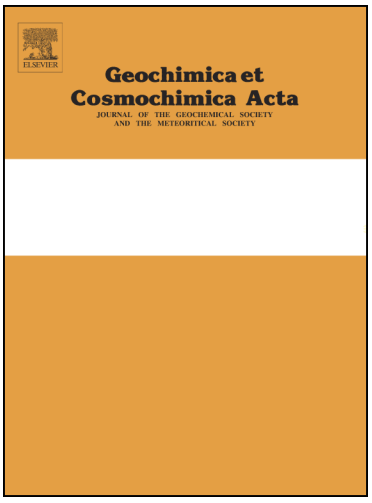
Article:
Saitta, Evan T., Vinther, Jakob, Crisp, Molly K. et al. (24 more authors) (2024) Non-avian dinosaur eggshell calcite can contain ancient, endogenous amino acids. Geochimica et Cosmochimica Acta. pp. 1-20. ISSN 0016-7037

https://doi.org/10.1016/j.gca.2023.11.016

Journal Pre-proofs



NON-AVIAN DINOSAUR EGGSHELL CALCITE CAN CONTAIN ANCIENT, ENDOGENOUS AMINO ACIDS

Evan T. Saitta, Jakob Vinther, Molly K. Crisp, Geoffrey D. Abbott, Lucy Wheeler, Samantha Presslee, Thomas G. Kaye, Ian Bull, Ian Fletcher, Xinqi Chen, Daniel Vidal, Fernando Sanguino, Ángela D. Buscalioni, Jorge Calvo, Paul C. Sereno, Stephanie L. Baumgart, Michael Pittman, Matthew J. Collins, Jorune Sakalauskaite, Meaghan Mackie, Federica Dal Bello, Marc R. Dickinson, Mark A. Stevenson, Paul Donohoe, Philipp R. Heck, Beatrice Demarchi, Kirsty E. H. Penkman

PII: S0016-7037(23)00578-1
DOI: https://doi.org/10.1016/j.gca.2023.11.016
Reference: GCA 13230

To appear in: *Geochimica et Cosmochimica Acta*

Received Date: 14 February 2023
Revised Date: 17 November 2023
Accepted Date: 20 November 2023

Please cite this article as: Saitta, E.T., Vinther, J., Crisp, M.K., Abbott, G.D., Wheeler, L., Presslee, S., Kaye, T.G., Bull, I., Fletcher, I., Chen, X., Vidal, D., Fernando Sanguino, Buscalioni, A.D., Calvo, J., Sereno, P.C., Baumgart, S.L., Pittman, M., Collins, M.J., Sakalauskaite, J., Mackie, M., Dal Bello, F., Dickinson, M.R., Stevenson, M.A., Donohoe, P., Heck, P.R., Demarchi, B., E. H. Penkman, K., NON-AVIAN DINOSAUR EGGSHELL CALCITE CAN CONTAIN ANCIENT, ENDOGENOUS AMINO ACIDS, *Geochimica et Cosmochimica Acta* (2023), doi: https://doi.org/10.1016/j.gca.2023.11.016

NON-AVIAN DINOSAUR EGGSHELL CALCITE CAN CONTAIN ANCIENT, ENDOGENOUS AMINO ACIDS

Evan T. Saitta[1,2*], Jakob Vinther[3,4], Molly K. Crisp[5], Geoffrey D. Abbott[6], Lucy Wheeler[5], Samantha Presslee[5], Thomas G. Kaye[7], Ian Bull[8], Ian Fletcher[9], Xinqi Chen[10], Daniel Vidal[1,11], Fernando Sanguino[11], Ángela D. Buscalioni[12], Jorge Calvo[13†], Paul C. Sereno[1], Stephanie L. Baumgart[1], Michael Pittman[14], Matthew J. Collins[15,16], Jorune Sakalauskaite[17], Meaghan Mackie[15,18], Federica Dal Bello[19], Marc R. Dickinson[5], Mark A. Stevenson[20], Paul Donohoe[6], Philipp R. Heck[2,21,22], Beatrice Demarchi[17], & Kirsty E. H. Penkman[5]

[1]Department of Organismal Biology & Anatomy, University of Chicago, Chicago, IL, USA
[2]Integrative Research Center, Field Museum of Natural History, Chicago, Illinois, USA
[3]School of Earth Sciences, University of Bristol, Bristol, UK
[4]School of Biological Sciences, University of Bristol, Bristol, UK
[5]Department of Chemistry, University of York, York, UK
[6]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK
[7]Foundation for Scientific Advancement, Sierra Vista, Arizona, USA
[8]School of Chemistry, University of Bristol, Bristol, UK
[9]Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey, UK
[10]Department of Mechanical Engineering and NU*ANCE* Center, Northwestern University, Evanston, Illinois, USA
[11]Grupo de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain
[12]Departamento de Biología, Universidad Autónoma de Madrid, Madrid, Spain
[13]Departamento de Geología y Petróleo, Grupo de Transferencia Proyecto Dino, Facultad de Ingeniería, Universidad Nacional del Comahue, Neuquén, Argentina
[14]School of Life Sciences, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China
[15]The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[16]McDonald Institute for Archaeological Research, University of Cambridge, UK
[17]Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy
[18]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark
[19]Department of Molecular Biotechnology and Health Sciences, University of Turin, Turin, Italy
[20]Department of Geography, Durham University, Durham, UK
[21]Robert A. Pritzker Center for Meteorics and Polar Studies, Field Museum of Natural History, Chicago, Illinois, USA
[22]Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois, USA
†Deceased
*Corresponding author

Abstract: Proteins are the most stable of the macromolecules that carry genetic information over long periods of time. Closed systems are more likely to retain endogenous proteins or their degradation products. Amino acid racemisation data in experimental and subfossil material suggests that mollusc shell and avian eggshell calcite crystals can demonstrate closed system behaviour, retaining endogenous amino acids. Here, Late Cretaceous (Campanian–

Maastrichtian) Argentine titanosaurian sauropod eggshells show dark, organic stains under light microscopy/photography and fluorescence imaging. Raman spectroscopy can yield bands consistent with various organic molecules, possibly including N-bearing molecules or geopolymers. Pyrolysis-gas chromatography-mass spectrometry reveals pyrolysates consistent with amino acids as well as aliphatic hydrocarbon homologues that are not present in modern eggshell, consistent with kerogen formation deriving from eggshell lipids. High-performance liquid chromatography reveals that their intra-crystalline fraction can be enriched in some of the most stable amino acids (Glx, Gly, Ala, and possibly Val) and are fully racemic, despite being some of the slowest racemising amino acids, indicating ancient origin. This preservation varies across localities, but similar ancient amino acid profiles were also observed in Late Cretaceous Spanish titanosaurians from several localities and Chinese putative hadrosaurid eggshell. These amino acid results are consistent with previous studies on degradation trends deduced from modern, thermally matured, sub-fossil, and ~3.8–6.5 Ma avian eggshell, as well as ~30 Ma calcitic mollusc opercula. Selective preservation of certain fully racemic amino acids, which do not racemise in-chain, and the concentration of free amino acids suggests likely complete hydrolysis of original peptides. Liquid chromatography-tandem mass spectrometry supports this hypothesis by failing to detect any non-contamination peptide sequences from the Mesozoic eggshell. These closed-system amino acids are possibly the most thoroughly supported non-avian dinosaur endogenous protein-derived constituents, at least those that have not undergone oxidative condensation with other classes of biomolecules. Biocrystal matrices can help preserve mobile organic molecules by trapping them (perhaps with the assistance of resistant organic polymers), but trapped organics are nevertheless prone to diagenetic degradation, even if such reactions might be slowed in exceptional circumstances. Future work should survey fossil biocalcite to determine variability in amino acid preservation.

Keywords: Fossils, Eggshell, Amino Acids, Proteins, Taphonomy

1. Introduction: Some biomolecules are highly stable and can survive deep into the geologic record with minimal alteration (Eglinton & Logan 1991; Briggs & Summons 2014), including steroids (Melendez *et al*. 2013) and pigments, such as porphyrins (Greenwalt *et al.* 2013) and melanin (Glass *et al.* 2012). In contrast, biomacromolecules that form from the organised condensation of monomers into polymers based upon the genetic code (e.g., nucleic acids and proteins) can irreversibly hydrolyse to their constituent monomers. However, these relatively unstable biomacromolecules are of the highest biological interest since they serve critical, complex functions in organisms and changes in their sequence and structure can provide insight into evolution, physiology, and ecology (e.g., Leonard *et al*. 2002).

Ancient DNA has been recovered from mammoth teeth in permafrost sediments as old as 1.1–1.2 Ma (van der Valk *et al*. 2021), nearing the expected upper limit of DNA survival in nature based on predicted half-life calculated from observed decay kinetics (Allentoft *et al*. 2012), and a recent report suggests the preservation of environmental DNA in permafrost up to possibly 2 Ma (Kjær *et al.* 2022). Early claims of preserved older DNA, including Mesozoic

consideration) would place their thermal age orders of magnitude older than the reports from arctic sites (Hedges 2002; McNamara *et al*. 2009; Demarchi *et al.* 2016).

One difficulty in searching for ancient proteins comes from environmental and laboratory contamination (Buckley *et al.* 2008, 2017; Bern *et al.* 2009). For example, amber might trap some ancient amino acids, but their composition and racemization patterns suggest that at least some are exogenous (Collins *et al*. 2009; McCoy *et al.* 2019; Barthel *et al*. 2020). The triboelectric (i.e., static electric) effect of amber (Freeman & March 1999) can attract exogenous proteins, especially with filamentous keratin interactions, such as feathers (McCoy *et al.* 2019). Examining intra-crystalline proteins deposited within biominerals mitigates contamination concerns. Unlike open-system bone (Bada *et al.* 1999; Reznikov *et al.* 2018; Saitta *et al.* 2019), typically denser calcium carbonate biominerals (e.g., mollusc shells [Penkman *et al.* 2008, 2013; Gries *et al.* 2009] and avian eggshells [Brooks *et al*. 1990; Crisp *et al.* 2013]) can act as a closed system for amino acids within the intra-crystalline voids of the calcite (Towe & Thompson 1972; Towe 1980; Collins & Riley 2000). Eggshell respiratory pores, which are orders of magnitude larger than the intra-crystalline voids which are proposed to entrap the protein (Gries *et al.* 2009), do not influence this property since it is the calcite crystals of the eggshell that trap these amino acids within them (Towe & Thompson 1972; Towe 1980; Brooks *et al*. 1990; Collins & Riley 2000; Crisp *et al.* 2013). The eggshell pores are simply larger regions in which these calcite crystal subunits are absent. To clarify, we are not arguing that the egg as a whole acts as a closed system, in which the endogenous amino acids are to be found within the region of embryonic development (since clearly the eggshell pores open this region to the external environment); they are instead trapped within the calcite crystals of the eggshell itself. Calcite is thermodynamically more stable than aragonite, the latter often recrystallizing as calcite during fossilisation (Benton 2001), making calcite the more promising biomatrix (Wehmiller *et al.* 1976; Harmon *et al.* 1983; Hearty & Aharon 1988; Hoang & Hearty 1989; Penkman *et al.* 2007, 2010).

Early research reported extremely ancient, thermally stable amino acids Glu, Ala, and Val from a ~360 Ma trilobite (Abelson 1954), which had *in vivo* calcite in the cuticle (Dalingwater 1973) and eye lenses (Towe 1973; although see a counter by Lindgren *et al.* 2019 arguing for secondary mineralization). However, the study reported a similar amino acid profile in open-system Jurassic *Stegosaurus* bone apatite (Abelson 1954), suggesting that some of the detected amino acids were possibly exogenous (Saitta *et al.* 2019; Liang *et al.* 2020). Since trilobites are long-extinct, examination of protein diagenesis and calcite system behaviour can be better characterized in extant materials such as eggshell and mollusc opercula, which have recent fossil records and modern tissues for use in comparative thermal maturation experiments. Well-supported closed system amino acids (i.e., not necessarily within a peptide chain) have been reported from ~30 Ma mollusc calcitic opercula (Penkman *et al.* 2013), while claims of intact peptide bonds within interprismatic proteins in 66 Ma Late Cretaceous mollusc shell with data obtained from photoemission electron spectromicroscopy have also been made (Myers *et al.* 2018).

Although calcite can act as a closed system for peptides and amino acids, degradation of trapped organics still proceeds. For example, in a survey of calcitic brachiopod shell, immunochemical signatures of modern shell peptides disappeared by ~2 Ma (Curry *et al.* 1991; Walton 1998; Collins *et al.* 2003). Peptide fragmentation, amino acid profiles, and racemisation patterns have been thoroughly studied in modern, sub-fossil, and ~3.8–6.5 Ma avian eggshell and compared to experimentally matured avian eggshell (Crisp 2013; Crisp *et al.* 2013; Demarchi *et al.* 2016, 2022). As eggshell peptides degrade over time and under higher environmental/experimental temperatures, D/L values along with relative concentrations of Glx, Gly, and Ala increase, while concentrations of Asx and Ser decrease. Among a consistent pattern of peptide degradation observed through a suite of eggshell samples, the oldest

independently authenticated peptide fragments are of an otherwise unstable, short, acidic region of the struthiocalcin protein preserved in ~3.8 Ma low-latitude ratite eggshell (Demarchi *et al.* 2016) and 6.5–9 Ma ratite eggshell from northwestern China (Demarchi *et al.* 2022). Even under warm burial histories, the high binding energy of this region of the peptide to calcite results in a unique 'molecular refrigeration' mechanism that drops the effective temperature around the peptide by ~30 K, reducing rates of hydrolysis (thermal age of low-latitude ~3.8 Ma peptide fragment equivalent to ~16 Ma at 10 °C) (Demarchi *et al.* 2016).

Non-avian dinosaur eggshell also consisted of calcite, with a somewhat similar structural organisation to avian eggshell, and can be found in large quantities at certain nesting sites, such as Late Cretaceous Auca Maheuvo in Argentina (Grellet-Tinner *et al.* 2006). Furthermore, they can contain endogenous biomolecules, such as stable porphyrin pigments (Wiemann *et al.* 2017). Even higher degrees of biomolecular preservation have been proposed in Auca Mahuevo eggshells, where immunochemistry was used as evidence for intact protein or protein-derived organics along the eggshell cross-section, including inter-crystalline regions considered to be outside of the closed system calcite crystals (Schweitzer *et al.* 2005). However, using immunochemistry to detect ancient, especially Mesozoic, proteins in fossils has been suggested to be susceptible to false positives (Montgelard *et al*. 1997; Buckley *et al.* 2017; Saitta & Vinther 2019). For example, allergies, such as those to nuts, are instances of inaccurate antigen detection by antibodies, and antibodies raised against parasitic blood flukes can cross-react with peanuts (Igetei *et al.* 2017). See Saitta & Vinther (2019) for suggested methodological improvements of such antibody studies of fossils to add further controls.

More recently, Late Cretaceous titanosaurian eggshell has been suggested to contain proteinaceous moieties using pyrolysis two-dimensional gas chromatography time-of-flight mass spectrometry (Py-GC×GC-TOFMS), based on the presence of nitrogen-bearing pyrolysates, including diketodipyrrole (Dhiman *et al*. 2021).

Therefore, using a variety of analytical techniques that can detect different components of organic molecular signals, this study aims to test the potential for preservation of original amino acids (and ultimately peptide sequence information) from Mesozoic calcite eggshell.

2. Materials and Methods:

To explore the potential for preservation of peptide sequences from dinosaur eggshell, we took a staged approach to the sample selection – initially analysing material that due to their collection histories were most amenable to destructive analysis and then progressing as successful results were obtained (Tables 1–2).

Initially we analysed two independently obtained South American titanosaurian eggshells that were separately commercially imported into the USA and Denmark in roughly the late 1990s to early 2000s and then donated for research in the late 2010s (Table 1). Through our repatriation to Argentina with the assistance of Asociación Paleontológica Argentina and the National Authority of the Application of the Law of Paleontological Heritage, these two samples now belong to the collection of the Museo Provincial Patagónico de Ciencias Naturales (MPCN) de la Ciudad de General Roca, Río Negro (see supplemental material). These two eggshells are best assigned to the Late Cretaceous (Middle Campanian–Early Maastrichtian, ~73–69 Ma) titanosaurian ootaxon *Megaloolithus megadermus* (also referred to as Tipo 1e) from the Allen Formation based on their diagnostic features (see supplemental material), such

histories, these samples were deemed amenable for highly destructive analyses using many methods. *M. megadermus* A and B are consistent with Argentine titanosaurian eggshells more generally in morphology and preservation, both in exterior ornamentation and internal calcite layering (see supplemental material).

We also studied two Late Cretaceous (Early–Middle Campanian, ~83–74.5 Ma) Argentine titanosaurian eggshells (Table 1) from the Auca Mahuevo Lagerstätte in the Anacleto Formation of the Río Colorado Subgroup in Neuquén Province, Argentina (Chiappe

reminiscent of petrol and burnt hair (i.e., an observation consistent with ancient organic preservation).

3.1. Light microscopy, LSF imaging, & photography; evidence of organic staining

The *M. megadermus* A fragment has a lightly coloured interior and exterior surface, and the exterior surface is covered in small, round ornamentation with what appears to be small amounts of lightly colored sediment in between the ornaments (Fig. 1A, C). The interior cross-section of the eggshell shows large regions of black calcite (i.e., consistent with organic impurities in the calcite) whose structure has been lost (Fig. 1B); however, there is a band of lightly coloured calcite deep in the interior of the eggshell cross-section (Fig. 1E). The black, astructural calcite does not fluoresce. The lightly coloured calcite fluoresces pale white/yellow. The infilling material within the pore spaces, possibly from the sediment matrix (see discussion of thin sections below), between ornaments and calcite crystal units fluoresces light blueish (Fig. 1 D, F). About half of the calcite in the eggshell appears to be black and astructural,

dark/recrystallized regions between these light regions. The light regions showed a much higher fluorescence background than the dark regions during Raman spectroscopy; this resulted in more noise and therefore the need to lower the excitation laser power relative to the analyses of the dark regions, making quantitative comparisons of spectral data between the two phases extremely difficult.

Both phases showed some peaks consistent with reference vibrations from calcite and quartz (likely from infilling sediment), but these are still relatively weak compared to the noise – a concerning spectral pattern to obtain from a calcite eggshell in light of our TOF-SIMS attempts that detected Ca ions (supplemental material). Peaks roughly consistent with potential non-cyclic, cyclic, and aromatic hydrocarbons and O-, N-, S-, or halogen-containing organic compounds (Fig. 2C, supplemental material) are of far lower confidence. The epoxy has a distinct spectrum from those of the *M. megadermus* A (supplemental material), although some peaks may be shared (Fig. 2C). Some of the pattern in the *M. megadermus* A spectra is likely due to artefactual quasi-periodic ripples resulting from intense sample luminescence interacting with the edge filter on the Raman spectroscopy equipment we used (Alleon *et al*. 2021; Wiemann & Briggs *et al.* 2022), especially in the light regions of the eggshell. To help account for sample luminescence, future work could run pure calcite and organic standards for comparison or use wavelet transform analysis (Alleon *et al*. 2021) or principal component analysis (Wiemann & Heck 2023). In the meantime, and considering the possible presence of artefactual quasi-periodic ripples in these spectra, we simply note here that the difference in luminescence between the light and dark regions of *M. megadermus* A indicate two different chemical compositions. Enigmatic bands in fossils, especially in the 1200–1800 cm^{-1} range, have also been hypothesized to reflect inorganic (e.g., carbonate), rather than organic, composition (Jurašeková *et al.* 2022).

Modern ostrich eggshell showed calcite and putative organic peaks (with less noise than the *M. megadermus* A), including potential non-cyclic, cyclic, and aromatic compounds, as well as hydrocarbons, O-, N-, S-, or even halogen-bearing organic molecules (supplemental material). The Raman spectrum of the outer (prismatic external) layer of the ostrich eggshell was noisier than those of the center column/palisade and inner mammillary cone layers and may have been more heavily influenced by the embedding epoxy resin. The distinctiveness of the ostrich spectra compared to the *M. megadermus* A spectra is further evidence that the epoxy embedding resin is not dominating the Raman data. However, the possibility that the ostrich eggshell calcite spectra have instrumental edge-filter artefacts due to its high fluorescence background or an inorganic composition that can influence Raman peaks of interest (Jurašeková *et al.* 2022) should also be considered (especially for the outer layer), even if the spectra are less noisy than those of the *M. megadermus* A.

3.3. Py-GC-MS; evidence of ancient organic material

Examining the total ion chromatograms from Py-GC-MS of modern chicken and *M. megadermus* A reveals how different decontamination methods can greatly affect results (supplemental material). This is particularly apparent in *M. megadermus* A, where more intensive decontamination decreased the organic content, evidenced by the more prominent column bleed at the end of the run and reduction of the intensity of some of the relatively later eluting peaks. Overall, it appears that organic content in *M. megadermus* A is lower than that in the modern chicken eggshell samples, evidenced by the prominence of the column bleed observed in *M. megadermus* A that was not observed in the modern chicken eggshell samples. However, minor variation in the mass of eggshell powder analysed could also influence this pattern, at least in part.

Comparing pyrolysates from the samples that had been dichloromethane (DCM) rinsed

confidence (i.e., relative to the background signal) than *M. megadermus* A and B discussed above.

Another observation of note is that the lightly coloured outer flakes of *M. megadermus* A that separated during powdering and were analysed separately are intermediate between the whole *M. megadermus* samples and the LACM samples (Fig. 4C), suggesting relatively depleted amino acid signal in this region of the eggshell.

Late Cretaceous Spanish titanosaurian (cf. *Megaloolithus*) eggshell shows variable THAA compositional profiles according to locality (Fig. 4D). Samples from two localities UAM3a (Bastús, Lleida, Catalonia) and UAM4a–b (Biscarri, Lleida, Catalonia) do show high levels of stable Gly and Ala, but do not fully match with expected THAA compositional profiles from ancient or thermally mature avian eggshell (e.g., relatively low Glx, absent Val, small amounts of Asx and Ser present, high Tyr present). In contrast, samples from the other three localities UAM1a–c (La Rosaca, Burgos), UAM2a (Requena, Valencia), and UAM5a (Portilla, Cuenca) show THAA compositional profiles that match closely with those expected from ancient and thermally matured avian eggshell as well as those observed from *M. megadermus* A and B studied here, namely a preponderance of Glx, high levels of Gly and Ala, consistent Val detection, and absent Asx and Ser. These three localities provide strong evidence for ancient, endogenous amino acids.

Likewise, the Spanish titanosaurian localities with THAA compositional profiles consistent with diagenetically altered avian eggshell (UAM1a–c, UAM2a, and UAM5a) have D/L ratios consistent with fully racemized amino acids, as well as nearly complete degradation of Ser into Ala (Table 3), similar to *M. megadermus* A and B. Similar D/L values are seen between FAA and THAA. In contrast, the localities with THAA compositional profiles less consistent with diagenetically altered avian eggshell (UAM3a, UAM4a–b) show inconsistent D/L and Ser/Ala values reflective of their low amino acid concentrations.

Late Cretaceous Chinese putative hadrosauridae eggshell likewise showed THAA compositional profiles (Fig. 4E) strongly suggestive of a subset of four ancient, endogenous amino acids (a preponderance of Glx, high levels of Gly and Ala, consistent Val detection) with absent Asx and Ser. Furthermore, the two replicates from each the two analysed fragments (UC1a–b) all yielded very similar THAA profiles, indicating replicability of the results.

The putative hadrosauridae eggshell show D/L values indicative of full racemization, as well as nearly complete degradation of Ser into Ala (Table 3), consistent with ancient amino acids like those in *M. megadermus* A and B. Similar D/L values are seen between FAA and THAA.

When comparing total THAA concentrations of the sum of 13 amino acids in picomoles/mg of (non-ethanol rinsed) bleached, 24-hr hydrolysed fossil eggshell (Fig. 4F), it is important to keep in mind that quantification at low values, with relatively few samples/replicates and an elevated baseline obscuring later eluting amino acids, makes such measurements imprecise. Fossil eggshell does have low estimated total THAA concentrations compared to modern, untreated avian eggshell, which are expected to be around ~5,000–13,000 picomoles/mg (Crisp et al. 2013). However, we can see that fossil eggshell whose THAA compositional profiles (Fig. 4B–E) more closely match with those expected from ancient and thermally mature avian eggshell (Fig. 4A) (i.e., *M. megaloolithus* A and B, three localities of Spanish titanosaurian [UAM1a–c, UAM2a, UAM5a], and Chinese putative hadrosauridae) tend to have higher total estimated THAA concentrations (Fig. 4F) than do fossil eggshell whose THAA compositional profiles do not as closely match with that expectation (i.e., Auca Mahuevo LACM 7324 A and B, two localities of Spanish titanosaurian [UAM3a, UAM4a–b]). Combined with the fact that the fossil eggshells which give robust results have total estimated THAA concentrations higher than expected from laboratory blanks

diketopiperazine (i.e., diketodipyrrole) they detected as a pyrolysis product could be consistent with low levels of amino acid preservation like those described here.

A study using Py-GC-MS on a thick fluid produced from modern feathers thermally matured at 250°C, 250 bars, and 24 h also yielded 2,5-diketopiperazine (Saitta *et al*. 2017), hinting that these pyrolysis products might also derive from free amino acid mixtures after hydrolysis of polypeptides, rather than from preserved proteins themselves. It might also be worth noting that the Dhiman *et al.* (2021) samples did not undergo bleach treatment as in our HPLC amino acid analysis, but instead had their outer surfaces cleaned with 5% HCl, then ultra-pure water, and finally ultrasonication in dichloromethane – so it should be considered as to whether this method is as efficient at removing inter-crystalline amino acids. Ultimately, it is better to triangulate results using multiple methods (e.g., pyrolysis and HPLC) than to draw conclusions from a single marker using one type of method (i.e., pyrolysis), and as such, we think our current results provide further insight into those of Dhiman *et al.* (2021).

How, then, might one best explore the evidence for putative ancient, proteinaceous moieties? Studies concluding protein preservation in fossils must consider several aspects of this claim (Hendy *et al*. 2018). Fossil proteins or protein-derived organics are those that have an appropriate *chemical signature*, *endogenicity* (McLoughlin 2011), and *antiquity*.

1. The composition of the organics must A) be consistent with protein or their degradation products generally (*chemical signature*) and B) should specifically be consistent with the composition expected from the *in vivo* proteins of the tissue or their degradation products (*chemical signature, endogenicity*).
2. The organics should be analysed for their degree of preservation (*antiquity*). Typically, older fossils would be expected to have greater degradation and alteration. Mechanisms explaining the observed degree of preservation must be supported (e.g., thermal ages or 'molecular cooling' of ~3.8–6.5 Ma eggshell peptide fragments; Demarchi *et al.* 2016, 2022).
3. The organics must localise in a manner that would be expected from endogenous protein sources as opposed to exogenous sources (*endogenicity*). The tissue matrix (e.g., biominerals of bone apatite or eggshell calcite) that any organics are fossilised in will dictate what patterns of organic influx or outflux are observed. Closed systems, as eggshell calcite can be, make interpreting these patterns far easier.

The three points above are further benefitted by amassing evidence obtained from multiple analytical methods, each with their own strengths and weaknesses, that help to triangulate/validate conclusions via consilience.

The results from the thick-shelled *M. megadermus* A and B (but to a lesser extent the low-concentrated amino acids in the Auca Mahuevo eggshells LACM 7324 A and LACM 7324 B as well as two localities of the Spanish titanosaurian eggshells) appear to meet these criteria. As such, the *M. megadermus* A and B deserve detailed discussion. Note that three localities of the Spanish titanosaurian eggshells as well as the Chinese hadrosaurid eggshell also showed strong evidence of selective, endogenous amino acid preservation with RP-HPLC, but *M. megadermus* A and B were analysed with a greater number of destructive methods given their collection histories.

4.1. Composition of protein-derived material

A) The chemical signature of the dinosaur eggshells match with that of organic, protein-derived material. There is a non-fluorescing (in bulk cross-section under LSF), black/brown

colouration typical of organic material, as well as a release of organic volatiles upon powdering (as evidenced by the strong, peculiar odour); characterisation of similar volatile organic compounds by GC-MS supported the existence of a closed system in ~3.8 Ma ratite eggshell (Demarchi *et al.* 2016). *M. megadermus* A also yields organic pyrolysis products that are at least consistent with the presence of amino acid-derived material, such as toluene, benzenes, and phenols (Fig. 3). Py-GCxGC-TOFMS of Late Cretaceous titanosaurian eggshell from India similarly yielded major pyrolysis products, largely localized to the eggshell rather than the sediment, that included benzenes and phenols (Dhiman *et al*. 2021). Those researchers attributed phenols to amino acid precursors (Stankiewicz *et al*. 1998; Dutta *et al*. 2007; Dhiman *et al*. 2021). The Indian titanosaurian eggshell also contained succinimide, diketodipyrrole (a type of diketopiperazine), and abundant nonadecenenitrile pyrolysis products (Dhiman *et al*. 2021) that are consistent with amino acid precursors (Saitta *et al*. 2017). Here, Raman spectroscopy bands at least consistent with various organic molecules, including N-bearing molecules, are present throughout the *M. megadermus* A cross section, but we also observe edge-filter artefacts (Alleon *et al.* 2021) and currently cannot exclude peak overlaps from inorganic compounds (Jurašeková *et al.* 2022) that do not allow for an unambiguous identification of peaks from biological organic compounds. Nevertheless, RP-HPLC shows that amino acids are present within many of the titanosaurian as well as the putative hadrosaurid eggshells' calcite.

B) As for the more precise nature of this organic signature consistent with protein-derived material, the THAA compositional profiles of most of the dinosaur eggshells (*M. megadermus*, putative hadrosaurid, three localities of Spanish titanosaur) closely match those expected from old, thermally mature avian eggshell (i.e., Glx, Gly, and Ala enriched, but Asx and Ser depleted) (Crisp *et al*. 2013), unsurprising given that birds are dinosaurs and non-avian

4.2. Preservation of protein-derived material

As only four amino acids (Glx, Gly, Ala, and possibly Val) show clear, consistent evidence of survival in all the variously treated *M. megadermus*, putative hadrosaurid, and three localities of Spanish titanosaurian THAA and FAA replicates (supplemental material), consistent with known half-lives and decomposition products (Vallentyne 1964) and degradation patterns of subfossil avian eggshell (Crisp *et al*. 2013), this is strongly suggestive of significant peptide bond hydrolysis and subsequent degradation of less stable amino acids. These amino acids tend to be thermally resistant/stable over deep time in avian eggshell (Crisp 2013; Crisp *et al.* 2013) and simple in structure (e.g., Gly, Ala, Val). They are the only amino acids unequivocally present in the dinosaur eggshells and are in low concentrations relative to modern avian eggshell (supplemental material), indicative of long-term diagenesis. Ala and Val have hydrophobic side chains, and insolubility might further enhance their preservation. Ser does not appear to be present in the dinosaur eggshells, and this amino acid is one of the least thermally stable, with the degradation of Ser contributing to Ala enrichment (Vallentyne 1964) in ancient or thermally mature eggshell. The amino acid compositional profiles from ~30 Ma mollusc shell (Penkman *et al.* 2013) show similarities to those detected in the titanosaurian eggshells, despite presumably different profiles of the original proteins, suggesting that amino acid thermal stability is key to preservation. Given such a decrease in the amino acid types, long phylogenetically informative peptides are not expected. This is analogous to taking a novel and selectively removing all but five letters; paragraphs, sentences, and words would be lost in the process. Furthermore, relatively little comparative literary criticism would be expected merely by comparing novels by their relative frequency of these remaining five letters.

The amino acids in the dinosaur eggshells are all fully racemic (Table 3), suggesting that they are very ancient. Furthermore, the amino acids detected in the dinosaur eggshells are among the slowest racemising and most stable amino acids (Smith & Evans 1980; Crisp *et al.* 2013). Since relative racemisation rates between different amino acids are consistent over a range of temperatures (Crisp *et al.* 2013), any endogenous amino acids are likely fully racemic regardless of the dinosaur eggshells' burial temperatures. Most amino acids can only racemise as free amino acids or N-terminally bound in-chain (Mitterer & Kriausakul 1984), with the exception of Ser (Demarchi *et al.* 2013a) and Asx (Stephenson & Clarke 1989) that can racemise internally bound in-chain; neither Ser or Asx are retained in the dinosaur eggshells. The fully racemic mixtures observed in the dinosaur eggshells suggest that the amino acids derive largely from free amino acids or dipeptides in the form of cyclic dipeptides (e.g., diketopiperazines formed under thermal polymerisation from even racemized amino acid reactants [Hartmann *et al*. 1981]), abiotically condensed dipeptides (i.e., secondarily synthesized from previously free amino acids [Cleaves *et al*. 2009]), or the final remnants of the original peptide chain. However, abiotic dipeptide synthesis would require significant geothermal heat (Cleaves *et al.* 2009) and, even though hydrolysis rates vary with environmental factors such as temperature, previously predicted rates of peptide hydrolysis are typically not supportive of original Mesozoic polypeptide survival by orders of magnitude (Nielsen-Marsh 2002). Gly, Ala, and Val in the replicates of *M. megadermus*, putative hadrosaurid, and three localities of Spanish titanosaurian show some consistency in having somewhat similar THAA and FAA concentrations, which would suggest high levels of peptide bond hydrolysis, supported by the similar D/L values retrieved from FAA and THAA, suggesting that very few peptide-bound L-amino acids persist. This similarity in THAA and FAA D/L values in the dinosaur eggshells is in contrast to younger proteinaceous samples whose FAA D/L values are greater than their THAA D/L values (Hare 1971; Smith & Evans

1980; Liardon & Lederman 1986), as most amino acids cannot readily racemise within a peptide chain (Hare 1971; Smith & Evans 1980; Liardon & Lederman 1986; Crisp *et al.* 2013). At low temperatures, such as would be expected during early taphonomic processes prior to any significant geothermal heating during diagenesis, hydrolysis is favoured over racemisation for many amino acids (Crisp *et al.* 2013; Demarchi *et al.* 2013b; Tomiak *et al.* 2013), meaning that the fully racemic amino acids detected here are likely indications of heavily hydrolysed proteins.

Detected Glx is predicted to be largely comprised of Glu since irreversible deamidation is a rapid degradation reaction, especially in acidic conditions (Hill 1965; Geiger & Clarke 1987; Wilson *et al.* 2012). The recrystallisation observed in *M. megadermus* A could be consistent with past acidic conditions (Plummer *et al.* 1978) but does not appear to have impacted the closed-system nature of the eggshell calcite. Additionally, given their role in eggshell mineralisation, one might also expect many acidic amino acids to be present prior to diagenetic alteration (Marin *et al.* 2007). Therefore, the detected Glx is best interpreted as an indicator of diagenetically altered, ancient Glu, rather than Gln.

The apparently complete hydrolytic cleavage of amino acids in the *M. megadermus* A, compounded by the loss of most of the unstable amino acids, is further supported by the failure of LC-MS/MS to detect any significant, non-contaminant peptides (Table 4). No homologous sequence to the highly stable region of struthiocalcin, as detected in ~3.8–6.5 Ma ratite eggshell (Demarchi *et al.* 2016, 2022) and preliminarily in 6.5–9 Ma ratite eggshell (Demarchi *et al.* 2022), was detected. Of course, one would not necessarily expect a titanosaurian to have a homolog to ratite struthiocalcin, given the vast evolutionary distance between them. However, struthiocalcin and related proteins are involved in eggshell mineralisation (Mann & Siedler 2004; Sánchez-Puig 2012; Ruiz-Arellano & Moreno 2014; Ruiz-Arellano *et al.* 2015) and make up ~20 % of the total organics in modern avian eggshell (Nys *et al.* 1999, 2004; Mann & Siedler 2004; Woodman 2012). If any endogenous peptides were to occur in the titanosaur, a similar negatively charged, Asp-rich sequence that binds tightly to calcite and has high preservation potential (Marin *et al.* 2007; Demarchi *et al.* 2016) might be a prime candidate. Importantly, most of the detected peptides in LC-MS/MS contain the labile amino acid Ser, as well as amide-bearing Asn and Gln residues (Table 4). Since Asn and Gln are expected to undergo fairly rapid deamidation, even in-chain (Hill 1965; Geiger & Clarke 1987; Wilson *et al.* 2012), if such peptides were indeed Mesozoic, one might predict them to be fully converted into Asp and Glu.

Furthermore, while modern avian eggshell yields several prominent nitrogen-bearing pyrolysis products, the same is not true for the *M. megadermus* A (Fig. 3). This likely indicates a far higher proteinaceous concentration in modern eggshell and, conversely, high amounts of degradation of original proteins in the titanosaurian eggshells, confirmed by the lower amino acid concentrations evident in the RP-HPLC data (Fig. 4; supplemental material). Similarly, Py-GCxGC-TOFMS of Late Cretaceous titanosaurian eggshell from India found a low abundance of N-bearing organic pyrolysis products compared to aromatic products, alongside a limited diversity of diketopiperazines (i.e., only detecting a single type, diketodipyrrole), and the authors attributed this to diagenetic degradation (Dhiman *et al*. 2021). In our study, modern ostrich eggshell appeared to yield Raman vibrations with greater signal/noise ratio than the *M. megadermus* A (i.e., cleaner spectra), even under the same excitation laser power (i.e., 20 mW). This greater noise is potentially consistent with relatively lower concentrations of organic molecules in the *M. megadermus* A than in the ostrich eggshell, although luminescence in the fossil sample during Raman spectroscopy can make such quantitative comparisons unreliable (Alleon *et al*. 2021). On a related note, the presence of various Raman bands in the *M. megadermus* A potentially consistent with halogen-bearing organic molecules (supplemental material) possibly indicates bonding of exogenous halogens to endogenous organic

might expect the dinosaur amino acids to be present in these more internal layers. The dark black/brown staining of the titanosaurian eggshells, consistent with the presence of endogenous organics, is often most prominent in the central regions of the eggshell cross-sections.

Calcite's birefringent, anisotropic optical properties (Ghosh 1999) allow for easy determination under cross-polarised light as to what portions of the *M. megadermus* A cross-section have been recrystallised, altering their orientation and leading to a loss of original eggshell morphology in its internal calcite layering. One might hypothesize that such recrystallisation could open the system, leading to a loss of endogenous amino acids. The recrystallised regions of the *M. megadermus* A are those that also have black colouration (Fig. 1) – consistent with the presence of organics. It has been experimentally demonstrated in ostrich eggshell that calcite can maintain closed system behaviour with respect to their intra-crystalline proteins between pH 5 and pH 9, at least, without affecting protein degradation and amino acid racemisation (Crisp *et al.* 2013). Recrystallisation, if induced by pH fluctuations, might have occurred to a degree that resulted in a loss of original eggshell structure but maintained the closed system behaviour of intra-crystalline proteins without completely dissolving the calcite (as seen in some diagenetically altered molluscan opercula [Preece & Penkman 2005]) or inducing acid hydrolysis of any organic geopolymers that possibly contributing to closed system behaviour (see following section).

Exogenous environmental amino acids might have become subsequently trapped in the recrystallised calcite. Based on our cumulative data, the amino acids are very ancient, so such re-entrapment would have to have occurred long ago. Given that recrystallisation could have occurred under significant diagenetic influence, the immediate burial environment might have been low in exogenous amino acids. Hypothetically, if exogenous amino acids were trapped late in diagenesis, the environmental THAA profile might be enriched in diagenetically stable amino acids. However, the THAA compositional profiles of the dinosaur eggshells match those predicted from ancient, thermally mature eggshell (i.e., ratios of Glx to Gly, Ala, and Val). The relatively high Glx concentration compared to moderate Gly and Ala concentrations in the titanosaurian eggshells is better explained by eggshell protein precursors than diagenetic biases. Gly is the simplest amino acid and, we hypothesize, might be expected to occur in the highest concentration if amino acid compositional profiles contained solely a diagenetic signal. For instance, one study found that open-system, Late Cretaceous dinosaur bone supporting an active microbiome can become heavily Gly dominated (Saitta *et al.* 2019) (although note that bone and eggshell amino acid composition differ *in vivo*, with high Gly content in bone). Furthermore, depending on the precise mechanism by which biocalcite crystals act as a closed system, re-entrapment of exogenous amino acids might be unlikely (see following section).

Raman spectroscopy revealed that both light and dark phases of the *M. megadermus* A possibly, but not unambiguously, contained Raman signals that were consistent with various organic molecules, including N-bearing molecules (supplemental material). If genuine, however, this would further mitigate the concern that all of the amino acids are hypothetically deriving from exogenous amino acids trapped in the recrystallized regions of the eggshell. Ultimately, given differences in luminescence between the two phases under Raman spectroscopy and the associated noise in the spectra (Fig. 2), quantitative comparisons of the concentrations of organics between the two phases is ill-advised. As such, the hypothesis that the majority of the amino acids and other organics are associated with the dark, low Raman luminescence regions of the eggshell remains open.

Although data is very limited, the intermediate quality of an ancient amino acid signature in the lightly coloured outer flakes of *M. megadermus* A that separated during powdering (in-between the strong signature of the whole *M. megadermus* A and B eggshells and the weak signature of LACM 7324 A and B [Fig. 1, supplemental material]) might indicate that the dark coloured regions of *M. megadermus* A contain the highest concentrations of

endogenous amino acids. This would be consistent with the general correlation between dark colour and high ancient organic content seen across many fossils and sediments (e.g., conodont colour alteration index [Epstein *et al.* 1976]), but further data are needed.

4.4. Non-protein organics in eggshell through fossilisation

What other endogenous organics might be present in the fossil eggshells, and might they contribute to the mechanism of preservation of the protein-derived material? Modern eggshells contain organics other than proteins. In avian eggshells and other biocalcite, their closed system behaviour may be purely a result of the calcite crystals themselves or a combination of calcite and recalcitrant organics within the biomineral pores (Crisp *et al.* 2013).

Modern avian eggshells contain endogenous phospholipids (Simkiss & Tyler 1958). Kerogen-like aliphatic compounds can form taphonomically via *in situ* polymerisation of labile lipids (e.g., fatty acids from hydrolysed phospholipids) during decay and diagenesis (Stankiewicz *et al.* 2000; Gupta *et al.* 2006a, 2006b, 2007a, 2007b, 2008, 2009). Kerogen signatures were detected in the *M. megadermus* A using Py-GC-MS under full scan mode, and these could have derived from endogenous phospholipids (although the potential for organic polymer consolidants, such as butvar or vinac, to contribute to this signature should be considered). Further analysis of the fossil eggshell kerogen using selected ion monitoring (SIM) scanning mode would allow for a useful comparison of carbon number between modern eggshell phospholipid fatty acid tails and the alkanes/alkenes detected in the fossil in order to estimate the extent of *in situ* polymerisation. For comparison, Py-GCxGC-TOFMS of Late Cretaceous titanosaurian eggshell from India detected a homologous series of n-alkane/alkenes from C_8 to C_{12} as major pyrolysis products (Dhiman *et al*. 2021), and those authors attributed these aliphatic compounds to *in situ* polymerization of eggshell lipids (Stankiewicz *et al*. 2000; Dhiman *et al*. 2021). Raman vibrations possibly from aliphatic organic compounds (e.g., hydrocarbons) were also detected in the *M. megadermus* A (supplemental material), consistent with alkane/alkene geopolymers, but are possibly overlapped by peaks from edge-filter artefacts and inorganic compounds.

Furthermore, protein breakdown products can react with oxidised lipids through Maillard-like reactions to condense into stable, browning compounds referred to as N-heterocyclic polymers (Hidalgo *et al.* 1999; Wiemann *et al.* 2018). Raman bands in the *M. megadermus* A consistent with cyclic or N-bearing organic molecules could support the presence of such nitrogenous polymers, although this assumes that they are not edge-filter spectral artefacts or peaks from inorganic compounds. Therefore, kerogen and/or N-heterocyclic polymers could contribute to the dark, organic colouration observed in the titanosaurian eggshells. The possibility that these lipid-derived or partly lipid-derived organic fossils help to trap endogenous amino acids should be investigated.

Polysaccharides are also present in the palisade/column and mammillary cone layers in modern avian eggshells (Baker & Balch 1962). Additionally, acid-mucopolysaccharide and protein complexes are present in avian eggshells (Simkiss & Tyler 1957). Melanoidins, condensation products formed from protein and polysaccharide degradation via Maillard reactions, can be present in fossils (Collins *et al.* 1992; Stankiewicz *et al.* 1997). Low molecular weight, aromatic structures comprise a significant portion of humic acids, formed through similar Maillard-like reactions (Hatcher *et al.* 1981; Hedges *et al.* 2000; Sutton & Sposito 2005). Therefore, the small, aromatic pyrolysis products detected in the *M. megadermus* A (as well as those detected in Late Cretaceous titanosaurian eggshell from India [Dhiman *et al*. 2021]) may be at least consistent with melanoidins. Melanoidin or humic acid-like organics might also contribute to the black colouration in the titanosaurian eggshells (Schroeter *et al*. 2019). Raman bands in the *M. megadermus* A possibly from aromatic and/or N-bearing organic compounds are consistent with melanoidins, but these are possibly influenced by quasi-

periodic artefacts (Alleon *et al*. 2021) or inorganic compounds (Jurašeková *et al.* 2022). Melanoidins can be bleach resistant, although they can be degraded using acid hydrolysis (Hoering 1980; Namiki 1988; Wang *et al.* 2011). Therefore, the potential presence of melanoidins might help to protect amino acids in the titanosaurian eggshells, shielding the so-called 'intra-crystalline' amino acid fraction from bleach oxidation but subsequently releasing them upon acid hydrolysis in the laboratory.

Kerogen can form early on in taphonomy during decay (Gupta *et al.* 2009) and humic acids can form in surface soils (Sutton & Sposito 2005). However, it is also possible that the dark-staining, non-protein organics in the titanosaurian eggshells formed after long periods of time and through diagenesis during deep burial, possibly consistent with their localisation to the recrystallized calcite in *M. megadermus* A (as evidenced by the dark colouration). Given observed rates of protein hydrolysis in eggshells (Crisp *et al.* 2013), it is reasonable to hypothesise that protein hydrolysis would typically occur before and contribute reactants to N-heterocyclic condensation products between amino acids and either sugars (i.e., producing melanoidins) or oxidised lipids. If recalcitrant organics like N-heterocyclic polymers or kerogen contribute to the retention of surviving endogenous amino acids, such a process might occur relatively early or late during the taphonomic process (i.e., at different points along the decomposition of proteins).

Based on the correlation between the black colouration and recrystallisation in the *M. megadermus* A, one might hypothesize that calcite dissolution promotes kerogen or N-heterocyclic polymer formation, freeing trapped reactants and allowing for them to mix more easily to ultimately condense into resistant organic geopolymers. Experimental production of melanoidin can be done using Gly as a reactant, but subsequent acid hydrolysis of the melanoidin product was observed to yield <1 % Gly, suggesting that Gly is ultimately modified

fossilisation process or depositional environment has yet been reported that is anhydrous throughout the entirety of the taphonomic process.

If fully hydrolysed free amino acids (a subset of the original amino acid composition of the starting proteins) are the only proteinaceous surviving remnants in Mesozoic fossils not subsequently condensed into a highly altered geopolymer, then this would preclude obtaining any peptide sequence information. However, the capacity of eggshell calcite to maintain a closed system deep into the fossil record, as suggested by the results here, indicates that a broader sampling in both number, locality, and age of Mesozoic eggshells will likely provide clearer insight into patterns of ancient amino acid preservation in this system. The concentrations of amino acids in the LACM Auca Mahuevo titanosaurian eggshells and Spanish titanosaurian eggshell from two of the localities are far lower than those of the thicker *M. megadermus* specimens as well as Spanish titanosaurian eggshell from the other three localities as well as the Chinese putative hadrosauridae eggshell (Fig. 1), indicating that amino acid preservation can vary between fossil eggshell of similar age and geologic provenance, calling for further study into the specific conditions that promote biomolecular preservation in biomineralized fossils. Although our results do not provide unambiguous indication of Phe preservation in the fossils (consistent with their low concentrations in untreated modern avian eggshell), the fact that the side chain of Phe bears a highly stable aromatic ring might confer it stability through fossilization in some cases. A similar argument could be made for Ile and its the simple hydrocarbon side chain.

Future work, like that reported here, on sub-fossil and fossil eggshell will help to calibrate experimental studies of organic degradation in closed systems. Short, intense thermal maturation experiments may sometimes be inappropriate to compare to specimens that have spent longer periods of time at relatively lower temperatures (Tomiak *et al.* 2013). For example, protein three-dimensional structure might affect rates of hydrolysis and racemisation (Collins *et al.* 1999) and denaturation can occur under elevated temperature more typical of experimental maturation than natural early taphonomic settings. The closed system conditions experienced by intra-crystalline amino acids helps to avoid confounding effects due to leaching of amino acids, pH changes, contamination, and microbial decay (Child *et al.* 1993; Walton 1998; Crisp *et al.* 2013), so a deep fossil record of eggshells allows for studying long-term protein degradation in completely natural closed system environments. However, for Mesozoic eggshell, it is reasonable to assume that some degree of diagenesis (or possibly even catagenesis) will have taken place. For example, geothermal gradients can potentially expose buried eggshell to, or above, denaturation temperatures of some proteins, i.e., 50–80 ºC (Roos 1995). Therefore, thermal maturation remains a useful experimental tool for studying organic degradation in fossils of appreciable age and thermal maturity.

Very ancient amino acids might yield insights into palaeobiology in addition to organic geochemistry, potentially preserving taxonomic signatures in the amino acid profiles of fossils, as seen in calcium carbonate biominerals (Jope 1967; King & Hare 1972; Andrews *et al.* 1984; Haugen *et al.* 1989; Kaufman *et al.* 1992; Hincke *et al.* 1995; Mann & Siedler 1999; Miller *et al.* 2000; Lakshminarayanan *et al.* 2002, 2003; Crisp *et al.* 2013; Demarchi *et al.* 2014). Such potential insight depends on the presence of sufficient variation in the original concentrations of stable amino acids of non-avian and avian dinosaur eggshells so as to be able to detect differences in original protein content after significant diagenesis and degradation. At the very least, endogenous, ancient amino acids and other fossil organics are good candidates for compound-specific stable isotope analysis (e.g., C, O, or N) without the likelihood of incorporated environmental isotopes altering the observed ratios, similar to a recent study of amino acid-specific nitrogen isotopes in modern bivalve shells (Huang *et al.* 2023).

As far as pushing the upper age limit for well-supported amino acids, calcified eggshell represents a fairly limited fossil record. Examining the fossils of other calcite biominerals, such

as mollusc/brachiopod shells or trilobite cuticles/eye lenses, might provide opportunities to detect demonstrably ancient, endogenous amino acids throughout the Palaeozoic.

5. Conclusions: Mesozoic dinosaur eggshells from multiple localities (*M. megadermus*, putative hadrosaurid, and three localities of Spanish titanosaur) show strong chemical evidence for the presence of highly stable ancient, endogenous amino acids in THAA compositional profiles, D/L ratios, and total estaimted THAA concentrations, although with varying degrees of preservation across localities (e.g., weak signals from Auca Mahuevo titanosaurians and two localities of Spanish titanosaurs). Although eggshell calcite is known to act as an extremely efficient closed system, these results are still about an order of magnitude older than the oldest reported eggshell amino acids and an estimated ~56–42 million years older (Titanosaurian eggshell UAM2a, Requena, Valencia, Spain [Table 1]) than the oldest reported amino acids in biocalcite fossils for which there is unambiguous evidence (~30 Ma mollusc opercula [Penkman *et al*. 2013]), potentially making these amino acids the best supported amino acids from non-avian dinosaur fossils. These results bolster excitement of the potential for eggshell calcite to aid in the study of ancient organic degradation. As for their level of preservation, the amino acids appear to be predominantly hydrolysed; this has negative implications for the likelihood of highly preserved Mesozoic peptides and proteins, especially from open systems like bone or integument. The closed system nature of eggshell calcite also highlights that there are two general aspects of molecular preservation in fossils: stability of the original molecule (e.g., against microbial/autolytic decay or diagenesis) and mobility of the molecule and its degradation products (e.g., solubility or the degree of openness of the matrix). However, the methods used here should be repeated on other Mesozoic eggshell samples (and surrounding sediment matrix controls) alongside the addition of analyses (e.g., principal component) of large amino acid datasets to better characterise diagenetic patterns in ancient eggshells. Eggshell calcite diagenesis and closed system behaviour might also possibly be further examined using methods applied to carbonate alteration in the geologic record, such as clumped isotope geochemistry (Eiler 2007) and Ca/Mg isotopic analysis (Fantle & Higgins 2014).

Acknowledgments: We thank Jennika Greer (University of Chicago) for assistance in aseptic polishing of the eggshells that were not embedded in a matrix, Sheila Taylor (University of York) for assistance in preparing the samples for RP-HPLC, Matthew Von Tersch (University of York) for photographing the LACM specimens, Erzsebet Thornberry (University of Bristol) for assistance in Py-GC-MS data analysis, Claudia Hildebrandt (University of Bristol) for assistance in polarized light microscopy, and Prof. Jesper Velgaard Olsen (Novo Nordisk Center for Protein Research, University of Copenhagen) for providing MS access and resources. Work supported by the University of Bristol Bob Savage Memorial Fund. We thank Maureen Walsh, Luis Chiappe, and Lina Candrella of the Natural History Museum of Los Angeles County for access to Auca Mahuevo specimens. We thank Rodolfo Coria (Universidad Nacional de Río Negro), Leonardo Salgado (Universidad Nacional de Río Negro), Juliana Sterli (Museo Paleontológico Egidio Feruglio), Pablo Tubaro (Museo Provincial Patagónico de Ciencias Naturales), Gabriela Costanzo, Ernesto Rodrigo Paz, José Luis Garrido, Asociación Paleontológica Argentina, and National Authority of the Application of the Law of Paleontological Heritage for assisting in our efforts to repatriate the independently acquired *M. megadermus* specimens to Argentina. We also thank two anonymous reviewers and the editors Robert Asher (University of Cambridge), Jérémy Anquetin (Jurassica Museum), and Guillaume Billet (Muséum national d'Histoire naturelle) at PCI Paleo for helpful comments on our previously drafted preprint. Finally, we thank three anonymous reviewers that commented upon the manuscript as published here. *Gallus gallus*

domesticus (Public Domain Dedication 1.0, https://creativecommons.org/publicdomain/zero/1.0/legalcode), *Struthio camelus* (Lukasiniho, Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported, https://creativecommons.org/licenses/by-nc-sa/3.0/legalcode, CC BY-NC-SA 3.0), hadrosaur (Iain Reid, Attribution 3.0 Unported, https://creativecommons.org/licenses/by/3.0/legalcode, CC BY 3.0), and titanosaurians (T. Tischler, Creative Commons Attribution-ShareAlike 3.0 Unported, https://creativecommons.org/licenses/by-sa/3.0/legalcode, CC BY-SA 3.0; Ryan Santos Soledade, CC0 1.0 Universal Public Domain Dedication, https://creativecommons.org/publicdomain/zero/1.0/legalcode; Scott Hartman, Attribution-NonCommercial-ShareAlike 3.0 Unported, https://creativecommons.org/licenses/by-nc-sa/3.0/legalcode, CC BY-NC-SA 3.0) silhouettes were obtained from phylopic.org with some color modifications.

Competing Interests: We declare no competing interests.

Appendix A. Supplementary Material: Within this supplementary appendix, the reader can find further details of methods, descriptions, figures, and tables as they relate to the following topics: materials, resin-embedded thin sections, light microscopy/LSF imaging/photography, RP-HPLC amino acid analysis, LC-MS/MS, Py-GC-MS, aseptic polishing protocol, TOF-SIMS, Raman spectroscopy, additional eggshell micrographs/photographs/records, and supplemental references.

Data Availability
Data are available through figshare at https://doi.org/10.6084/m9.figshare.23784300.

References:
Abelson, P.H., 1954. Amino acids in fossils. Science 119, 576.
Abbott, G.D., Fletcher I.W., Tardio S., Hack E., 2017. Exploring the geochemical distribution of organic carbon in early land plants: a novel approach. Phil. Trans. R. Soc. B373, 20160499.
Abbott, G.D., Bashir, F.Z., 2021. Petrified Flora and Fauna from the Carboniferous – Peering into the Past 300 Million Years Ago, in: Hüglin, S., Gramsch, A., Seppänen, L. (Eds.), Petrification Processes in Matter and Society. Themes in Contemporary Archaeology, Springer, Cham, pp. 45–52.
Allard, M.W., Young, D., Huyen, Y., 1995. Technical comment: detecting dinosaur DNA. Science 268, 1192–1192.

Bada, J.L., Wang, X.S., Hamilton, H., 1999. Preservation of key biomolecules in the fossil record: current knowledge and future challenges. Philos. Trans. R. Soc. Lond. B Biol. Sci. 354, 77–87.

Baker, J.R., Balch, D.A., 1962. A study of the organic material of hen's-egg shell. Biochem. J. 82, 352–361.

Barroso-Barcenilla F, Cambra-Moo, O, Segura, M., 2010. Estudio preliminar sobre Geología y Tafonomía del yacimiento paleontológico de "Lo Hueco" (Cretácico Superior, Cuenca, España). Bol. R. Soc. Esp. Hist. Nat. Sec. Geol. 104, 57–70.

Collins, M.J., Riley, M.S., 2000. Amino acid racemization in biominerals: the impact of protein degradation and loss, in Goodfriend, G. A., Collins, M. J., Fogel, M. L., Macko, S. A., Wehmiller, J. F. (Eds), Perspectives in Amino Acid and Protein Geochemistry. Oxford University Press, Oxford, pp. 120–142.

Collins, M.J., Westbroek, P., Muyzer, G., De Leeuw, J.W., 1992. Experimental evidence for condensation reactions between sugars and proteins in carbonate skeletons. Geochim. Cosmochim. Acta 56, 1539–1544.

Collins M.J., Walton, D., King, A., 1998. The Geochemical Fate Of Proteins., in: Stankiewicz, B.A., Van Bergen, P.F. (Eds), Nitrogen-Containing Macromolecules in the Bio- and Geosphere. Oxford University Press, Oxford, pp. 74–77.

Collins, M.J., Waite, E.R., Van Duin, A.C.T., 1999. Predicting protein decomposition the case of aspartic-acid racemization kinetics. Philos. Trans. R. Soc. Lond. B Biol. Sci. 354, 51–64.

Collins, M.J., Walton, D., Curry, G.B., Riley, M.S., Von Wallmenich, T.N., Savage, N.M., Muyzer, G., Westbroek, P., 2003. Long-term trends in the survival of immunological epitopes entombed in fossil brachiopod skeletons. Org. Geochem. 34, 89–96.

Greenwalt, D.E., Goreva, Y.S., Siljeström, S.M., Rose, T., Harbach, R.E., 2013. Hemoglobin-derived porphyrins preserved in a Middle Eocene bloodengorged mosquito. Proc. Natl. Acad. Sci. U.S.A. 110, 18496–18500.

King Jr., K., Hare, P.E., 1972. Amino acid composition of the test as a taxonomic character for living and fossil planktonic foraminifera. Micropaleontology 18, 285–293.

Kjær, K.H., Winther Pedersen, M., De Sanctis, B., De Cahsan, B., Korneliussen, T.S., Michelsen, C.S., Sand, K.K., Jelavić, S., Ruter, A.H., Schmidt, A.M.A., Kjeldsen, K. K., Tesakov, A.S., Snowball, I., Gosse, J.C., Alsos, I.G., Wang, Y., Dockter, C., Rasmussen, M., Jørgensen, M.E., Skadhauge, B., Prohaska, A., Kristensen, J.A., Bjerager, M., Allentoft, M.E., Coissac, E., PhyloNorway Consortium, Rouillard, A., Simakova, A., Fernandez-Guerra, A., Bowler, C., Macias-Fauria, M., Vinner, L., Welch, J.J., Hidy, A.J., Sikora, M., Collins, M.J., Durbin, R., Larsen, N.K., Willerslev, E., 2022. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. Nature 612, 283–291.

Lakshminarayanan, R., Kini, R.M., Valiyaveettil, S., 2002. Investigation of the role of ansocalcin in the bio-mineralization in goose eggshell matrix. Proc. Natl. Acad. Sci. U.S.A. 99, 5155–5159.

Lakshminarayanan, R., Valiyaveettil, S., Rao, V.S., Kini, K.M., 2003. Purification, characterization, and in vitro mineralization studies of a novel goose eggshell matrix protein, ansocalcin. J. Biol. Chem. 278, 2928–2936.

of the type Cromerian of West Runton, Norfolk, UK. Quat. Int. 228, 25–37.
Penkman, K.E., Preece, R.C., Bridgland, D.R., Keen, D.H., Meijer, T., Parfitt, S.A., White, T.S., Collins, M.J., 2013. An aminostratigraphy for the British Quaternary based on Bithynia opercula. Quat. Sci. Rev. 61, 111–134.

Geochronol. 16, 87–109.
Towe, K.M., 1980. Preserved organic ultrastructure: an unrealiable indicator for Paleozoic amino acid biogeochemistry, in: Hare, P.E., Hoering, T.C., King Jr., K. (Eds.), Biogeochemistry of Amino Acids. Wiley, New York, pp. 65–74.
Towe, K.M., 1973. Trilobite eyes: calcified lenses in vivo. Science 179, 1007–1009.
Towe, K.M., Thompson, G.R., 1972. The structure of some bivalve shell carbonates prepared by ion-beam thinning. Calcif. Tissue Res. 10, 38–48.

Zischler, H., Hoss, M., Handt, O., Von Haeseler, A., Van Der Kuyl, A.C., Goudsmit, J., 1995. Technical comment: detecting dinosaur DNA. Science 268, 1192–1193.

Table and Figure Legends:

Table 1. Summary of fossil eggshell samples studied. *Sample underwent extensive methodological analyses.

Table 2. Methodological triangulation employed in this study.

Table 3. Bleached fossil eggshell amino acid racemisation and [Ser]/[Ala] values rounded to the nearest hundredth and then averaged across replicates, with standard deviations (in italics) reported for samples with more than one replicate. NA indicates that amino acid concentration was consistently below detection limit or that standard deviation cannot be calculated because only one replicate is above detection limit. *Data from elution time > 58 min is of low accuracy due to elevated baseline values. LACM 7324 A and B sample replicates are presented alongside their subsample number (i.e., 1 or 2). Unlike the separately presented LACM 7324 A and B fragments that derive from the same locality, UC and UM eggshell from the same locality with multiple fragments (i.e., denoted a, b, and c) were averaged together in this table for simpler presentation.

Table 4. Peptides detected by LC-MS/MS in the bleached *M. megadermus* A (Turin and Copenhagen replicates) and their significant matches to known proteins. Note that the asparagine and glutamine are undeamidated in peptide DNIQGITK (Histone 4), supporting its modern origins. Underlined methionines are oxidised. Although there are various ways to calculate the significance of a putative peptide sequence from mass spectral data, PEAKS software first uses a linear discriminant function (LDF) to calculate peptide-spectrum match quality (i.e., determining the most likely database peptide match for each spectrum and discriminating against false identifications) using factors like *de novo* sequence-database sequence similarity and the matching of spectral peaks and the fragment ions. The LDF score is then converted to a P-value such that the P-value equals the probability that a false identification has a greater score than the observed score (i.e., greater P-values indicate greater probabilities that the peptide-spectrum match is due to random chance). The P-value is then converted according to $-10*\log_{10}$(P-value) to yield what is called a -10lgP for easier interpretation. A greater -10lgP indicates a more significant result, and typically speaking, values above 20 are considered significant since they correspond to a P-value of 0.01 (Zhang *et al*. 2012).

Figure 1. Dinosaur eggshell analysed in this study under light microscopy/photography. A–L, titanosaurian *M. megadermus* A. A, large fragment of *M. megadermus* A viewed from the exterior surface showing ornamentation as well as some underlying black, amorphous calcite revealed when surface layers flaked off during splitting with a pestle. B, amorphous, black calcite viewed from exterior that was exposed. Exterior surface ornamentation under white light, C, and LSF, D. Cross section through the entire eggshell with the exterior surface to the top of the panel under white light, E, and LSF, F. Thin section of entire eggshell cross-section with exterior surface to the left of the panel under plane-, G, and cross-, H, polarised light. Thin section of eggshell exterior ornamentation with the exterior surface to the left of the panel under plane-, I, and cross-, J, polarised light. Thin section of recrystallised interior calcite under plane-, K, and cross-, L, polarised light. M–P, titanosaurian *M. megadermus* B. M, *M.*

megadermus B viewed from the exterior surface showing ornamentation. N, a weathered edge of the eggshell revealing palisade/column crystals. O–P, freshly broken edge of the eggshell showing brown staining of the calcite crystals with the exterior surface to the top of the panel. Q–T titanosaurian LACM 7324. Q–R, LACM 7324 A. Q, interior surface. R, largely freshly broken edge showing brown staining of calcite crystals with interior surface to the top of the panel. S–T, titanosaurian LACM 7324 B. S, exterior surface. T, freshly broken edge showing brown staining of calcite crystals with interior surface to the top of the panel. U–Z, Spanish titanosaurian from five localities (cf. *Megaloolithus*). U, UAM1b. V, UAM2a. W, UAM3a. X–Y, UAM4a with cross section. Z, UAM5a. AA–AB, Chinese putative hadrosauridae. AA, UC1a viewed from the exterior surface. AB, UC1b viewed from cross section.

Figure 2. Raman spectroscopy of *M. megadermus* A resin-embedded thin section. A, transmitted light micrograph with area mapped outlined in red. Dark regions appear transparent, whereas light regions appear brown. The five general regions from which spectra were taken in panel C are labelled with their two-letter abbreviation (note that these do not indicate the precise points where the spectra were taken). B, Whole-spectrum map (i.e., all wavenumbers) under ~100 μW laser power. C, Spectra from the dark/recrystallized (20 mW laser power) and light/non-recrystallized (~100 μW laser power) regions. Inorganic reference peak positions (Handbook of Raman Spectra for Geology, Laboratoire de Géologie de Lyon, Université de Lyon) shown with grey and brown vertical lines. Vertical blue lines indicate the prominent peaks detected in the surrounding epoxy embedding resin, which could contribute in part to certain peaks in the eggshell. Some of the peaks may be genuine organic vibrations, but are strongly reminiscent of artefactual quasi-periodic ripples, especially in the light regions.

Figure 3. Comparison of identified pyrolysis products in, A, modern chicken (ethanol rinsed before powdering) and, B, *M. megadermus* A (not ethanol rinsed before powdering) eggshell after DCM rinsing and Soxhlet extraction.

Figure 4. THAA compositional profiles of modern, experimental, and ancient eggshell. A,

Sample name	*M. megadermus* A* (MPCN-PV-900.1; Thin section: MPCN-PV-900.3)	*M. megadermus* B (MPCN-PV-900.2)	LACM 7324 A	LACM 7324 B	UC 1a (LH PV 51; Long Hao collection)	UC 1b (LH PV 51; Long Hao collection)	UAM1a-c Titanosaur (cf. *Megaloolithus*)	UAM2a Titanosaur (cf. *Megaloolithus*)	UAM3a Titanosaur (cf. *Megaloolithus*)	UAM4a-b Titanosaur (cf. *Megaloolithus*)	UAM5a Titanosaur (cf. *Megaloolithus*)
Amino acid evidence	Strong	Strong	Weak	Weak	Strong	Strong	Strong	Strong	Weak	Weak	Strong
Origin	Commercial (USA)	Commercial (Denmark)	Collected by LACM		Collected by UC		Collected by UAM				
Ootaxon	*Megaloolithus megadermus*		*Fusioolithus baghensis*		Hadrosauridae?		*Megaloolithus siruguei*?		*Megaloolithus mammillare?*	*Megaloolithus siruguei?*	
Collection	Museo Provincial Patagónico de Ciencias Naturales (General Roca, Río Negro, Argentina)		Natural History Museum of Los Angeles County (Los Angeles,		University of Chicago		Universidad Autónoma de Madrid				

		California, USA)						
Locality	Bajos de Santa Rosa (Berthe II), Río Negro Province, Argentina	Auca Mahuevo, Neuquén Province, Argentina	San Ge Quam locality, Central Junggar, Xinjiang, China	La Rosaca, Burgos, Spain	Requena, Valencia, Spain	Bastús, Lleida, Catalonia	Biscarri, Lleida, Catalonia	Portilla, Cuenca, Spain
Formation	Allen	Anacleto	Ailikehu	Calizas de Lychnus	Sierra Perenchiza	Arén	Tremp	Villalba de la Sierra
Age	Late Cretaceous; Middle Campanian – Early Maastrichtian; ~73–69 Ma	Late Cretaceous; Early–Middle Campanian; ~83–74.5 Ma	Late Cretaceous; Maastrichtian; ~72–66 Ma	Late Cretaceous; Maastrichtian; ~72–66 Ma	Late Cretaceous; Santonian–Campanian; ~86–72 Ma	Late Cretaceous; Campanian–Maastrichtian; ~84–66 Ma	Late Cretaceous; Late Maastrichtian; ~67.6–66 Ma	Late Cretaceous; Early Campanian–Maastrichtian; ~84–66 Ma
Relevant sources	Mohabey 1998; Fernández 2014; Fernández & Khosla 2015; Dhiman *et al.* 2019; Khosla & Lucas 2020; Fernández *et al.* 2022	Chiappe *et al.* 1998, 2003, 2005; Dingus *et al.* 2000; Grellet-Tinner *et al.* 2004; Garrido 2010; Fernández &	Pei-ji 1983; Zhang 2010	Moratalla & Melero 1987; Moratalla 1993; Vinaed-Llynaud & López-Martinez 1997; Izquierdo *et al.* 1999; Company 2004; Gil *et al.* 2004; Barroso-Barcenilla *et al*. 2010; Sellés & Galobart 2014; Company 2019; Sanguino *et al*. 2022				

		Khosla 2015		

Technique	**Signal analyzed**	**Potential insight into protein preservation**	**Example references**
Light microscopy / photography	Plane or crossed polarized, transmitted or reflected light	Integrity of calcite crystal structure (i.e., system dynamics); localization of dark organic material	Hirsch & Quinn 1990
LSF	Fluoresced light	Localization of non-fluorescing organic material	Kaye *et al.* 2015
Raman spectroscopy	Raman-active molecular vibrations	Presence and localization of molecules consistent with amino acids, proteins, or organic geopolymers, assuming no quasi-periodic artefacts	Wiemann *et al.* 2018; Alleon *et al.* 2021
Py-GC-MS	Pyrolysis decomposition products of molecules	Presence of molecules consistent with amino acids, proteins, or organic geopolymers	Saitta *et al.* 2017
TOF-SIMS (supplemental material)	Secondary ions from fragmented molecules	Presence and localization of molecules consistent with amino acids, proteins, or organic geopolymers	Orlando *et al.* 2013

RP-HPLC	13 primary amino acids in their relevant chiral forms	Amino acid concentration, composition, racemization extent, and hydrolysis extent; endogenicity of amino acids and any preserved peptides	Crisp *et al.* 2013
LC-MS/MS	Peptide sequences	Endogenicity of any recovered peptides; if endogenous, evolutionary information	Demarchi *et al.* 2016

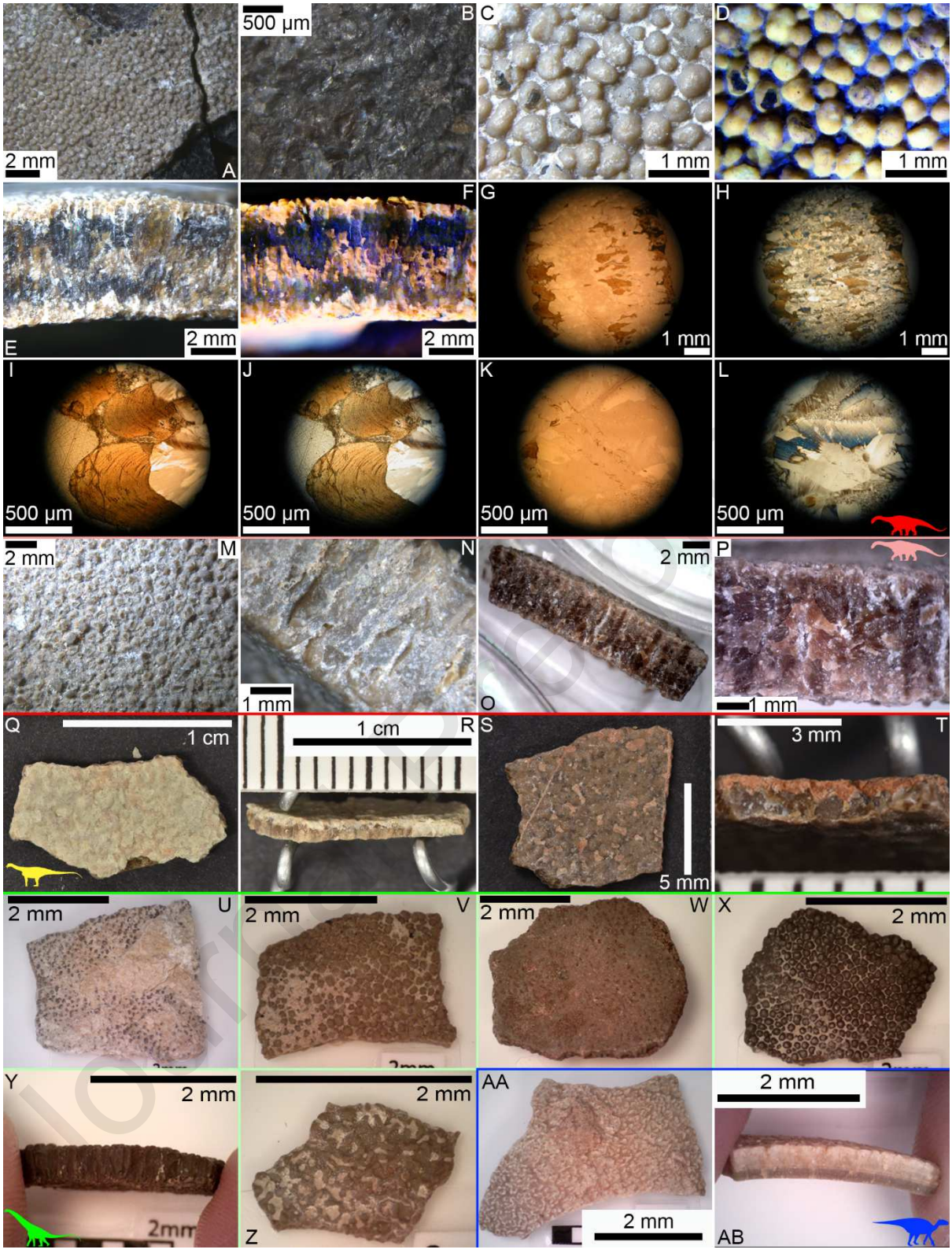
Treatment	Sample	Total analytical replicates	Glx D/L	Ala D/L	Val D/L*	Asx D/L	Ser D/L	[Ser]/[Ala]
Bleached FAA	*M. megadermus* A	1	1.03	0.93	1.22	0	NA	0
	M. megadermus B	4	1.015	0.93	1.255	0.98	0.955	0.01
		Standard deviations	*0.0208166*	*0.011547005*	*0.1410673*	*0.057154761*	*0.595119036*	*0.008164966*
Ethanol rinsed before powdering, bleached FAA	*M. megadermus* A	1	1.05	0.97	1.11	NA	NA	0
Bleached, 24-hr hydrolysis THAA	*M. megadermus* A	1	1.04	0.96	1.29	NA	NA	0
	M. megadermus B	3	0.99	0.69	1.186666667	0.23	0.043333333	0.093333333

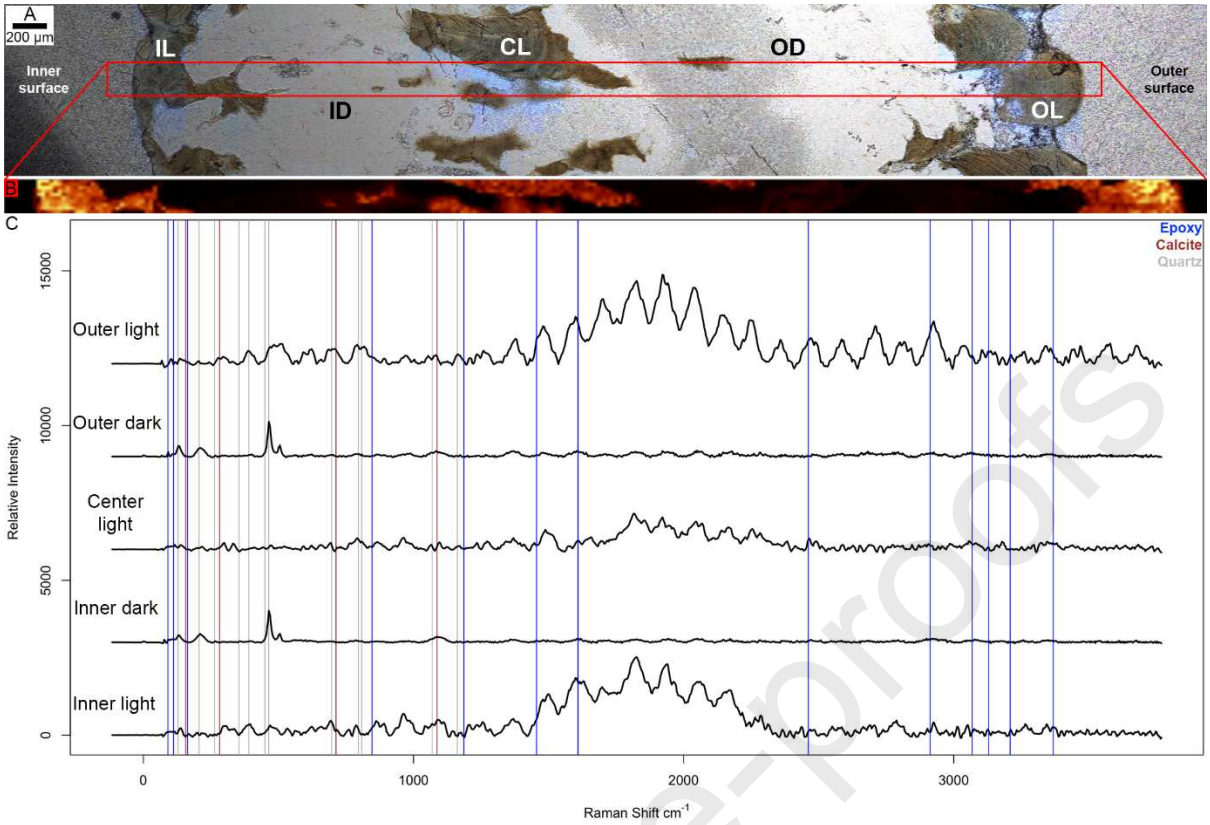
		Standard deviations	*1.35974E-16*	*0.01732050808*	*0.270985854*	*0*	*0.051316014*	*0.005773503*
Ethanol rinsed before powdering, bleached, 24-hr hydrolysis THAA	*M. megadermus* A	1	1.02	0.93	1.14	NA	NA	0
Bleached FAA	LACM 7324 A1 & A2	4	0.2	1.015	0	NA	0	9.0075
		Standard deviations	*0.23323807 6*	*0.04725815 6*	*0*	*NA*	*NA*	*17.995005 56*
	LACM 7324 B1 & B2	4	0.1125	1.0125	0	NA	NA	9.25
		Standard deviations	*0.225*	*0.00957427 1*	*0*	*NA*	*NA*	*18.5*
Bleached, 24-hr hydrolysis THAA	LACM 7324 A1 & A2	4	0.7775	0.92	0.3175	0	0	9.5525
		Standard deviations	*0.01258305 7*	*0.06976149 8*	*0.10812801 1*	*0*	*0*	*18.965001 76*
	LACM 7324 B1 & B2	4	0.585	0.895	0.3975	0	0	9.87
		Standard deviations	0.03316624 8	0.02886751 3	0.06238322 4	0	0	19.420005 15
Bleached FAA	UC1a-b	4	1.035	1.06	1.81	1.015	0	0.00225

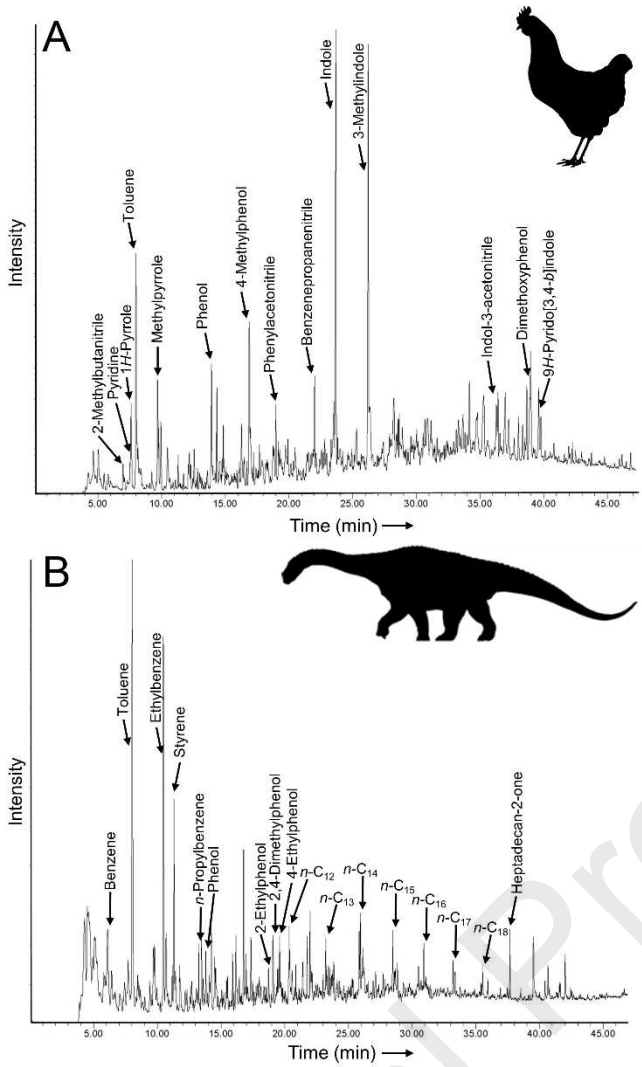
		Standard deviations	*0.042 0317 34*	*0.035 5902 61*	*0.068 7992 25*	*0.116 7618 66*	*0*	*0.001 7078 25*
Bleached, 24-hr hydrolysis THAA	UC1a-b	4	1.047 5	1.045	2.73	0.75	0.15	0.006
		Standard deviations	*0.015*	*0.038 7298 33*	*0.147 1960 14*	*0.102 3067 28*	*0.212 1320 34*	*0.009 5219 05*
Bleached FAA	UM1a-c	6	1.045	1	1.096 6666 67	0.933 3333 33	0	0.000 5
		Standard deviations	*0.017 6068 17*	*0.016 7332 01*	*0.045 4606 06*	*0.092 2315 93*	*NA*	*0.001 2247 45*
Bleached, 24-hr hydrolysis THAA	UM1a-c	6	1.055	1.011 6666 67	1.145	0.801 6666 67	0.133 3333 33	0.002 1666 67
		Standard deviations	*0.005 4772 26*	*0.020 4124 15*	*0.025 8843 58*	*0.043 5507 37*	*0.230 9401 08*	*0.002 5625 51*
Bleached FAA	UM2a	2	1.045	0.985	1.4	0.9	NA	0
		Standard deviations	*0.021 2132 03*	*0.007 0710 68*	*0*	*0.014 1421 36*	*NA*	*0*
Bleached, 24-hr hydrolysis THAA	UM2a	2	1.055	1.025	1.605	0.83	0.24	0.005
		Standard deviations	*0.007 0710 68*	*0.021 2132 03*	*0.063 6396 1*	*0*	*NA*	*0.007 0710 68*
Bleached FAA	UM3a	3	NA	NA	NA	NA	NA	NA

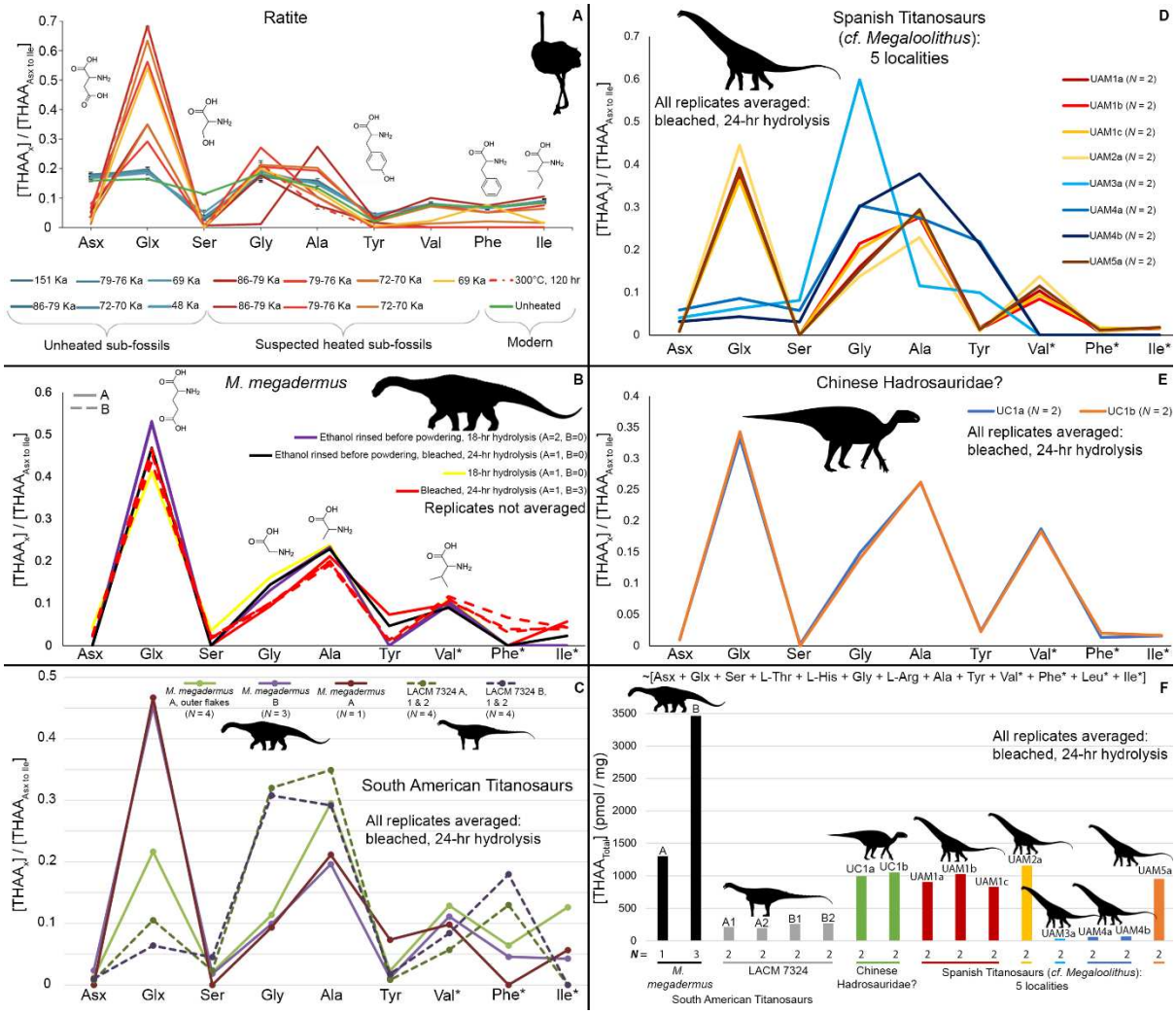
		Standard deviations	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Bleached, 24-hr hydrolysis THAA	UM3a	2	0.375	0.355	NA	0.075	0	0.705
		Standard deviations	*0.007 0710 68*	*0.021 2132 03*	*NA*	*0.106 0660 17*	*0*	*0.035 3553 39*
Bleached FAA	UM4a-b	4	0.732 5	NA	NA	1.002 5	0	0
		Standard deviations	*0.618 5130 02*	*NA*	*NA*	*0.059 0903 26*	*NA*	*0*
Bleached, 24-hr hydrolysis THAA	UM4a-b	4	0.307 5	6.027 5	NA	0.36	0	0.15
		Standard deviations	*0.112 3610 25*	*2.999 4930 13*	*NA*	*0.076 1577 31*	*0*	*0.071 1805 22*
Bleached FAA	UM5a	2	1.05	0.945	1.195	0.94	NA	0
		Standard deviations	*0.014 1421 36*	*0.063 6396 1*	*0.021 2132 03*	*0.014 1421 36*	*NA*	*0*
Bleached, 24-hr hydrolysis THAA	UM5a	2	1.05	1	1.34	0.91	0	0.001
		Standard deviations	*0*	*0.014 1421 36*	*0.028 2842 71*	*0.070 7106 78*	*NA*	*0.001 4142 14*

Sample preparation	Protein name	Peptide	-10lgP	Number of Spectra
Turin	Histone H4 [*Gallus gallus*]	TVTAMDVVYALK	31.18	2
		ISGLIYEETR	25.28	1
		DNIQGITK	25.2	1
	Isoform 2 of Histone H2B type 2-F [*Homo sapiens*]	AMGIMNSFVNDIFER	37.15	2
	Keratin, type I cytoskeletal 9 [*Homo sapiens*]	SRSGGGGGGGLGSGGSIRSSY	30.04	1
	Keratin, type II cytoskeletal 4 [*Homo sapiens*]	LALDIEIATYR	27.43	1
	POTE ankyrin domain family member I [*Homo sapiens*]	AGFAGDDAPR	21.13	1
Copenhagen	Keratin, type I cytoskeletal 9 [*Homo sapiens*]	GGGGGGGLGSGGSIRSS	24.14	1
		SRSGGGGGGGLGSGGSIRSSY	23.09	1
		SGGGGGGGLGSGGSIR	21.02	1


A
2 mm
B
500 µm
C
1 mm
D
1 mm
E
2 mm
F
2 mm
G
1 mm
H
1 mm
I
500 µm
J
500 µm
K
500 µm
L
500 µm
M
2 mm
N
1 mm
O
2 mm
P
1 mm
Q
1 cm
R
1 cm
S
5 mm
T
3 mm
U
2 mm
V
2 mm
W
2 mm
X
2 mm
Y
2 mm
2mm
Z
2 mm
AA
2 mm
AB
2 mm







Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: