Chemical Geology 663 (2024) 122266

Contents lists available at ScienceDirect

Chemical Geology

journal homepage: www.elsevier.com/locate/chemgeo

Phosphorus in belemnites: Extraction, quantification, and variability

Ailsa C. Roper [a,*], Yijun Xiong [a], Yafang Song [a], Crispin T.S. Little [a], Simon W. Poulton [a], Paul B. Wignall [a], Clemens V. Ullmann [b], Robert J. Newton [a]

[a] School of Earth and Environment, University of Leeds, Leeds LS2 9JT, United Kingdom
[b] Camborne School of Mines, Tremough Campus, University of Exeter, Penryn TR10 9EZ, United Kingdom

ARTICLE INFO

Editor: Dr. Vasileios Mavromatis

Keywords:
Phosphorus
Carbonate
Belemnite
Nutrient cycling
Method development
Variability
Calcium carbonate
Calcite
Carbonate associated phosphorus

ABSTRACT

Phosphorus is generally considered the ultimate limiting nutrient for marine primary productivity over geological timescales and plays a key role in modulating several biogeochemical cycles. Most established methods for investigating P cycling do not provide direct evidence for water-column P concentrations, but recent work on carbonate associated phosphorus (CAP) has shown there is a potential to record ancient dissolved P concentrations. We present a method to extend the application of carbonate associated P measurements to belemnites, and we quantify variability in P contents within and between belemnite specimens from two stratigraphic levels in Jurassic rocks from the Yorkshire coast, UK, as well as in modern analogues of belemnites.

We show that there is little difference in P measurements between different preparative methods in uncontaminated belemnite calcite samples. In samples with higher levels of contaminant phases of P (P other than CAP), or a greater extent of diagenetic alteration, cleaning with non-oxidative methods and dissolution in weak acids (acetic) was found to minimise the impact of contamination on the measured P contents. P concentrations vary within and between specimens, but variations are not a result of taxonomic differences, and overall P measurements are reproducible between replicate samples, within individual belemnites and within stratigraphic levels. There were statistically significant differences in belemnite P concentrations between the stratigraphic levels studied here, indicating the potential for this technique to be used to measure changes in belemnite CAP through time.

1. Introduction

Phosphorus is an essential nutrient for life, used in all living organisms to build cell membranes, RNA, DNA, and hard tissues. P is therefore a key nutrient controlling marine primary productivity (e.g., Delaney, 1998; Tyrrell, 1999; Paytan and McLaughlin, 2007; Filippelli, 2008) and is generally considered the ultimate limiting nutrient on geological timescales (Tyrrell, 1999). Understanding ancient P cycling is vital to understand the coevolution of the environment and life, as P availability controls the longer-term rate of oxygen production via its influence on organic carbon burial, hence influencing the evolution of the redox state of the oceans and atmosphere (e.g., Van Cappellen and Ingall, 1996; Lenton and Watson, 2000; Bjerrum and Canfield, 2002; Papineau, 2010; Ruttenberg, 2003; Planavsky et al., 2010; Lenton et al., 2012; Reinhard et al., 2016; Li et al., 2013; Alcott et al., 2022). On shorter timescales, increases in P supply or recycling have been suggested as a driver of ocean anoxic events and coincident extinctions across several intervals of geological time (e.g., Monteiro et al., 2012; Schobben et al., 2020; Hülse et al., 2021; Qiu et al., 2022).

Despite the environmental importance of P cycling, measuring changes of past ocean bioavailable P concentrations has proved challenging. Unlike other nutrient elements such as nitrogen or carbon, P has only one natural stable isotope, so isotope mass balance cannot be used to track changes in P fluxes. Established methods for investigating P cycling throughout Earth's history, such as P phase partitioning (Ruttenberg, 1992; Thompson et al., 2019), P/Fe ratios of iron formations (Bjerrum and Canfield, 2002; Planavsky et al., 2010), phosphorites distribution (Baturin and Bezrukov, 1979), and bulk rock P concentrations (Reinhard et al., 2016), have been successful in providing some constraints on the distributions and cycling of P in the past, but have their own limitations (e.g., Dodd et al., 2021). The distribution of phosphorite deposits and iron-formation P/Fe ratios have limited spatial

* Corresponding author.
E-mail addresses: A.C.Roper@leeds.ac.uk (A.C. Roper), Y.Xiong1@leeds.ac.uk (Y. Xiong), yf.song@ustc.edu.cn (Y. Song), C.T.S.Little@leeds.ac.uk (C.T.S. Little), S.Poulton@leeds.ac.uk (S.W. Poulton), P.B.Wignall@leeds.ac.uk (P.B. Wignall), c.ullmann@exeter.ac.uk (C.V. Ullmann), R.J.Newton@leeds.ac.uk (R.J. Newton).

https://doi.org/10.1016/j.chemgeo.2024.122266
Received 11 April 2024; Received in revised form 1 July 2024; Accepted 6 July 2024
Available online 8 July 2024

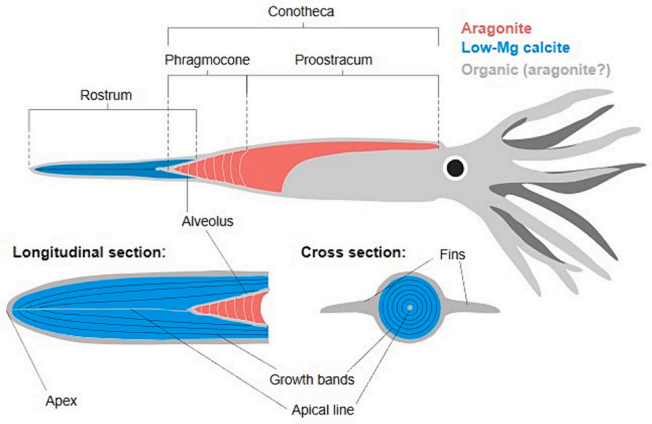

Fig. 1. Diagram of a living belemnite, as well as longitudinal and cross-sectional views of the rostrum. The rostrum is commonly preserved and hence may be subject to geochemical analyses. The phragmocone region is occasionally also preserved, but is commonly replaced by secondary minerals and sediment. The soft parts of the belemnite are rarely preserved, and are only known from a small number of specimens. Simplified after Hoffmann and Stevens (2020).

and temporal distributions. In addition, phosphorite deposits only represent elevated concentrations of P at the sediment water interface or in sediment porewaters, limiting nuanced interpretations of P availability in the water column, whilst P/Fe ratios in iron formations can be subject to diagenetic effects which can result in remobilisation and loss of P (Baturin and Bezrukov, 1979; Bjerrum and Canfield, 2002). Bulk rock concentrations offer insight into the delivery and retention of P in sediments, but contain no information on phase partitioning and may only be poorly related to the concentration of P in the water column due to variability in the redox-influenced recycling of P between the sediments and bottom waters (Ingall et al., 1993; Ruttenberg, 1992, 2003; Thompson et al., 2019; Alcott et al., 2022). P phase partitioning (Ruttenberg, 1992; Thompson et al., 2019) is perhaps the most sophisticated of these techniques and can provide valuable insight into P cycling in the sediment, but, as with bulk rock analysis, it is decoupled from water column phosphate bioavailability and can only provide inferences about P cycling in the water column (Ruttenberg, 1992). There is therefore a need for the development of a direct water-column P concentration proxy, which could be used to complement existing sediment data to better understand the phosphorus cycle (e.g., Dodd et al., 2021, 2023).

The P content of carbonates may be related to the concentration of phosphate in the solution from which the carbonate precipitates (e.g., Dodge et al., 1984; Shotyk et al., 1995; Kumarsingh et al., 1998; Montagna et al., 2006; LaVigne et al., 2008, 2010; Anagnostou et al., 2011; Mallela et al., 2013; Chen et al., 2019; Ingalls et al., 2020; Dodd et al., 2021; Ingalls et al., 2022; Richardson et al., 2022; Dodd et al., 2023; Roest-Ellis et al., 2023). This indicates the potential for carbonate-associated phosphorus (CAP) to be used as a proxy for water-column phosphate. Thus far, CAP has been investigated in a range of modern and ancient carbonates, including biogenic carbonates from foraminifera, gastropods, the brittle star family (Ophiocomidae), and corals (Dodge et al., 1984; Shotyk et al., 1995; Kumarsingh et al., 1998; Montagna et al., 2006; LaVigne et al., 2008, 2010; Anagnostou et al., 2011; Mallela et al., 2013; Chen et al., 2019; Dodd et al., 2021), natural abiotic carbonates such as dolomites and ooids (Ingalls et al., 2020; Dodd et al., 2021; Richardson et al., 2022; Dodd et al., 2023), and synthetic, laboratory-precipitated carbonates (Dodd et al., 2021; Richardson et al., 2022).

Here, we focus on different methods of extracting P from belemnite rostra. Belemnite rostra are the calcitic internal skeletons of belemnites, which are an extinct group of nektic animals belonging to the phylum Mollusca and the class Cephalopoda (Fig. 1). Belemnite rostra are potentially an ideal candidate for reconstructing changes in water column P concentrations for the Jurassic and Cretaceous periods as they are abundant in the fossil record of this period and have been shown to contain phosphate. P has been documented in belemnite calcite, but its occurrence has not yet been systematically studied (Longinelli et al., 2002; Gröcke et al., 2011; Doguzhaeva et al., 2013; Hoffmann et al., 2016; Hoffmann and Stevens, 2020). P is likely to be incorporated in several different ways into belemnite rostra, including phosphate associated with the carbonate, potentially as a substitution for the carbonate ion (Roest-Ellis et al., 2020; Richardson et al., 2022). Additionally, belemnite calcite, like most other forms of biogenic carbonate, contains around 1% organic matter which may host organic P (Wierzbowski and Joachimski, 2009). Other possible, but less likely, reservoirs include iron-bound P or micro-inclusions of apatite, which can be present in calcite (Richardson et al., 2022).

Our approach considers several important methodological aspects related to the extraction of belemnite P. Initially, we report the impact of different cleaning and dissolution methods on the extraction and quantification of P. Previous studies of elemental ratios in biogenic carbonates have either forgone any pre-cleaning step, or have used a range of cleaning methods to reduce contamination (defined here as phosphorus from phases other than CAP) from other sources, such as leaches with ultra-pure water, or oxidative cleaning with hydrogen peroxide or sodium hypochlorite (e.g., Shotyk et al., 1995; Penkman et al., 2008; Zhang et al., 2020; Riding, 2021 and references therein). Subsequently, we utilise a P phase partitioning approach on bulk samples to identify the dominant forms of P in belemnite calcite. These different forms of P will respond differently to the precise dissolution method applied. Previous studies of biogenic carbonates, including studies of belemnites, as well as CAP in other carbonate phases, have used strong acids, such as nitric acid or hydrochloric acid (e.g., Shotyk et al., 1995; Kumarsingh et al., 1998; Penkman et al., 2008), as well as weak acids, such as acetic acid (e.g., Ingalls et al., 2020; Dodd et al., 2021; Ingalls et al., 2022) to dissolve carbonate. Previous CAP studies have suggested that stronger acids are more likely to leach P from organic matter and other P containing minerals, such as different forms of apatite, as well as dolomite (which can be associated with P-containing iron-(oxyhydr)oxide minerals), thereby over-estimating CAP (Sælen, 1989; Ingalls et al., 2020, 2022). As CAP studies have not previously been investigated on belemnite calcite, hydrochloric and nitric acid were trialled, in addition to acetic acid, to determine which method was most suitable for belemnite calcite.

We also investigate variability in CAP within and between belemnites, as well as investigating the effect of factors such as size and interspecies variability. Previous work on belemnites has demonstrated that concentrations of some elements vary systematically within and between individuals (McArthur et al., 2007a; Li et al., 2012, 2013; Ullmann et al., 2015). Several studies have shown differences in some elemental ratios, such as Mg/Ca, between different species of belemnite, but some elemental ratios exhibited a subordinate species control (such as Na/Ca, Sr/Ca; e.g., McArthur et al., 2007; Li et al., 2012, 2013; Ullmann et al., 2015). Other studies have shown systematic spatial differences in elemental ratios within a belemnite rostrum, including variations between darker areas of calcite which represent more organic-rich growth rings, and variation along the length of the rostrum for ratios such as Mg/Ca and Sr/Ca (McArthur et al., 2007; Ullmann et al., 2015; Hoffmann et al., 2016). Our ultimate aim with this approach is to evaluate whether belemnites may be used to investigate water column P concentrations in Earth's past.

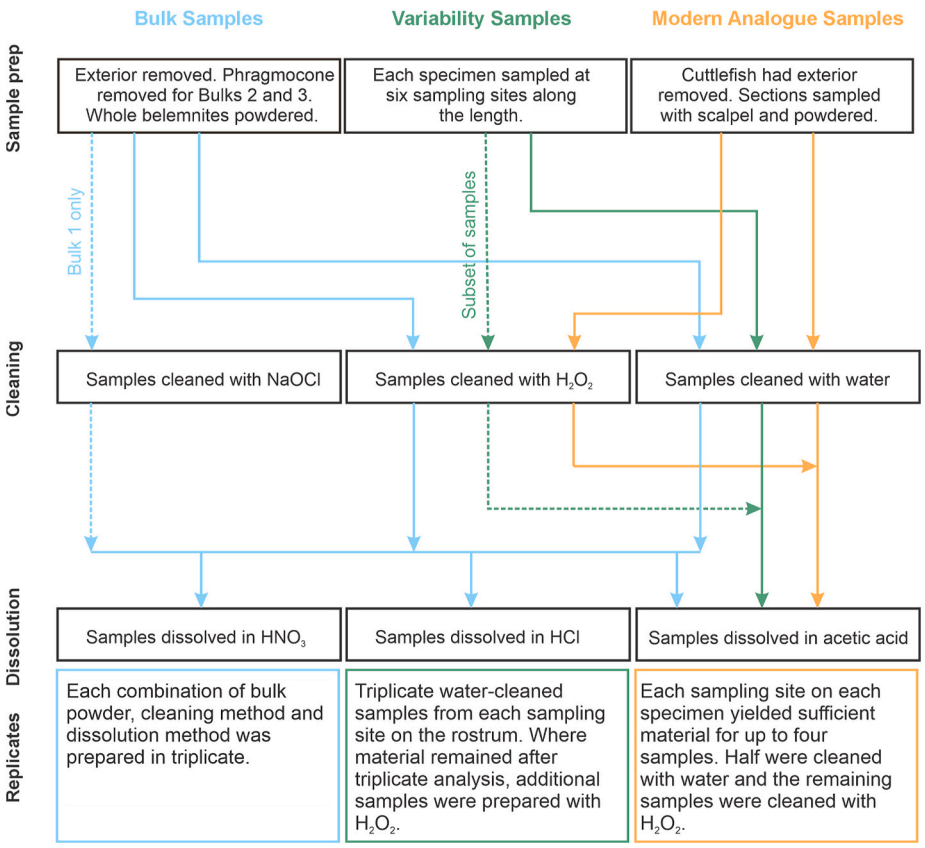

Fig. 2. Schematic for the method and variability tests applied to the Bulk Samples, Variability Samples and Modern Analogue Samples.

Table 1
Number of individuals, belemnite species, ammonite subzones and sampling locations for analysed samples from Penny Nab and Wine Haven, to investigate the variability of elemental ratios within and between specimens and stratigraphic levels.

Sampling Location	Ammonite zone	Ammonite subzone	Belemnite species	Number of individuals
Penny Nab, Staithes	*Amaltheus margaritatus*	*Amaltheus gibbosus*	*Bairstowius* sp.	1
			Parapassaloteuthis sp.	3
			Passaloteuthis pessula	2
Wine Haven, Robin Hood's Bay	*Uptonia jamesoni*	*Phricodoceras taylori*	*Passaloteuthis pessula*	1
			Nannobelus delicatus	3

2. Materials and Methods

2.1. Materials

2.1.1. Runswick Bay, Cowbar Nab and Ravenscar belemnites

Homogenous bulk powders were created from three large (~30 g) specimens of belemnite rostra collected from different stratigraphic levels on the Yorkshire coast, UK. Bulk 1 was created from the rostrum of a Lower Toarcian *Passaloteuthis bisulcata* belemnite, collected from the *Dactylioceras tenuicostatum* ammonite zone of the Grey Shale Member of the Whitby Mudstone Formation at Runswick Bay, North Yorkshire. Bulk 2 was created from a Lower Toarcian *Acrocoelites oxyconus* belemnite, collected from the *Dactylioceras commune* ammonite subzone of the Alum Shale Member, Whitby Mudstone Formation at Ravenscar, North Yorkshire. Bulk 3 was created from an Upper Pliensbachian *Passaloteuthis pessula* belemnite, collected from the *Amaltheus stokesi* ammonite subzone at the Cowbar Nab section in Staithes, North Yorkshire. For each Bulk Sample, the exterior of the belemnite was abrasively removed by wet grinding on a 75 μm full faced diamond grinding disc (8 in./200 mm) on an ATM Saphir 330 grinding machine. For Samples 2 and 3, the region around the alveolus (the socket which contains the phragmocone) was also removed, however for Bulk 1 the alveolar region, including the phragmocone, was included to create a Bulk Sample which can be considered 'spiked' with contaminant P derived from the infilling sediment (Fig. 1). For each Bulk Sample, the remaining material was crushed by hand in a steel pestle and mortar then sieved to ensure a grain size of < 180 μm.

2.1.2. Penny Nab and Wine Haven belemnites

Ten belemnites from two stratigraphic levels were analysed to investigate variability in elemental ratios within and between specimens and from different timepoints. Six specimens comprising three species were collected from the *Amaltheus gibbosus* ammonite subzone in the Upper Pliensbachian from the Penny Nab section at Staithes, North Yorkshire. Four specimens comprising two species were collected from the *Phricodoceras taylori* ammonite subzone, in the Lower Pliensbachian from the Wine Haven section at Robin Hood's Bay, North Yorkshire, which were approximately 5 million years older than the specimens from Penny Nab. In each case, specimens were collected from the same bedding plane to ensure that they were as contemporaneous as possible.

Each belemnite rostrum was sampled six times along its length using an abrasive diamond tipped Dremel drill bit (Fig. 2). Attempts were made to minimise inclusion of contaminants and diagenetically altered material by avoiding sampling the belemnite exterior, phragmocone and apical line, as well as targeting visually clear calcite (Ullmann et al., 2015) (Fig. 1). In most cases the resulting powder was extremely fine ($<$ 180 μm), but for some samples additional powdering was completed by

Table 2
Average limits of detection (LOD), limits of quantification (LOQ) and percentage uncertainties for ICP-OES runs for Mg, P, Fe, Mn, Sr and Ca values. LOD and LOQ values are calculated as 3 times and 10 times the standard deviation of six blank samples respectively. The percentage uncertainty is the calculated 95% confidence interval of 6 repeated quality control standard measurements.

	Mg	P	Fe	Mn	Sr	Ca
LOD / mg L^{-1}	3E-03	3E-03	2E-03	2E-04	2E-04	0.40
LOQ / mg L^{-1}	9E-03	1E-02	5E-03	7E-04	5E-04	1.35
% Uncertainty	1.4	1.8	2.1	1.0	0.9	1.5

hand using an agate pestle and mortar. Collectively, samples from Wine Haven and Penny Nab are referred to as Variability Samples (Table 1).

2.1.3. Modern Analogue Samples

Cuttlefish and Ram's Horn squid (*Spirula spirula*) are close living relatives of belemnites, and they both have hard internal carbonate shells, similar to a belemnite rostrum. Modern cuttlebones from *Sepia officinalis* and shells from the Ram's Horn squid (*Spirula spirula)* were analysed to provide further evidence of phosphate inclusion into cephalopod carbonate frameworks, and to evaluate variability within and between contemporaneous individuals. Unlike belemnite rostra, which are formed of calcite, cuttlebones and Ram's Horn squid shells are formed of aragonitic carbonate, which limits their utility as a modern analogue for belemnite rostra. Despite this, their similarity to belemnite rostra as a hard internal carbonate shell makes them the best placed materials for modern comparison. Two *Sepia officinalis* cuttlebones were donated by the Marine Biological Association, UK (MBA). The cuttlebones were from two immature individuals (one aged 8 months and one aged 13 months) which were raised together in natural seawater from the Plymouth Sound, UK, and thus differences in environmental conditions are expected to be minimal. The cuttlebones were taken from deceased individuals and frozen by the MBA. The specimens were then defrosted and dried in an oven at 70 °C for 72 h. The exterior of the cuttlebone was removed using a scalpel to reveal the porous interior. Samples of the interior were taken using a scalpel across the width of the cuttlebones but avoiding the apical line. Each sample incorporated multiple growth bands providing an averaged signal. Samples were then powdered using an agate pestle and mortar.

Five *Spirula spirula* samples were collected from the Marokopa coastline of North Island, New Zealand, including whole shells and fragments. *S. spirula* have an aragonitic, open planispiral, internal chambered shell. For each of the five *S. spirula* specimens, sections were removed beginning from the outer end of the open planispiral shell. In the case of fragments, a determination was made as to which end was most likely the outer end. Each section was powdered using an agate pestle and mortar.

2.2. Method

2.2.1. Experimental Design

A range of three cleaning methods (water, hydrogen peroxide (H_2O_2) and sodium hypochlorite (NaOCl)) were trialled, alongside three dissolution methods (acetic acid, nitric acid and hydrochloric acid). These methods were selected for this investigation as they are commonly used methods for analysing palaeontological samples of biogenic carbonates (Shotyk et al., 1995; Kumarsingh et al., 1998; Penkman et al., 2008; Zhang et al., 2020; Ingalls et al., 2020; Dodd et al., 2021; Riding, 2021 and references therein; Ingalls et al., 2022).

For each specimen type, 10–20 mg of powdered carbonate were weighed out. Different preparative methods were trialled on the Bulk Samples and a subset of the Variability Samples. For Bulk 1, three different sample cleaning methods were trialled: ultrapure water, H_2O_2 and NaOCl (Fig. 2). NaOCl was only trialled on samples of Bulk 1 and was found to give similar results to H_2O_2 cleaning. Because of this similarity and because it's more difficult to purchase at the higher purity grades required for trace element analyses, it was not used for subsequent tests. For Bulk 2 and 3 Samples, two different cleaning methods were used: ultrapure water and H_2O_2 (Fig. 2). For each cleaning method, three different dissolution methods were tested: acetic acid, hydrochloric acid, and nitric acid (Fig. 2). For each combination of sample, cleaning method and dissolution method, Bulk Samples were analysed in triplicate (Fig. 2).

Variability Samples were cleaned with ultra-pure water and dissolved in acetic acid then analysed in triplicate, where there was sufficient material from each location on the rostra (Fig. 2). Where there was remaining material after the triplicate analysis, additional 10–20 mg samples were cleaned with hydrogen peroxide to investigate the impact of oxidative cleaning on intra- and inter-specimen differences in elemental ratios (Fig. 2). All Variability Samples were dissolved in acetic acid (Fig. 2).

For Modern Analogue specimens, each site on the specimens yielded sufficient material for up to four samples (10–20 mg aliquots). Half of the samples from each site on the specimens were cleaned with ultrapure water, then the remaining sample were cleaned with H_2O_2 (Fig. 2). All Modern Samples were dissolved in acetic acid (Fig. 2).

For each combination of cleaning method and dissolution method, samples of reference material JCp-1 were analysed. JCp-1 is a reference material of coral *Porites* sp. prepared by the Geological Survey of Japan (Okai et al., 2007). During the analysis of Variability and Modern Samples, additional analyses of Bulk Samples 2 and 3 were conducted to provide an additional standard for comparison between runs and methods.

2.2.2. Cleaning methods

For samples cleaned with water, ultrapure water (2 mL) was added to the 10–20 mg samples and reference material and sonicated for 15 mins. Samples were centrifuged and the supernatant removed. This was repeated a further four times to ensure removal of soluble contaminants, then samples were air dried for at least 48 h.

For samples cleaned with hydrogen peroxide, H_2O_2 (2 mL, 9.8 M, Analytical Reagent Grade, Fisher Chemicals) was added to each sample and samples were agitated for 72 h on a shaker table. Samples were centrifuged, and the supernatant pipetted away. The samples were then rinsed five times with ultrapure water using the method above (5× addition/removal of 2 mL ultrapure water), given an additional rinse with methanol (2 mL, Analytical Reagent Grade, Fisher Chemicals) using the same method, then air dried for at least 48 h.

For samples cleaned with sodium hypochlorite, NaOCl (2 mL, 12%, GPR RECTAPUR, VWR Chemicals) was added to each sample, and samples were agitated for 48 h. Samples were centrifuged, and the supernatant liquid removed. Samples were rinsed with water and methanol using the same method as samples cleaned with H_2O_2, then air dried for at least 48 h.

2.2.3. Acidification

Cleaned and dried powders were accurately weighed to the nearest microgram into approximately 5–6 mg samples. Bulk Samples were dissolved in either HCl (1 mL, 1 M, Analytical Reagent Grade, VWR Chemical), nitric acid (1 mL, 0.7 M, Trace Metal Grade, Fisher Chemical) or acetic acid (1 mL, 5 M, Trace Metal Grade, Fisher Chemical). The acidified samples were agitated for 15 mins then centrifuged. Variability and Modern Analogue samples were prepared via dissolution in acetic acid (1 mL, 5 M, Trace Metal Grade, Fisher Chemical).

Additional samples for Bulks 1, 2 and 3 were also prepared using the H_2O_2 cleaning method. The cleaned and dried powders were accurately weighed into approximately 2 mg and 10 mg samples and dissolved in acetic acid (1 mL, 5 M, Trace Metal Grade, Fisher Chemical), using the same procedure as the 5–6 mg samples, to investigate the impact of sample size.

For all acidified samples, 0.3 mL of the supernatant solution was

Table 3
R^2 values for the correlation between P/Ca and either Fe/Ca or Mn/Ca in samples from Penny Nab and Wine Haven which either had no screening criteria applied, or which had been screened to remove samples with a high level of diagenetic alteration (using Mn/Ca < 300 μmol mol^{-1} and Fe/Ca < 600 μmol mol^{-1}).

Data	Penny Nab section		Wine Haven section	
	Fe/Ca	Mn/Ca	Fe/Ca	Mn/Ca
All data	0.10	0.19	< 0.01	0.04
Screened samples	< 0.01	0.04	0.44	0.42

added to a matrix solution (HNO_3, 8.6 mL, 0.1 M, Trace Metal Grade, Fisher Chemical) and an internal standard (Y, 0.1 mL, 10 ppb, 0.1 M HNO_3 matrix). Samples and blanks were analysed on a Thermo Scientific iCAPQc ICP-MS for Li, B, Na, Al, K, Ba and U concentrations, and a Thermo Scientific iCAP 7400 Radial ICP-OES for Ca, Mg, P, Fe, Mn, S and Sr concentrations in the Cohen Laboratory at the University of Leeds. For these analyses the Limits of Detection (LOD), Limits of Quantificatio (LOQ) and measurement % uncertainties based on replicate measurements are displayed in Table 2.

Some previous studies of phosphorus have used repeated high salinity washes to remove adsorbed phosphate from any remaining material not dissolved in acid, such as the $MgCl_2$ washes used in the SEDEX phosphorus phase partitioning method (Ruttenberg, 1992; Thompson et al., 2019) and the NaCl washes used in the CAP analysis of bulk rock samples (Dodd et al., 2021, 2023). Belemnite samples are almost entirely comprised of soluble calcite, compared to a higher percentage of insoluble material in bulk rock samples. For most samples there was no visible material left after dissolution, and thus high salinity washes were not necessary.

2.2.4. Phosphorus phase Partitioning

A sequential extraction protocol was applied to the Bulk Sample powders to identify and quantify the pools of P present in the belemnite specimens, using the method described in Thompson et al. (2019). This method targets five different pools of P: ferric (oxyhydr)oxide-bound P (P_{Fe}); magnetite-bound P (P_{Mag}); authigenic P, dominantly comprising carbonate fluorapatite, biogenic apatite and CAP (P_{Aut}); organic bound P (P_{Org}); and crystalline P, which dominantly comprises detrital apatite, but can also include a proportion of recrystallized carbonate fluorapatite (P_{Cryst}) (Thompson et al., 2019).

During the first phase of the extraction, P_{Fe} was dissolved in a solution of sodium citrate/sodium dithionate/sodium bicarbonate (room temperature, 8 h). The P_{Aut} pool was then extracted from the residue using a sodium acetate solution (room temperature, 6 h), and then P_{Cryst} was extracted in HCl (1 M, room temperature, 16 h). P_{Mag} was extracted in an ammonium oxalate solution (room temperature, 6 h) before P_{Org} was measured. To determine P_{Org}, the remaining sample was dried (100 °C, overnight) then ashed in a furnace (550 °C, 2 h). Samples were then extracted in HCl (1 M, room temperature, 16 h) to determine P_{Org}. The resultant solutions from all extractions were analysed to determine the concentration of P, predominantly utilising a spectrophotometric method at 880 nm (Strickland and Parsons, 1968) (particularly for $MgCl_2$ washes of samples to remove readsorbed P after extraction), with P_{Fe} extractions analysed by ICP-OES (as these extractions are likely to contain chemicals which interfere with colour development in the molybdate blue method) (Strickland and Parsons, 1968; Thompson et al., 2019).

2.2.5. Pyrite analysis

Pyrite concentrations were determined as oxidative cleaning methods are likely to oxidise pyrite to form iron oxyhydroxide minerals, which phosphate is easily adsorbed onto, potentially affecting the concentrations of phosphate released or retained by the powder. The amount of iron in pyrite (Fe_{Py}) in the Bulk Sample powders was measured following a chromous chloride distillation protocol (Canfield et al., 1986) and the concentration of pyrite was calculated from the weight of precipitated silver sulfide produced by the extraction.

2.2.6. Statistical tests

Where comparisons were made between samples or preparative methods, an independent two-sample *t*-test assuming unequal variance was conducted with a 95% confidence interval to determine if any differences were statistically significant. In some cases, as multiple *t*-tests have been performed on the same datasets, it is possible that statistically significant results with *P* values close to the threshold (0.05) are false positive results, and there are no statistically significant differences between the datasets. For each *t*-test conducted, the mean and standard deviation of each dataset compared are given, along with the *P* value, *t* value and degrees of freedom.

3. Results and discussion

3.1. Screening for diagenetic alteration

Data for the Variability Samples from Penny Nab and Wine Haven were screened to minimise the impact of diagenetic alteration on the investigations of variability within the belemnites. Bulk Samples were not screened as they are expected to contain diagenetic or contaminant material. For the Penny Nab and Wine Haven samples, Mn/Ca and Fe/Ca ratios were used to screen samples for an unacceptably high degree of diagenetic alteration. Limits of Fe/Ca < 600 μmol mol^{-1} (Fe < 828 ppm) and Mn/Ca < 300 μmol mol^{-1} (Mn < 411 ppm) were applied, resulting in the exclusion of roughly 10% of samples. These limits were selected based on correlations between P/Ca, Mn/Ca and Fe/Ca to minimise the impact on P/Ca ratios by diagenetic and contaminant effects, while ensuring samples where P/Ca values were not impacted by these effects were not unnecessarily excluded, as described below. The resultant limits are relatively high compared to limits set in some previous studies of belemnite calcite, although limits vary from study to study (e.g., Ullmann and Korte, 2015; Stevens et al., 2022).

In both screened and unscreened data, there was a correlation between Fe/Ca and Mn/Ca, particularly in samples from the Wine Haven section, indicating that screened samples have still undergone a degree of diagenetic alteration, the impact of which was assessed using P/Ca ratios. To determine the impact on P/Ca ratios of the degree of diagenetic alteration in the screened samples, P/Ca ratios were compared to Mn/Ca and Fe/Ca ratios (Table 3). In unscreened samples, there was no statistically significant correlation between P/Ca and either Fe/Ca or Mn/Ca ($R^2 \leq 0.19$) (Table 3). After screening conditions had been applied, the R^2 value for correlation between P/Ca and both Fe/Ca and Mn/Ca decreased for samples from Penny Nab ($R^2 < 0.04$) but increased for samples from Wine Haven, with R^2 values indicating a weak correlation between P/Ca and both Mn/Ca ($R^2 = 0.42$, slope = 0.96) and Fe/Ca ($R^2 = 0.44$, slope = 0.37) (Table 3). This suggests that the screening criteria were successful in eliminating samples where diagenesis has influenced P/Ca ratios in material from Penny Nab. For screened samples from Wine Haven, it should be considered that P/Ca values may still be impacted by diagenesis.

3.2. Phosphorus form and content of belemnites

For each Bulk Sample, the different pools of P were summed together to estimate the total P content. Based on the phosphorus phase partitioning results, Bulk 1, 2 and 3 samples contained 730 ppm (0.073 wt%), 332 ppm (0.033 wt%) and 372 ppm (0.037 wt%), respectively (Fig. 3). This is at the lower end of previous measurements of belemnite phosphorus, but within the established range of 100–1800 ppm. (0.01–0.18 wt%, Longinelli et al., 2002), potentially suggesting variability between specimens. For all samples, the dominant pool was P_{Aut}, which accounted for 90–95% of the P (Fig. 3). Of the remaining pools, the

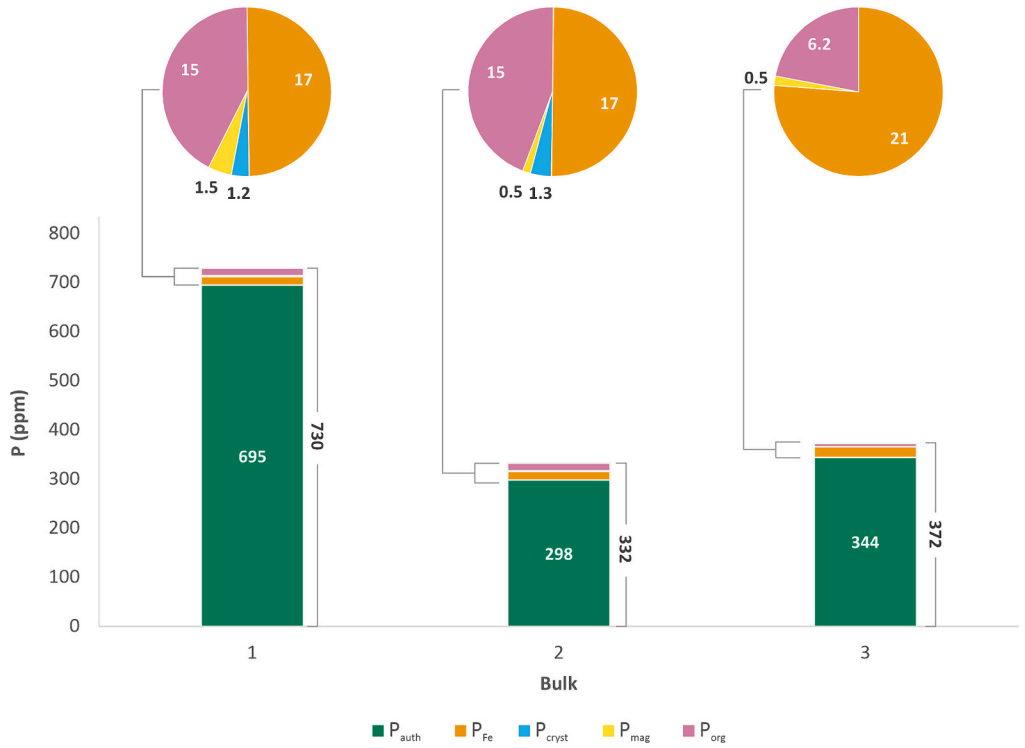

Fig. 3. Different pools of P in Bulk Samples 1, 2 and 3. Pie charts show only the proportions of the non-authigenic pools of P measured. For each sample and pool of P, the amount of P measured in ppm is indicated.

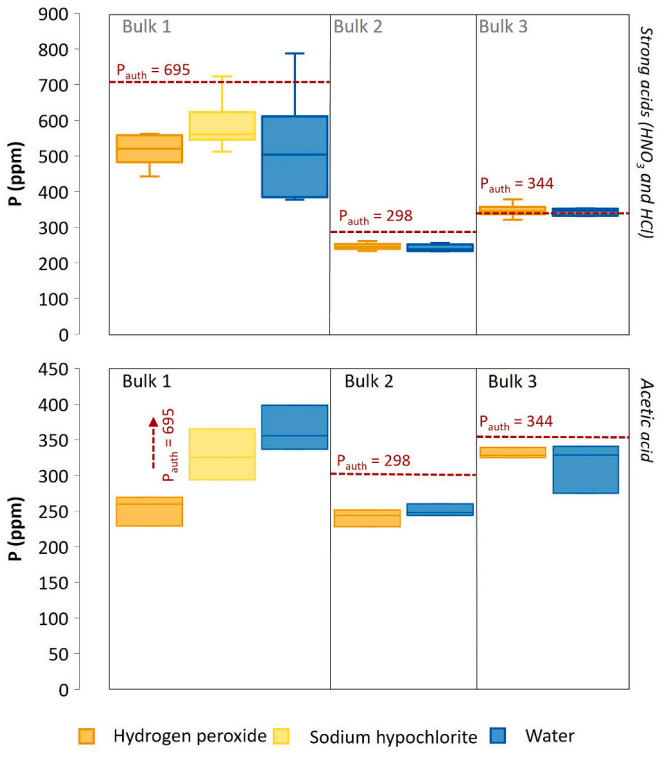

Fig. 4. Measured P concentrations for different cleaning methods (hydrogen peroxide, sodium hypochlorite or water) trialled on Bulk Samples 1, 2 and 3, which were then dissolved in strong acid (combined HCl and HNO_3 data) or acetic acid. Also marked is the measured P_{Aut} content of the samples based on phosphorus phase partitioning data. For each strong acid boxplot, the number of samples (n) is 6, while for acetic acid $n = 3$.

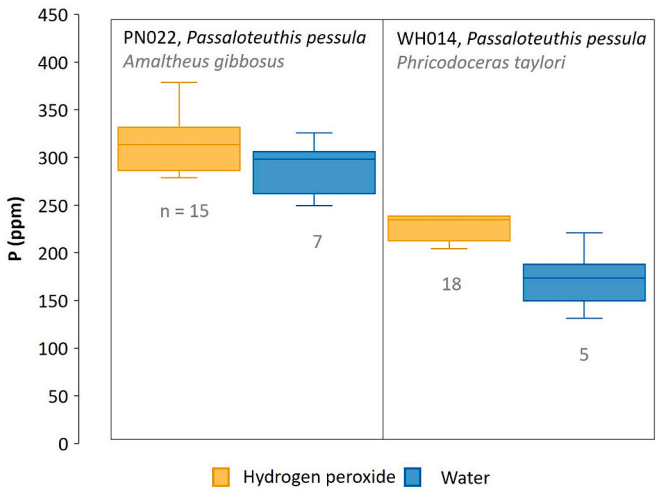

Fig. 5. P concentrations in two belemnites, one from the *Amaltheus gibbosus* subzone in the Penny Nab section and one from the *Phricodoceras taylori* ammonite subzone in the Wine Haven section, for samples cleaned with hydrogen peroxide (orange) and samples cleaned with water (blue). For each boxplot, the number of samples (n) is indicated.

samples contained 17–21 ppm P_{Fe}, 6–15 ppm P_{Org}, and only trace quantities of P_{Cryst} crystalline and P_{Mag} (Fig. 3).

Bulk 1, which is the 'spiked' bulk sample, had the largest total P content measured by the sequential extraction method, with over double the quantity of P present in Bulks 2 and 3 (Fig. 3). Despite this, the concentrations of P_{Fe}, P_{Cryst}, P_{Mag} and P_{Org} in Bulk 1 were similar to those in Bulks 2 and 3, with the higher total P content in Bulk 1 being due to much higher concentrations of P_{Aut} (Fig. 3). The Bulk 1 sample contained 0.30 wt% pyrite, while Fe_{Py} in Bulks 2 and 3 was below detection (< 0.001 wt%).

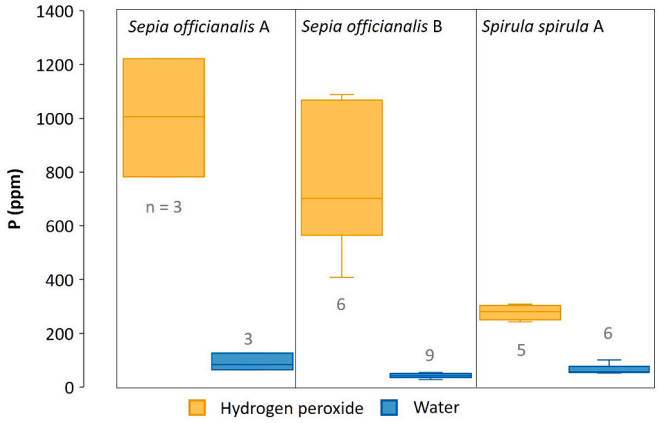

Fig. 6. P concentrations in three Modern Analogue specimens, two *Sepia officinalis* and one *Spirula spirula*, for samples cleaned with hydrogen peroxide (orange) and samples cleaned with water (blue); all samples were dissolved in acetic acid. For each boxplot, the number of samples (*n*) is indicated.

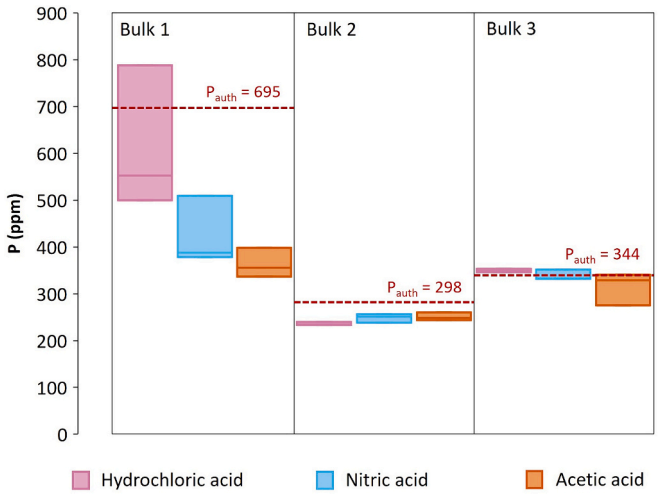

Fig. 7. Measured P concentrations for different dissolution methods (hydrochloric acid, nitric acid and acetic acid) trialled on samples of Bulk Sample powders 1, 2 and 3 which had been cleaned with water. Also marked is the measured P_{Aut} content of the samples based on phosphorus phase partitioning data. For each boxplot the number of samples (*n*) is 3.

3.3. Impact of cleaning method

Oxidative (H_2O_2 and NaOCl) and non-oxidative (water) cleaning methods were applied to the Bulk Samples. For Bulks 2 and 3 there was no statistically significant difference in P concentrations between samples cleaned with water and samples cleaned with hydrogen peroxide (Fig. 4). For Bulk 1, there were no statistically significant differences between samples cleaned with NaOCl and samples cleaned with hydrogen peroxide for any of the dissolution methods. For Bulk 1 samples dissolved in strong acids, there were no statistically significant differences between samples cleaned with oxidative cleaning methods and water (Fig. 4). For Bulk 1 samples dissolved in acetic acid, samples cleaned with water ($\bar{x}$ = 364 ppm, s = 32 ppm, n = 3) had 25% higher average P concentrations than samples cleaned with oxidative cleaning methods ($\bar{x}$ = 290 ppm, s = 49 ppm, n = 6); t(6) = 2.45, p = 0.035.

The majority of the Variability Samples were prepared with the water cleaning method, whilst a smaller subset was cleaned with hydrogen peroxide. Data from two specimens, one from the Penny Nab section and one from the Wine Haven section, with sufficient ($n \geq 3$) oxidatively cleaned samples are compared in Fig. 5. In the specimen from Penny Nab, there was no significant difference between samples cleaned with water ($\bar{x}$ = 286 ppm, s = 25 ppm, n = 15) and samples cleaned with hydrogen peroxide ($\bar{x}$ = 308 ppm, s = 22 ppm, n = 7); t(13) = 2.16, p = 0.059 (Fig. 5). However, in the specimen from Wine Haven, samples cleaned with hydrogen peroxide ($\bar{x}$ = 227 ppm, s = 23 ppm, n = 5) had 33% higher P concentrations than samples cleaned with water ($\bar{x}$ = 171 ppm, s = 23 ppm, n = 1); t(10) = 2.23, p = 7.11E-05 (Fig. 5).

Oxidative and non-oxidative cleaning methods were also applied to *Sepia officinalis* cuttlebone and *Spirula spirula* shell samples. In Modern Analogue Samples there were statistically significant differences between samples cleaned with hydrogen peroxide and samples cleaned with water for all specimens (Fig. 6). For Cuttlefish A and B and Spirula A, samples cleaned with hydrogen peroxide had P concentrations 10 times, 17 times and 3.3 times higher than samples cleaned with water, respectively (Fig. 6).

Where there are differences in P concentrations between the different cleaning methods, the data is consistent with the release and retention of non-CAP contaminant phases during oxidative cleaning, most likely from organic matter. In Modern Analogue Samples, which had the most dramatic difference between cleaning methods, cleaning with hydrogen peroxide gave substantially higher measured P concentrations, probably because these samples had high concentrations of organic matter compared to the ancient belemnite material. Cuttlebones are made of aragonite with high organic matter content (3–4.5 wt%), comprised of mostly chitinous protein complexes (Birchall and Thomas, 1983; Checa et al., 2015). *S. spirula* are also formed of aragonite, and contain organic matter that is mostly conchiolin (Mutvei, 1964; Florek et al., 2009; Hoffmann et al., 2018; Checa et al., 2022). Organic matter contents of *S. spirula* have not been reported but are likely similar to the chambered nautilus (*Nautilus pompilius*), which has an organic matter content of around 1 wt% (Petrochenkov et al., 2018). In oxidatively cleaned Modern Analogue Samples, hydrogen peroxide would have broken down the organic matter, releasing the organic-bound P into solution. This free P likely readsorbed onto the remaining powder, and was not removed by subsequent water washes, increasing the measured P concentration (Ruttenberg, 1992; Thompson et al., 2019). In water-cleaned samples, the organic matter was not oxidised, and the organic P remained bound in the acid-resistant protein complexes. As a weak acid was used, the amount of organic matter dissolved was minimal, giving lower overall P concentrations.

For the Variability Samples, some specimens gave higher measured P values for samples cleaned with hydrogen peroxide than for samples cleaned with water. This is likely also due to the oxidation of contaminant P phases by oxidative cleaning, in the same way as observed for the Modern Analogue Samples. In belemnite samples, the dominant form of contamination is likely to be apatite, potentially from sediment contamination or diagenetic alteration, with small quantities of acid-resistive organic matter (Sælen, 1989). As this was only observed for some specimens, the impact of this effect probably varies with the degree of alteration and contamination of the sample, which differs between specimens.

In Bulk Samples, there was no difference between cleaning methods for Bulks 2 and 3. This may be due to lower levels of alteration and contamination in these samples, resulting in low levels of contaminant release with oxidative cleaning methods. Alternatively, the oxidising process may be creating additional iron oxyhydroxides from pyrite oxidation, of which Bulk 1 had a substantially higher concentration. These iron oxyhydroxides would retain released contaminant P by adsorption during subsequent washes, before being dissolved and increasing the measured P of Bulk 1. For Bulk 1 samples dissolved in acetic acid, samples cleaned with water gave a 25% higher average P concentration than samples cleaned with hydrogen peroxide (Fig. 7). This was different to the trends observed in the Variability Samples, where samples dissolved in acetic acid and cleaned with hydrogen peroxide gave higher P measurements.

Table 4
CV_{rep} values in % for each combination of Bulk Sample powder, cleaning method and dissolution method. In each case n is 2–3.

Bulk Sample	NaOCl			H_2O_2			Water		
	HCl	HNO_3	Acetic	HCl	HNO_3	Acetic	HCl	HNO_3	Acetic
1	7.1	0.9	10.9	9.3	5.5	8.3	7.1	17.2	8.7
2				4.0	3.6	5.0	1.5	3.7	3.4
3				0.3	8.1	2.2	1.0	3.3	11.1

Table 5
CV_{ind} values in % for each combination of Modern Analogue Sample and cleaning method. In each case samples were dissolved in acetic acid. For each combination n is 2–3.

Organism	Specimen Identifier	H_2O_2	Water
Sepia officinalis	A	22	34
	B	35	21
Spirula spirula	A	10	29
	B	–	10
	C	5.8	25
	D	–	–
	E	11	–

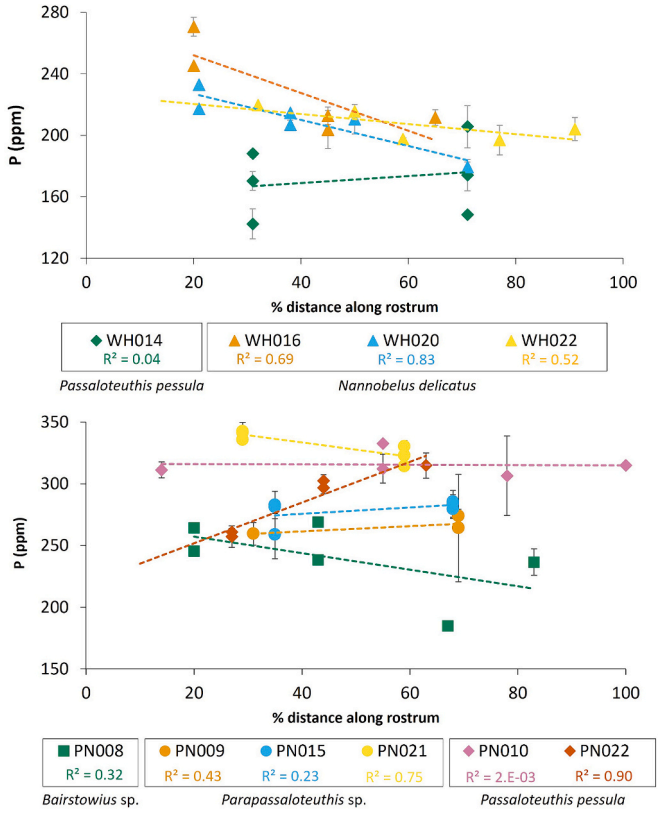

Fig. 8. P concentration versus sampling distance along the belemnite rostrum from the apex as a % of the belemnite length. 0% represents the apex and 100% represents the base of the rostrum. Data is for ten belemnite specimens, four from Wine Haven and six from Penny Nab. All samples were prepared with water cleaning and acetic acid dissolution. Three specimens are *Passaloteuthis pessula* (diamonds), three are *Nannobelus delicatus* (triangles), three are *Parapassaloteuthis* sp. (circles) and one is *Bairstowius* sp. (square). For each individual specimen the linear best fit (dashed line) is shown in a colour corresponding to the datapoints, along with an R^2 value for the line. Error bars (pale grey) representing one standard deviation calculated from 2 to 3 replicates are included where there were sufficient replicates.

3.4. Impact of dissolution method

For all samples and dissolution methods, the average amount of P measured was lower than or similar to the measured amount of P_{Aut}, as determined by phosphorus phase partitioning (Fig. 7). P_{Aut} in the phosphorus phase partitioning method represents carbonate-associated phosphorus, as well as the more soluble carbonate fluorapatite and biogenic apatite, but excludes the P_{Det} fraction (which contains less soluble igneous and metamorphic apatite). The similar magnitudes of the measured dissolved P and the phosphorus phase partitioning P_{Aut} fraction suggests that all the acids somewhat selectively dissolve the authigenic pool of P, rather than the non-authigenic forms, although these are only present at very low concentrations.

For Bulks 2 and 3, there were no statistically significant differences between the dissolution methods (Fig. 7). For Bulk 1, there was no statistically significant difference between the strong acids trialled (nitric and hydrochloric), but both gave consistently higher P concentrations than samples dissolved in acetic acid by a factor of around 1.5 (Fig. 7). The amount of P measured in Bulk 1 samples dissolved in acetic acid ($\bar{x} = 364$ ppm, $s = 32$ ppm, $n = 3$ [water cleaned]) was similar to the quantities measured in Bulk 2 ($\bar{x} = 251$ ppm, $s = 9$ ppm, $n = 3$ [acetic dissolved, water cleaned]) and Bulk 3 ($\bar{x} = 315$ ppm, $s = 35$ ppm, $n = 3$ [acetic dissolved, water cleaned]), which were not spiked in the same way as Bulk 1 (Fig. 7). Despite containing infilled phragmocone material from the alveolar region, the phosphorus phase partitioning data indicates that Bulk 1 does not have significantly higher concentrations of organic and apatite bound P compared to Bulks 2 and 3. This suggests that the difference in measured P is due to higher concentrations of contaminant carbonate fluorapatite. Previous studies have demonstrated that weak acids such as acetic acid are less able to target and dissolve minerals such as carbonate fluorapatite in belemnites, and other forms of contaminant P compared to strong acids (Ingalls et al., 2020, 2022). Weak acids have also been shown to be less able to dissolve organic-bound P compared to strong acids, particularly as organic matter in belemnites has been shown to have high acid resistivity – although as all Bulk Samples seem to have similar concentrations of organic P this is less likely to be a factor (Sælen, 1989). The lower measured P concentrations in Bulk 1 samples dissolved in acetic acid confirms that acetic acid is more selective, preferentially dissolving the more acid-soluble CAP rather than contaminant phases such as carbonate fluorapatite. This demonstrates that, while there is no difference between dissolution methods in uncontaminated samples, in samples with diagenetic alteration or contamination by sediment, acetic acid can effectively discriminate between the different forms of P in the sample. This makes it an ideal acid for investigating carbonate associated P in samples with high levels of contaminant P phases, such as diagenetically altered biogenic samples and bulk rock carbonate samples. We suggest that the precise concentration of acetic acid is not critical, as long as there is excess free H^+ compared to the sample carbonate content, it is likely that more dilute acetic acid will also effectively dissolve CAP and discriminate against contaminant phases.

3.5. Reproducibility of results

Coefficients of variation (CV) were used to investigate the variability and reproducibility of the results by scaling the standard deviations to the mean values. These were calculated as:

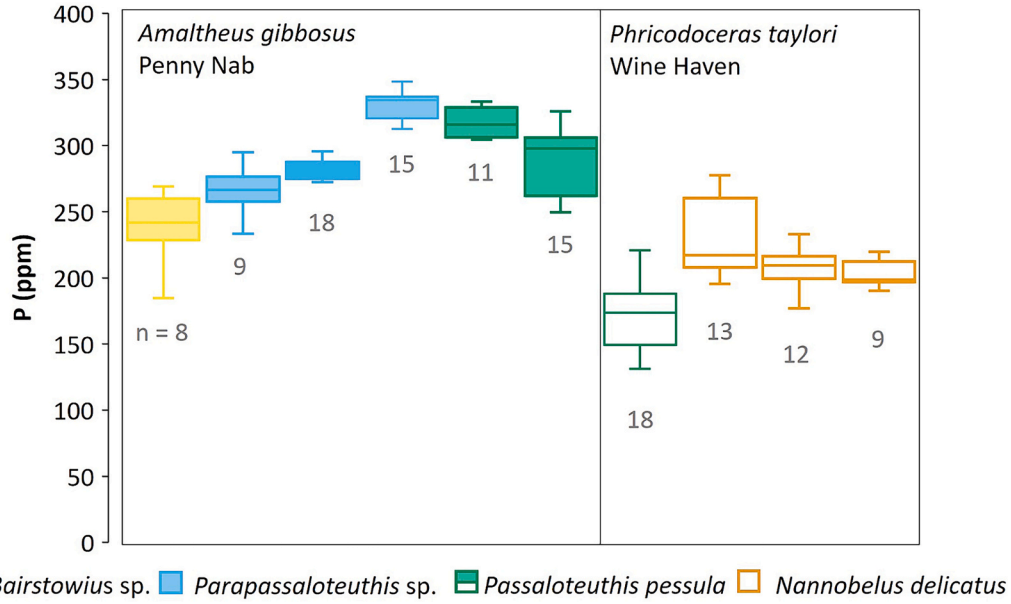

Fig. 9 - P. Concentrations for 6 individuals from the Penny Nab section (*Amaltheus gibbosus* subzone) [filled boxes] and 4 individuals from the Wine Haven section (*Phricodoceras taylori* subzone) [unfilled boxes]. Each colour represents a belemnite species, and each box corresponds to an individual specimen. Each belemnite was sampled at six points along the rostra. Samples were cleaned with water, disolved in acetic acid and, where there was sufficient material, samples were prepared and analysed in triplicate. The total number of samples (*n*) for each specimen after diagenetic screening criteria was applied is displayed.

$$CV\ (\%) = 100 \times \frac{s}{\bar{x}}$$

where s is the standard deviation and $\bar{x}$ is the mean of at least three replicate analyses. CVs were calculated for replicate samples (Cv_{rep}), individual whole specimens (CV_{ind}), and stratigraphic levels (CV_{lvl}).

For Bulk Samples, each combination of sample material, cleaning method and dissolution method gave a high degree of reproducibility. For each combination, CV_{rep} values ranged from 0.3% to 17.2%, with most CV_{rep} values < 10% indicating reproducible results (Table 4). The only combination where the CV_{rep} values exceeded 10% was for Bulk 1 samples cleaned with water and dissolved in nitric acid, which gave a CVs of 17.2% (Table 4). Bulk 1 samples gave generally higher CV_{rep} values than Bulk 2 or 3 samples for equivalent preparative methods, which are ascribed to variable dissolution or amounts of contaminant phases between the replicates (Table 4). Oxidatively cleaned and water cleaned samples gave similar CV values, indicating that cleaning methods did not impact the reproducibility of the results.

For Variability Samples, CV_{rep} values were calculated for replicate samples cleaned with water, but not for samples cleaned with hydrogen peroxide due to insufficient numbers of replicate analyses. Water-cleaned Variability Samples had low average CV_{rep} values ($\bar{x}_{CV}$ = 4.4%; s_{CV} = 6.6; n = 32), similar to Bulk Samples 2 and 3. CV_{ind} values were also calculated for each specimen, using replicate samples from multiple sampling points along the length of the rostra, for both water-cleaned and hydrogen peroxide-cleaned samples. The average CV_{ind} values for the specimens were higher than for the replicate samples (water: $\bar{x}_{CV}$ = 9.8%; s_{CV} = 6.2%; n = 10; hydrogen peroxide: $\bar{x}_{CV}$ = 5.7%; s_{CV} = 2.5; n = 4), but still relatively low, suggesting that single analyses may be sufficient to characterise individuals.

For Modern Analogue Samples, there were insufficient replicate samples to calculate CVs. Instead, where there were sufficient (≥ 3) samples from the same specimen, CV_{ind} values were calculated using all samples analysed across that specimen. CV_{ind} values for both *Sepia officinalis* and two *Spirula spirula* were calculated, for both water-cleaned and hydrogen peroxide-cleaned material, and CV_{ind} values for water-cleaned samples were calculated for two further *S. Spirula* specimens. For both *S. officinalis* cuttlebones, cleaning with either method gave high CV_{ind} values (Table 5). For the *S. spirula* shells, cleaning with hydrogen peroxide gave lower CV_{ind} values than cleaning with water (Table 5). For *S. spirula*, the CVs were mostly low (< 20%) indicating a high reproducibility for P concentrations within individual shells, with the exception of shells A and C cleaned with water (Table 5). This is likely due to *Spirula spirula* containing lower concentrations of organic-bound P compared to *Sepia officinalis*, and therefore having less variability in organic contamination. Variable concentrations of organic matter oxidation and dissolution is unlikely to be a factor in belemnite analysis, as belemnite calcite contains only a small fraction of organic matter, with the phosphorus phase partitioning results indicating that organic-bound P accounts for only around 1.5–5% of belemnite P.

3.6. Variability within and between belemnite specimens

Our data suggest that P measurements are reproducible within belemnite specimens, but further investigations were conducted to examine spatial variability in P contents in belemnite rostra, as well as potential taxonomic effects. The percentage distance along each rostrum was compared to the average P concentration for each sample site from each belemnite. For most individuals, P concentrations were found to be slightly elevated towards the apex of the rostrum (Fig. 8). This was seen in all *Nannobelus delicatus* belemnites from Wine Haven (R^2 for each individual ≥ 0.52), but not in the *Passaloteuthis pessula* belemnite from the same section, although this may be due to poor distribution of the sampling sites for this specimen (Fig. 8). For belemnites from the Penny Nab section, similar trends were seen with samples from the *Bairstowius* sp. belemnite and one of the *Parapassaloteuthis* sp. belemnites, with R^2 values of 0.32 and 0.75, respectively. One of the *Passaloteuthis pessula* individuals demonstrated the opposite trend, with a positive correlation between distance from the apex and P concentration (R^2 = 0.90). A slight positive correlation between distance from the apex and P concentration was also observed in the remaining two *Parapassaloteuthis* sp. specimens (R^2 = 0.43 and 0.23), though the magnitude of the variation was very slight (slope = 0.21 and 0.26 respectively). For the remaining *Passaloteuthis pessula* specimen, there was no coherent trend between P concentration and sample position on the rostra. The differences observed in the relationship between sampling location and P concentration between specimens may be related to the distribution of

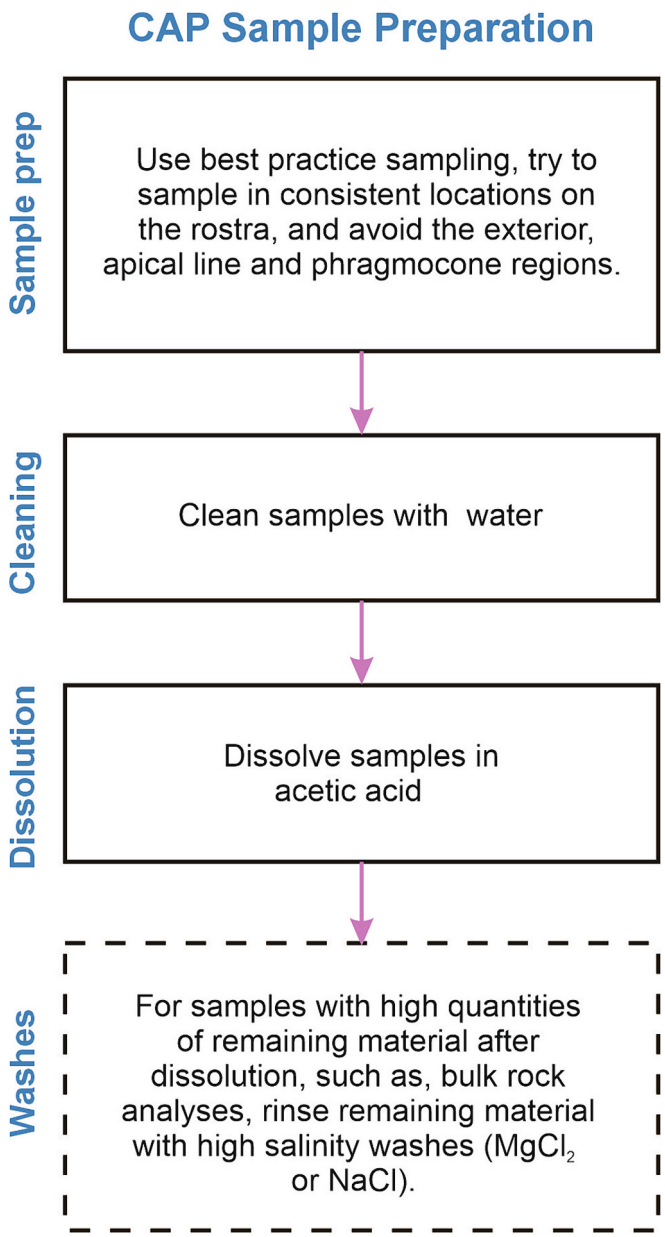

Fig. 10. Flow diagram of the ultimate preferred analysis scheme for carbonate-associated phosphorus measurements, based on the tests from this study.

sampling sites, or the relationship may be species-specific. For example, it may be that variation in P concentrations along the length of the belemnite is species-specific, and the trend is more pronounced in some species, such as *Nannobelus delicatus*. Alternatively, the control may be related to variations in P content on the short axis, perpendicular to the apical line. This positioning may be critical in relation to P uptake and may vary with ontogenetic differences between species. Further research on the three-dimensional distribution of P and other elements would be advantageous to explain the observed trends in P concentration along the rostra.

These trends may also be linked to growth rate or biological age of the individual, as sampling through the growth rings at different points along the rostrum could constitute a primitive growth study. Near the apex of the belemnite rostrum the growth bands in the calcite are likely closer together, and the record of the early part of the belemnite's life gets progressively lost towards the apex as growth bands converge, compared to sampling sites further away from the apex. It has been shown that other element concentrations such as Mg and Sr are affected by growth rate and growth bands, so it is likely that P also varies in this way (Ullmann and Pogge Von Strandmann, 2017). To minimise the impact of variable P concentrations within belemnite rostra, where possible sampling should be conducted at similar sampling sites in each specimen, as well as avoiding the apical line and the phragmocone. It should be noted that even taking into consideration the variations within individual belemnites, the CV values were low enough to indicate high reproducibility – so fragmented and broken belemnite material could still be used even when sampling site on the rostrum cannot be controlled.

The variance in P concentration between belemnites is larger than the variance within each specimen (Fig. 9). For each stratigraphic level studied, a CV was calculated (CV_{lvl}) to give an indication of the variability in P concentrations between belemnites from the same stratigraphic level. For samples from the *Amaltheus gibbosus* ammonite subzone, the CV_{lvl} value was 11.5% ($n = 76$); for samples from the *Phricodoceras taylori* ammonite subzone, the CV_{lvl} value was 18.7% ($n = 52$). These CVs are higher than the CV_{ind} values for variability in P concentrations within individual belemnites but are still low and indicate high reproducibility in P values within a stratigraphic level.

The impact of taxonomy on the P concentrations within a stratigraphic level was investigated. Previous studies of belemnite biogeochemistry have indicated that concentrations of some elements (e.g., Mg) vary between species, whilst some (e.g., Na, Sr) appear unaffected (McArthur et al., 2007). 4). The Mg and Sr data collected in this study indicated statistically significant differences between species and between stratigraphic levels for both Mg and Sr (Supplementary Fig. 3). A *t*-test showed that there was no significant difference between P concentrations in samples from *Parapassaloteuthis* sp. ($\bar{x} = 295$ ppm, $s = 31$ ppm, $n = 42$) and samples from *Passaloteuthis pessula* ($\bar{x} = 298$ ppm, $s = 26$ ppm, $n = 26$) in samples from the *Amaltheus gibbosus* ammonite subzone; $t(60) = 2.00$, $p = 0.600$. Therefore, no significant taxonomic effect was found in determining P concentrations for the species included in this test, although further testing across a wider range of species is necessary to confirm this. This suggests that samples which cannot be assigned to a particular genus or species may still be used for analysis, including samples that are externally weathered, broken or fragmented.

The largest difference observed was between individuals from different stratigraphic levels. A further *t*-test was performed to compare P concentrations between the stratigraphic levels. There was a significant difference in P concentrations between samples from the *Amaltheus gibbosus* ammonite subzone ($\bar{x} = 290$ ppm, $s = 33$ ppm, $n = 76$) and samples from the *Phricodoceras taylori* ammonite subzone ($\bar{x} = 194$ ppm, $s = 36$ ppm, $n = 52$); $t(103) = 1.98$, $p = 6.86E\text{-}28$. The good degree of reproducibility within stratigraphic levels, combined with the statistically significant difference in P concentrations between stratigraphic levels, indicates that this is a promising method to build stratigraphic profiles of belemnite P concentrations through time. Alternatively, P concentrations correlating with Mn and Fe concentrations in Wine Haven samples but not in samples from Penny Nab, may indicate that different extents of diagenetic alteration between the two levels is responsible for the difference in belemnite P concentrations, and specimens from further levels would need to be analysed to investigate this.

Differences between individuals were also investigated in Modern Analogues of belemnites rostra – *S. officinalis* cuttlebones and *S. spirula* shells (Fig. 6). In *S. officinalis* cuttlefish, there was found to be no statistically significant difference in P concentrations between the two cuttlebones (Cuttlefish A: $\bar{x} = 91$ ppm, $s = 31$ ppm; Cuttlefish B: $\bar{x} = 42$ ppm, $s = 9$ ppm) in samples cleaned with water; $t(2) = 0.115$, $p = 4.30$ (Fig. 6). This is unsurprising as the specimens were of the same species, and raised in the same tank, with similar environmental conditions (similar ammonia, nitrate, phosphate levels and pH), with the only difference being that specimen B had a greater biological age and likely experienced higher average temperatures.

For *S. spirula*, P concentration from two specimens were compared,

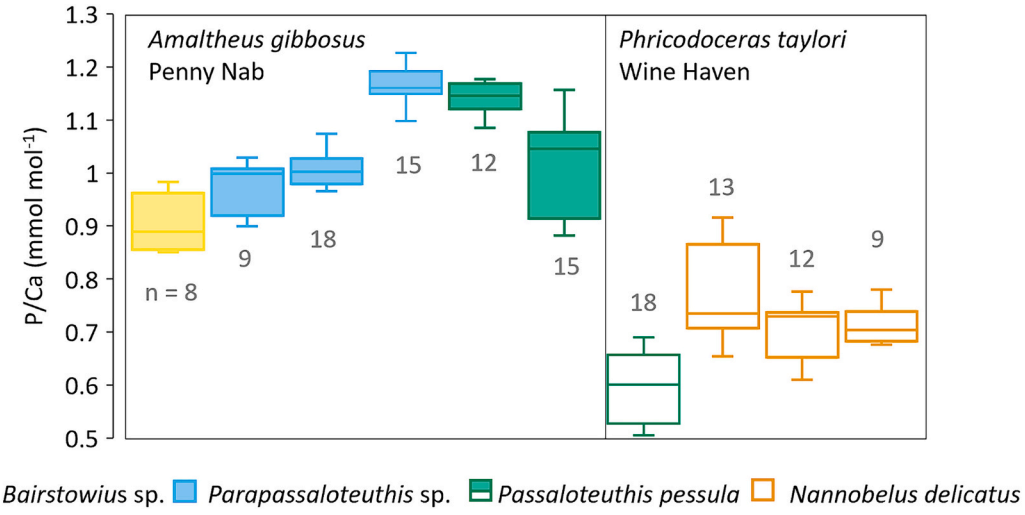

Fig. 11. P/Ca ratios for samples of 6 individual belemnites from the Penny Nab section (*Amaltheus gibbosus* subzone) [filled boxes] and 4 individuals from the Wine Haven section (*Phricodoceras taylori* subzone) [unfilled boxes]. Each colour represents a belemnite species, and each box corresponds to an individual specimen. Each belemnite was sampled at six points along the rostra and all samples were prepared with water cleaning and acetic acid dissolution. Where there was sufficient material, samples were prepared and analysed in triplicate. The total number for each specimen after diagenetic screening criteria was applied is displayed.

with the specimens selected on the basis of having ≥ 3 samples processed by the same cleaning method. It was found that there was no statistically significant difference in P concentrations between the specimens, with one specimen ($\bar{x}$ = 65 ppm; s = 19 ppm; n = 6) giving approximately similar P concentrations to another specimen ($\bar{x}$ = 59 ppm, s = 15 ppm, n = 3) in samples cleaned with water; t(5) = 2.57, p = 0.639. This may be due to greater natural variation in P values between individuals, related to differences in the individuals' behaviours. Each of the *S. spirula* shells were found on beaches in New Zealand, suggesting similar habitats. As wild specimens, however, they are likely to have experienced environmental variability, such as differences in water column depth and location of habitat, which can lead to different average temperatures or different sources of food. They also potentially encompass a range of ages, sexes and sizes which cannot be fully determined by examining the shells alone, particularly as many samples were partial shells.

In this case, the *S. spirula* shells are likely to be a better analogue for the belemnite rostra than the cuttlebone as they have a lower proportion of organic matter. Additionally, the belemnites would have experienced a more natural range of environmental conditions and greater variation in biology. For example, they likely occupied a greater range of different habitats, had different and more varied feeding habits, and there would have been a greater range of biological ages.

To summarise, we suggest that a consistent method be applied in the analysis of CAP to minimise the inclusion of P from other P-bearing phases in the sample. We recommend that this method should include water-cleaning samples and dissolution in acetic acid, to minimise the impact of contaminant P phases (Fig. 10). Additionally, we recommend that for belemnites, consistent sampling methodology should be applied, to reduce the impact of spatial variations of P within the rostra on the CAP measurement (Fig. 10).

3.7. P/Ca ratios

In addition to examining absolute concentrations of P within belemnite calcite, P/Ca ratios were examined as ratios to calcium are an established method of examining CAP (e.g., Montagna et al., 2006; LaVigne et al., 2008, 2010; Anagnostou et al., 2011; Mallela et al., 2013; Chen et al., 2019; Ingalls et al., 2020; Ingalls et al., 2022). P/Ca ratios for belemnites in this study varied between 0.5 and 1.3 mmol mol^{-1}, except for Bulk 1 which had P/Ca ratios between 1.7 and 5.4 mmol mol^{-1} (Fig. 11). Throughout the study, it was found that P/Ca ratios display similar relationships within and between belemnites as absolute P values, with statistically significant differences in P/Ca values between stratigraphic levels (*Amaltheus gibbosus* zone: $\bar{x}$ = 1.042 mmol mol^{-1}, s = 0.106 mmol mol^{-1}, n = 77; *Jamesoni taylori* zone: $\bar{x}$ = 0.684 mmol mol^{-1}, s = 0.09, n = 52; t(119) = 1.98, p = 4.1E-40), but not between different species (*Parapassaloteuthis* sp.: $\bar{x}$ = 1.054 mmol mol^{-1}, s = 0.089, n = 42; *Passaloteuthis pessula*: $\bar{x}$ = 1.070, s = 0.097, n = 27; t(51) = 2.01, p = 0.508) (Fig. 11). This shows that P/Ca values reflect absolute belemnite P concentration, and both are an appropriate method of displaying belemnite CAP values. We suggest that further investigations of belemnite phosphorus rely on P/Ca values, as this minimises errors associated with weighing of samples.

4. Conclusions

Overall, our data demonstrates that different preparative methods have limited impact on CAP measurements for uncontaminated fossil carbonates. The impact of different preparative methods is largest where there are significant P-containing contaminant phases present in the sample, such as organic matter and apatite. In these cases, or where there is no information on contaminating phases, acetic acid is preferred as it discriminates against these more effectively. Water cleaning is also preferred as it limits oxidation of contaminant phases such as organics and reduced iron-minerals, as well as limiting the dissolution of apatite/carbonate fluorapatite, all of which can contribute to an increase in the measured P content. In samples where there is no significant phosphorus contribution from non-carbonate phases, different preparative methods do not produce significantly different results.

These observations are particularly relevant for studies of carbonate associated P in cases where contaminant phases may be expected to be higher, such as bulk rock analyses. In cases where significant material remains after carbonate dissolution, high salinity washes with $MgCl_2$ or NaCl solutions should be used to remove any readsorbed P (e.g., Ruttenberg, 1992; Thompson et al., 2019; Dodd et al., 2021, 2023).

In belemnites, P concentrations varied within and between specimens, but we demonstrate that P measurements are reproducible between replicate samples, within individual belemnites and within stratigraphic levels. There were statistically significant differences in belemnite P concentrations between the stratigraphic levels studied here, indicating the potential for this technique to be used to measure changes in belemnite CAP through time.

CRediT authorship contribution statement

Ailsa C. Roper: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yijun Xiong:** Writing – review & editing, Investigation. **Yafang Song:** Writing – review & editing, Investigation. **Crispin T.S. Little:** Writing – review & editing, Supervision. **Simon W. Poulton:** Writing – review & editing, Supervision. **Paul B. Wignall:** Writing – review & editing, Supervision, Resources. **Clemens V. Ullmann:** Writing – review & editing, Supervision, Conceptualization. **Robert J. Newton:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

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.

Data availability

An Excel file containing all the data and statistical test results used in this paper is avaliable in the supplementary information.

Acknowledgements

We thank the North York Moors National Park Authority for funding the project associated with this publication. We also thank the Jed Atkinson and the Marine Biological Association for assistance procuring samples, and Stephen Reid and Andy Hobson for providing lab and analytical services.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemgeo.2024.122266.

References

Alcott, L.J., Mills, B.J.W., Bekker, A., Poulton, S.W., 2022. Earth's Great Oxidation Event facilitated by the rise of sedimentary phosphorus recycling. Nat. Geosci. 15 (3), 210–215.

Anagnostou, E., Sherrell, R.M., Gagnon, A., Lavigne, M., Field, M.P., McDonough, W.F., 2011. Seawater nutrient and carbonate ion concentrations recorded as P/Ca, Ba/Ca, and U/ca in the deep-sea coral *Desmophyllum dianthus*. Geochim. Cosmochim. Acta 75, 2529–2543.

Baturin, G.N., Bezrukov, P.L., 1979. Phosphorites on the sea floor and their origin. Mar. Geol. 31 (3–4), 317–332.

Birchall, J. D. and Thomas, N. L. (1983) 'On the architecture and function of cuttlefish bone.', J. Mater. Sci., 18, pp2081–2086.

Bjerrum, C.J., Canfield, D.E., 2002. Ocean productivity before about 1.9 Gyr ago limited by phosphorus adsorption onto iron oxides. Nature 417 (6885), 159–162.

Canfield, D.E., Raiswell, R., Westrich, J.T., Reaves, C.M., Berner, R.A., 1986. The use of Chromium Reduction in the Analysis of Reduced Inorganic Sulfur in Sediments and Shales Redox proxies. Chem. Geol. 54, 149–155.

Checa, A.G., Cartwright, J.H.E., Sánchez-Almazo, I., Andrade, J.P., Ruiz-Raya, F., 2015. The cuttlefish Sepia officinalis (Sepiidae, Cephalopoda) constructs cuttlebone from a liquid-crystal precursor. Sci. Rep. 5, 11513.

Checa, A.G., Grenier, C., Griesshaber, E., Schmahl, W.W., Cartwright, J.H.E., Salas, C., Oudot, M., 2022. The shell structure and chamber production cycle of the cephalopod *Spirula* (Coleoidea, Decabrachia). Mar. Biol. 169 (132).

Chen, M., Martin, P., Goodkin, N.F., Tanzil, J., Murty, S., Wiguna, A.A., 2019. An assessment of P speciation and P:Ca proxy calibration in coral cores from Singapore and Bali. Geochim. Cosmochim. Acta 267, 113–123.

Delaney, M.L., 1998. Phosphorus accumulation in marine sediments and the oceanic phosphorus cycle. Glob. Biogeochem. Cycles 12 (4), 563–572.

Dodd, M.S., Zhang, Z., Li, C., Algeo, T.J., Lyons, T.W., Hardisty, D.S., Loyd, S.J., Meyer, D.L., Gill, B.C., Shi, W., Wang, W., 2021. Development of carbonate-associated phosphate (CAP) as a proxy for reconstructing ancient ocean phosphate levels. Geochim. Cosmochim. Acta 301, 48–69.

Dodd, M.S., Shi, W., Li, C., Zhang, Z., Chenga, M., Gu, H., Hardisty, D.S., Loyd, S.J., Wallace, M.W., Hood, A., Lamothe, K., Mills, B.J.W., Poulton, S.W., Lyons, T.W., 2023. Uncovering the Ediacaran phosphorus cycle. Nature 618, 974–980.

Dodge, R.E., Jickells, T.D., Knap, A.H., Boyd, S., Bak, R.P.M., 1984. Reef-building coral skeletons as chemical pollution (phosphorus) indicators. Mar. Pollut. Bull. 15 (5), 178–187.

Doguzhaeva, L.A., Weis, R., Delsate, D., Mariotti, N., 2013. Embryonic shell structure of Early-Middle Jurassic belemnites, and its significance for belemnite expansion and diversification in the Jurassic. Lethaia 47, 49–65.

Filippelli, G.M., 2008. 'The Global Phosphorus Cycle: past, Present, and Future. Elements 4 (2), 89–95.

Florek, M., Fornal, E., Gómez-Romero, P., Zieba, E., Paszkowicz, W., Lekki, J., Nowak, J., Kuczumow, A., 2009. Complementary microstructural and chemical analyses of *Sepia officinalis* endoskeleton. Mater. Sci. Eng. C 29 (4), 1220–1226.

Gröcke, D.R., Hori, R.S., Trabucho-Alexandre, J., Kemp, D.B., Schwark, L., 2011. An open ocean record of the Toarcian oceanic anoxic event. Solid Earth 2 (2), 245–257.

Hoffmann, R., Stevens, K., 2020. The palaeobiology of belemnites – foundation for the interpretation of rostrum geochemistry. Biol. Rev. 95 (1), 94–123.

Hoffmann, R., Richter, D.K., Neuser, R.D., Jöns, N., Linzmeier, B.J., Lemanis, R.E., Fusseis, F., Xiao, X., Immenhauser, A., 2016. Evidence for a composite organic–inorganic fabric of belemnite rostra: Implications for palaeoceanography and palaeoecology. Sediment. Geol. 341, 203–215.

Hoffmann, R., Lemanis, R.E., Wulff, L., Zachow, S., Lukeneder, A., Klug, C., Keupp, H., 2018. Traumatic events in the life of the deep-sea cephalopod mollusc, the coleoid *Spirula spirula*. Deep-Sea Res. I Oceanogr. Res. Pap. 142, 127–144.

Hülse, D., Lau, K.V., van de Velde, S.J., Arndt, S., Meyer, K.M., Ridgwell, A., 2021. End-Permian marine extinction due to temperature-driven nutrient recycling and euxinia', *Nature Geoscience*, 14(11), pp. 862-867. Ruttenberg, K. C. (2003) 'The Global Phosphorus Cycle. Treatise Geochem. 8-9, 585–643.

Ingall, E.D., Buston, R.M., Van Cappellen, P., 1993. Influences of water column anoxia on the burial and preservation of carbon and phosphorus in marine shales. Geochim. Cosmochim. Acta 57, 303–316.

Ingalls, M., Blättler, C.L., Higgins, J.A., Magyar, J.S., Eiler, J.M., Fischer, W.W., 2020. P/Ca in Carbonates as a Proxy for Alkalinity and Phosphate Levels. Geophys. Res. Lett. 47 (21).

Ingalls, M., Grotzinger, J.P., Present, T., Rasmussen, B., Fischer, W.W., 2022. Carbonate-Associated Phosphate (CAP) Indicates Elevated Phosphate Availability in Neoarchean Shallow Marine Environments. Geophys. Res. Lett. 49 (6).

Kumarsingh, K., Laydoo, R., Chen, J.K., Siung-Chang, A.M., 1998. Historic records of phosphorus levels in the reef-building coral Montastrea annularis from Tobago, West Indies. Mar. Pollut. Bull. 36 (12), 1012–1018.

LaVigne, M., Field, M.P., Anagnostou, E., Grottoli, A.G., Wellington, G.M., Sherrell, R.M., 2008. Skeletal P/Ca tracks upwelling in Gulf of Panamá coral: evidence for a new seawater phosphate proxy. Geophys. Res. Lett. 35 (5).

LaVigne, M., Matthews, K.A., Grottoli, A.G., Cobb, K.M., Anagnostou, E., Cabioch, G., Sherrell, R.M., 2010. Coral skeleton P/ca proxy for seawater phosphate: Multi-colony calibration with a contemporaneous seawater phosphate record. Geochim. Cosmochim. Acta 74, 1282–1293.

Lenton, T.M., Watson, A.J., 2000. Redfield revisited: 2. What regulates the oxygen content of the atmosphere? Glob. Biogeochem. Cycles 14 (1), 249–268.

Lenton, T.M., Crouch, M., Johnson, M., Pires, N., Dolan, L., 2012. First plants cooled the Ordovician. Nat. Geosci. 5 (2), 86–89.

Li, Q., McArthur, J.M., Atkinson, T.C., 2012. Lower Jurassic belemnites as indicators of palaeo-temperature. Palaeogeogr. Palaeoclimatol. Palaeoecol. 315-316, 38–45.

Li, Q., McArthur, J.M., Doyle, P., Janssen, N., Leng, M.J., Müller, W., Reboulet, S., 2013. Evaluating Mg/ca in belemnite calcite as a palaeo-proxy. Palaeogeogr. Palaeoclimatol. Palaeoecol. 388, 98–108.

Longinelli, A., Wierzbowski, H., Di Matteo, A., 2002. $\delta^{18}O(PO_4^{3-})$ and $\delta^{18}O(CO_3^{2-})$ from belemnite guards from Eastern Europe: implications for palaeoceanographic reconstructions and for the preservation of pristine isotopic values. Earth Planet. Sci. Lett. 209 (3–4), 337–350.

Mallela, J., Lewis, S.E., Croke, B., 2013. Coral Skeletons provide Historical evidence of Phosphorus Runoff on the Great Barrier Reef. PLoS One 8 (9).

McArthur, J.M., Doyle, P., Leng, M.J., Reeves, K., Williams, C.T., Garcia-Sanchez, R., Howarth, R.J., 2007. Testing palaeo-environmental proxies in Jurassic belemnites: Mg/Ca, Sr/Ca, Na/Ca, $\delta^{18}O$ and $\delta^{13}C$. Palaeogeogr. Palaeoclimatol. Palaeoecol. 252 (3–4), 464–480.

Montagna, P., McCulloch, M., Taviani, M., Mazzoli, C., Vendrell, B., 2006. Phosphorus in cold-water corals as a proxy for seawater nutrient chemistry. Science 312 (5781), 1788–1791.

Monteiro, F.M., Pancost, R.D., Ridgwell, A., Donnadieu, Y., 2012. Nutrients as the dominant control on the spread of anoxia and euxinia across the Cenomanian-Turonian oceanic anoxic event (OAE2): Model-data comparison. Paleoceanography 27 (4).

Mutvei, H., 1964. On the Shell of *Nautilus* and *Spirula* with Notes on the Shell Secretion in Non-cephalopod Molluscs. Almqvist & Wiksell, Sweden.

Okai, T., Suzuki, A., Kawahata, H., Terashima, S., Imai, N., 2007. Preparation of a New Geological survey of Japan Geochemical Reference Material: Coral JCp-1. Geostand. Newslett. 26 (1), 95–99.

Papineau, D., 2010. Global Biogeochemical changes at both ends of the Proterozoic: Insights from Phosphorites. Astrobiology 10 (2), 165–181.

Paytan, A., McLaughlin, K., 2007. The oceanic phosphorus cycle. Chem. Rev. 107 (2), 563–576.

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

Petrochenkov, D.A., Veligzhanin, A.A., Frey, D.I., Chernyshov, A.A., 2018. Riddle of the Nautilus: specific Structural Features of its Shell. Oceanology 58 (1), 45–52.

Planavsky, N.J., Rouxel, O.J., Bekker, A., Lalonde, S.V., Konhauser, K.O., Reinhard, C.T., Lyons, T.W., 2010. The evolution of the marine phosphate reservoir. Nature 467 (7319), 1088–1090.
Qiu, Z., Zou, C., Mills, B.J.W., Xiong, Y., Tao, H., Lu, B., Liu, H., Xiao, W., Poulton, S.W., 2022. A nutrient control on expanded anoxia and global cooling during the late Ordovician mass extinction. Commun. Earth Environ. 3 (1), 1–9.
Reinhard, C.T., Planavsky, N.J., Gill, B.C., Ozaki, K., Robbins, L.J., Lyons, T.W., Fischer, W.W., Wang, C., Cole, D.B., Konhauser, K.O., 2016. Evolution of the global phosphorus cycle. Nature 541 (7637), 386–389.
Richardson, J.A., Roest-Ellis, S., Phillips, B.L., Strauss, J.V., Webb, S.M., Tosca, N.J., 2022. Characterization and Geological Implications of Precambrian Calcite-Hosted Phosphate. Geophys. Res. Lett. 49 (17) e2022GL100328.
Riding, J.B., 2021. A guide to preparation protocols in palynology. Palynol 45 (1), 1–110.
Roest-Ellis, S., Strauss, J.V., Tosca, N.J., 2020. Experimental constraints on nonskeletal $CaCO_3$ precipitation from Proterozoic seawater. Geology 49 (5), 561–565.
Roest-Ellis, S., Richardson, J.A., Phillips, B.L., Mehra, A., Webb., S. M., Cohen, P. A., Strauss, J. V. and Tosca, N. J., 2023. Tonian Carbonates Record Phosphate-Rich Shallow Seas. Geochem. Geophys. 24 (5) e2023GC010974.
Ruttenberg, K.C., 1992. Development of a sequential extraction method for different forms of phosphorus in marine sediments. Liminol Oceanogr 37 (7), 1460–1482.
Ruttenberg, K.C., 2003. The Global Phosphorus Cycle. Treatise on Geochem. 8-9, 585–643.
Sælen, G., 1989. Diagenesis and construction of the belemnite rostrum. Palaeontology 32 (4), 765–797.
Schobben, M., Foster, W.J., Sleveland, A.R.N., Zuchuat, V., Svensen, H.H., Planke, S., Bond, D.P.G., Marcelis, F., Newton, R.J., Wignall, P.B., Poulton, S.W., 2020. A nutrient control on marine anoxia during the end-Permian mass extinction. Nature Geo. 13 (9), 640–646.
Shotyk, W., Immenhauser-Potthast, I., Vogel, H.A., 1995. Determination of nitrate, phosphate and organically bound phosphorus in coral skeletons by ion chromatography. J. Chromatogr. A 706 (1–2), 209–213.
Stevens, K., Mutterlose, J., Ohnemus, B., Idakieva, V., Ivanov, M., 2022. Microstructures of Early Cretaceous Belemnite Rostra and their Diagenesis. Cretac. Res. 137.
Strickland, J.D.H., Parsons, T.R., 1968. A practical handbook of seawater analysis. Bull. Fish Res. Board Can. 167, 1–311.
Thompson, J., Poulton, S.W., Guilbaud, R., Doyle, K.A., Reid, S., Krom, M.D., 2019. Development of a modified SEDEX phosphorus speciation method for ancient rocks and modern iron-rich sediments. Chem. Geol. 524, 383–393.
Tyrrell, T., 1999. The relative influences of nitrogen and phosphorus on oceanic primary production. Nature 400 (6744), 525–531.
Ullmann, C.V., Korte, C., 2015. Diagenetic alteration in low-Mg calcite from macrofossils: A review. Geo. Q. 59 (1), 3–20.
Ullmann, C.V., Pogge Von Strandmann, P.A.E., 2017. The effect of shell secretion rate on Mg/ca and Sr/ca ratios in biogenic calcite as observed in a belemnite rostrum. Biogeosciences 14 (1), 89–97.
Ullmann, C.V., Frei, R., Korte, C., Hesselbo, S.P., 2015. Chemical and isotopic architecture of the belemnite rostrum. Geochim. Cosmochim. Acta 159, 231–243.
Van Cappellen, P., Ingall, E.D., 1996. Redox Stabilization of the Atmosphere and Oceans by Phosphorus-Limited Marine Productivity. Science 271 (5248), 493–496.
Wierzbowski, H., Joachimski, M.M., 2009. Stable isotope, element distribution, and growth rings of belemnopsid belemnite rostra: proxies for belemnite life habitat. Palaios 24 (6), 377–386.
Zhang, N., Lin, M., Yamada, K., Kano, A., Liu, Q., Yoshida, N., Matsumoto, R., 2020. The Effect of H_2O_2 Treatment on Stable Isotope Analysis ($\delta^{13}C$, $\delta^{18}O$ and Δ_{47}) of Various Carbonate Minerals. Chem. Geol. 532.