L., Mass, T., Haramaty, L. et al. (2013). Proteomic analysis of skeletal organic matrix from the stony coral *Stylophora pistillata*. *Proc. Natl. Acad. Sci. U.S.A.* 110: 3788–3793.

Emiliani, C. (1955). Mineralogical and chemical composition of the tests of certain Pelagic Foraminifera. *Micropaleontology* 1: 377–380.

Evans, J.S. (2013). 'Liquid-like' biomineralization protein assemblies: A key to the regulation of non-classical nucleation. *CrystEngComm* 15: 8388.

Falini, G., Albeck, S., Weiner, S. et al. (1996). Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science* 271: 67–69.

Farre, B. and Dauphin, Y. (2009). Lipids from the nacreous and prismatic layers of two Pteriomorpha Mollusc shells. *In: EGU General Assembly Conference Abstracts* 11: 5447.

Freeman, C.L., Harding, J.H., Quigley, D., and Rodger, P.M. (2011). Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation. *J. Phys. Chem. C* 115: 8175–8183.

Gaffey, S.J. (1988). Water in skeletal carbonates. *J. Sediment. Res.* 58: 397–414.

Gebauer, D., Völkel, A., and Cölfen, H. (2008). Stable prenucleation calcium carbonate clusters. *Science* 322: 1819–1822.

Gebauer, D., Kellermeier, M., Gale, J.D. et al. (2014). Pre-nucleation clusters as solute precursors in crystallisation. *Chem. Soc. Rev.* 43: 2348–2371.

Gordon, L.M. and Joester, D. (2011). Nanoscale chemical tomography of buried organic-inorganic interfaces in the chiton tooth. *Nature* 469: 194–197.

Gotliv, B.-A., Kessler, N., Sumerel, J.L. et al. (2005). Asprich: A novel aspartic acid-rich protein family from the prismatic shell matrix of the bivalve *Atrina rigida*. *Chembiochem* 6: 304–314.

Goulletquer, P. and Wolowicz, M. (1989). The shell of *Cardium edule*, *Cardium glaucum* and *Ruditapes philippinarum*: Organic content, composition and energy value, as determined by different methods. *J. Mar. Biol. Assoc. U.K.* 69: 563–572.

Gries, K., Kröger, R., Kübel, C. et al. (2009). Investigations of voids in the aragonite platelets of nacre. *Acta Biomater.* 5: 3038–3044.

Hare, P.E., Hoering, T.C., and King, K. (eds.) (1980). *Biogeochemistry of Amino Acids*. New York: Wiley.

Hendley, C.T., Tao, J., Kunitake, J.A.M.R. et al. (2015). Microscopy techniques for investigating the control of organic constituents on biomineralization. *MRS Bulletin* 40: 480–489.

Hendy, E.J., Tomiak, P.J., Collins, M.J. et al. (2012). Assessing amino acid racemization variability in coral intra-crystalline protein for geochronological applications. *Geochim. Cosmochim. Acta* 86: 338–353.

Hendy, J., Welker, F., Demarchi, B. et al. (2018). A guide to ancient protein studies. *Nat. Ecol. Evol.* 2: 791–799.

Hodson, M.E., Benning, L.G., Demarchi, B. et al. (2015). Biomineralisation by earthworms: An investigation into the stability and distribution of amorphous calcium carbonate. *Geochem. Trans.* 16: 4.

Höss, M., Jaruga, P., Zastawny, T.H. et al. (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Res.* 24: 1304–1307.

Hudson, J.D. (1967). The elemental composition of the organic fraction, and the water content, of some recent and fossil mollusc shells. *Geochim. Cosmochim. Acta* 31: 2361–2378.

Illies, M.R., Peeler, M.T., Dechtiaruk, A.M., and Ettensohn, C.A. (2002). Identification and developmental expression of new biomineralization proteins in the sea urchin *Strongylocentrotus purpuratus*. *Dev. Genes Evol.* 212: 419–431.

Jonuks, T., Oras, E., Best, J., and Demarchi, B. (2018). Multi-method analysis of avian eggs as grave goods: Revealing symbolism in conversion period burials at Kukruse, NE Estonia. *Environ. Archaeol.* 23: 109–122.

Knoll, A.H. (2003). Biomineralization and evolutionary history. *Rev. Mineral. Geochem.* 54: 329–356.

Kocot, K.M., Aguilera, F., McDougall, C. et al. (2016). Sea shell diversity and rapidly evolving secretomes: Insights into the evolution of biomineralization. *Front. Zool.* 13: 23.

Lakshminarayanan, R., Kini, R.M., and Valiyaveettil, S. (2002). Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix. *Proc. Natl. Acad. Sci. U.S.A.* 99: 5155–5159.

Lécuyer, C. and O'Neil, J.R. (1994). Stable isotope compositions of fluid inclusions in biogenic carbonates. *Geochim. Cosmochim. Acta* 58: 353–363.

Li, H., Xin, H.L., Kunitake, M.E. et al. (2011). Calcite prisms from mollusk shells (*Atrina rigida*): Swiss-cheese-like organic–inorganic single-crystal composites. *Adv. Funct. Mater.* 21: 2028–2034.

Liu, C., Xie, L., and Zhang, R. (2015). Heterogeneous distribution of dye-labelled biomineralizaiton proteins in calcite crystals. *Sci. Rep.* 5: 18338.

Liu, C., Du, J., Xie, L., and Zhang, R. (2017). Direct observation of nacre proteins in the whole calcite by super-resolution microscopy reveals diverse occlusion patterns. *Cryst. Growth Des.* 17: 1966–1976.

Livingston, B.T., Killian, C.E., Wilt, F. et al. (2006). A genome-wide analysis of biomineralization-related proteins in the sea urchin *Strongylocentrotus purpuratus*. *Dev. Biol.* 300: 335–348.

Lowenstam, H. (1980). Bioinorganic constituents of hard parts. In: *Biogeochemistry of Amino Acids* (eds. P.E. Hare, T.C. Hoering and K. King Jr.), 3–16. New York: Wiley.

Lowenstam, H.A. (1981). Minerals formed by organisms. *Science* 211: 1126–1131.

Lowenstam, H.A. and Weiner, S. (1989). *On Biomineralization*. Oxford University Press.

Luquet, G., Fernández, M.S., Badou, A. et al. (2012). Comparative ultrastructure and carbohydrate composition of gastroliths from Astacidae, Cambaridae and Parastacidae freshwater crayfish (Crustacea, Decapoda). *Biomolecules* 3: 18–38.

Mann, K. (2004). Identification of the major proteins of the organic matrix of emu (*Dromaius novaehollandiae*) and rhea (*Rhea americana*) eggshell calcified layer. *Br. Poult. Sci.* 45: 483–490.

Mann, K. (2015). The calcified eggshell matrix proteome of a songbird, the zebra finch (*Taeniopygia guttata*). *Proteome Sci.* 13: 29.

Mann, K. and Siedler, F. (1999). The amino acid sequence of ovocleidin 17, a major protein of the avian eggshell calcified layer. *Biochem. Mol. Biol. Int.* 47: 997–1007.

Mann, K. and Siedler, F. (2004). Ostrich (*Struthio camelus*) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications. *Biochim. Biophys. Acta* 1696: 41–50.

Mann, K., Weiss, I.M., André, S. et al. (2000). The amino-acid sequence of the abalone (*Haliotis laevigata*) nacre protein perlucin. *FEBS J.* 267: 5257–5264.

Mann, S. (1983). Mineralization in biological systems. In: *Inorganic Elements in Biochemistry*, 125–174. Springer.

Marie, B., Luquet, G., Pais De Barros, J.-P. et al. (2007). The shell matrix of the freshwater mussel *Unio pictorum* (Paleoheterodonta, Unionoida). Involvement of acidic polysaccharides from glycoproteins in nacre mineralization. *FEBS J.* 274: 2933–2945.

Marie, B., Luquet, G., Bédouet, L. et al. (2008). Nacre calcification in the freshwater mussel *Unio pictorum*: Carbonic anhydrase activity and purification of a 95 kDa calcium-binding glycoprotein. *Chembiochem* 9: 2515–2523.

Marie, B., Joubert, C., Tayalé, A. et al. (2012). Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proc. Natl. Acad. Sci. U.S.A.* 109: 20986–20991.

Marin, F., Luquet, G., Marie, B., and Medakovic, D. (2007). Molluscan shell proteins: Primary structure, origin, and evolution. *Curr. Top. Dev. Biol.* 80: 209–276.

Marin, F., Le Roy, N., and Marie, B. (2012). The formation and mineralization of mollusk shell. *Front. Biosci.* 4: 1099–1125.

Miller, M.F., 2nd and Wyckoff, R.W. (1968). Proteins in dinosaur bones. *Proc. Natl. Acad. Sci. U.S.A.*, 60: 176–178.

Miyamoto, H., Miyashita, T., Okushima, M. et al. (1996). A carbonic anhydrase from the nacreous layer in oyster pearls. *Proc. Natl. Acad. Sci. U.S.A.* 93: 9657–9660.

Miyamoto, H., Miyoshi, F., and Kohno, J. (2005). The carbonic anhydrase domain protein nacrein is expressed in the epithelial cells of the mantle and acts as a negative regulator in calcification in the mollusc *Pinctada fucata*. *Zoolog. Sci.* 22: 311–315.

Murdock, D.J.E. and Donoghue, P.C.J. (2011). Evolutionary origins of animal skeletal biomineralization. *Cells Tissues Organs* 194: 98–102.

Nakahara, H. and Bevelander, G. (1970). An electron microscope study of the muscle attachment in the mollusc *Pinctada radiata*. *Tex. Rep. Biol. Med.* 28: 279–286.

Ostrom, P.H., Schall, M., Gandhi, H. et al. (2000). New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. *Geochim. Cosmochim. Acta* 64: 1043–1050.

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

Penkman, K.E.H., Preece, R.C., Bridgland, D.R. et al. (2011). A chronological framework for the British Quaternary based on *Bithynia opercula*. *Nature* 476: 446–449.

Pierini, F., Demarchi, B., Turner, J., and Penkman, K. (2016). Pecten as a new substrate for IcPD dating: The Quaternary raised beaches in the Gulf of Corinth, Greece. *Quat. Geochronol.* 31: 40–52.

Ramos-Silva, P., Marin, F., Kaandorp, J., and Marie, B. (2013). Biomineralization toolkit: The importance of sample cleaning prior to the characterization of biomineral proteomes. *Proc. Natl. Acad. Sci. U.S.A.* 110: E2144–E2146.

Rao, A., Seto, J., Berg, J.K. et al. (2013). Roles of larval sea urchin spicule SM50 domains in organic matrix self-assembly and calcium carbonate mineralization. *J. Struct. Biol.* 183: 205–215.

Rao, A., Berg, J.K., Kellermeier, M., and Gebauer, D. (2014). Sweet on biomineralization: Effects of carbohydrates on the early stages of calcium carbonate crystallization. *Eur. J. Mineral.* 26: 537–552.

Reyes-Grajeda, J.P., Moreno, A., and Romero, A. (2004). Crystal structure of ovocleidin-17, a major protein of the calcified *Gallus gallus* eggshell: Implications in the calcite mineral growth pattern. *J. Biol. Chem.* 279: 40876–40881.

Robach, J.S., Stock, S.R., and Veis, A. (2005). Transmission electron microscopy characterization of macromolecular domain cavities and microstructure of single-crystal calcite tooth plates of the sea urchin *Lytechinus variegatus*. *J. Struct. Biol.* 151: 18–29.

Saitta, E.T., Liang, R., Lau, C.Y. et al. (2019). Cretaceous dinosaur bone contains recent organic material and provides an environment conducive to microbial communities. *eLife* 8: e46205.

Salamon, M., Tuross, N., Arensburg, B., and Weiner, S. (2005). Relatively well preserved DNA is present in the crystal aggregates of fossil bones. *Proc. Natl. Acad. Sci. U.S.A.* 102: 13783–13788.

Sarashina, I. and Endo, K. (1998). Primary structure of a soluble matrix protein of scallop shell: Implications for calcium carbonate biomineralization. *Am. Mineral.* 83: 1510–1515.

Schaffer, J. (1889). Über den feineren Bau fossiler Knochen. *Sitzungsberichte Kaiserlich Akad Wissenschaft Wien Math Naturwissenschaft Klasse* 3–7: 319–382.

Schroeter, E.R., DeHart, C.J., Cleland, T.P. et al. (2017). Expansion for the *Brachylophosaurus canadensis* collagen i sequence and additional evidence of the preservation of Cretaceous protein. *J. Proteome Res.* 16: 920–932.

Schweitzer, M.H., Marshall, M., Carron, K. et al. (1997). Heme compounds in dinosaur trabecular bone. *Proc. Natl. Acad. Sci. U.S.A.* 94: 6291–6296.

Sigel, A., Sigel, H., and Sigel, R.K.O. (2008). *Biomineralization: From Nature to Application*. John Wiley & Sons.

Smith, G.G. and Evans, R.C. (1980). The effect of structure and conditions on the rate of racemization of free and bound amino acids. In: *Biogeochemistry of Amino Acids* (eds. P.E. Hare, T.C. Hoering and K. King Jr.), 257–282. New York: Wiley.

Stern, B. (1996). Biomineral lipids in living and fossil molluscs. PhD thesis, Newcastle University.

Stewart, J., Allen, R.B., Jones, A.K.G. et al. (2014). Walking on eggshells: A study of egg use in Anglo-Scandinavian York based on eggshell identification using ZooMS. *Int. J. Osteoarchaeol.* 24: 247–255.

Sykes, G.A., Collins, M.J., and Walton, D.I. (1995). The significance of a geochemically isolated intracrystalline organic fraction within biominerals. *Org. Geochem.* 23: 1059–1065.

Towe, K.M. and Thompson, G.R. (1972). The structure of some bivalve shell carbonates prepared by ion-beam thinning. *Calcif. Tissue Res.* 10: 38–48.

Vielzeuf, D., Garrabou, J., Baronnet, A. et al. (2008). Nano to macroscale biomineral architecture of red coral (*Corallium rubrum*). *Am. Mineral.* 93: 1799–1815.

Walton, D. (1998). Degradation of intracrystalline proteins and amino acids in fossil brachiopods. *Org. Geochem.* 28: 389–410.

Wang, L. and Nilsen-Hamilton, M. (2013). Biomineralization proteins: From vertebrates to bacteria. *Front. Biol.* 8: 234–246.

Weber, E. and Pokroy, B. (2015). Intracrystalline inclusions within single crystalline hosts: From biomineralization to bio-inspired crystal growth. *CrystEngComm* 17: 5873–5883.

Wehmiller, J.F. (1980). Intergeneric differences in apparent racemization kinetics in mollusks and foraminifera: Implications for models of diagenetic racemization. In: *Biogeochemistry of Amino Acids* (eds. P.E. Hare, T.C. Hoering and K. King Jr.), 341–345. New York: Wiley.

Wehmiller, J.F. and Miller, G.H. (2000). Aminostratigraphic dating methods in quaternary geology. *Quat. Geochronol.*: 187–222.

Weiner, S. and Dove, P.M. (2003). An overview of biomineralization processes and the problem of the vital effect. *Rev. Mineral. Geochem.* 54: 1–29.

Weiss, I.M., Kaufmann, S., Mann, K., and Fritz, M. (2000). Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc *Haliotis laevigata*. *Biochem. Biophys. Res. Commun.* 267: 17–21.

Welker, F. (2018). Palaeoproteomics for human evolution studies. *Quat. Sci. Rev.* 190: 137–147.

Wood, R., Ivantsov, A.Y., and Zhuravlev, A.Y. (2017). First macrobiota biomineralization was environmentally triggered. *Proc. Biol. Sci.*, 284: 20170059.

Younis, S., Kauffmann, Y., Bloch, L., and Zolotoyabko, E. (2012). Inhomogeneity of nacre lamellae on the nanometer length scale. *Cryst. Growth Des.* 12: 4574–4579.

Zhang, Q., Liu, L., Zhu, F. et al. (2014). Integrating de novo transcriptome assembly and cloning to obtain chicken Ovocleidin-17 full-length cDNA. *PLoS One* 9: e93452.