The shell matrix of the european thorny oyster, Spondylus gaederopus: microstructural and molecular characterization

The shell matrix of the European thorny oyster, *Spondylus gaederopus*: microstructural and molecular characterization

List of authors:

Jorune Sakalauskaite[1,2], Laurent Plasseraud[3], Jérôme Thomas[2], Marie Albéric[4], Mathieu Thoury[5], Jonathan Perrin[6], Frédéric Jamme[6], Cédric Broussard[7], Beatrice Demarchi[1], Frédéric Marin[2]

Affiliations

1. Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy;

2. Biogeosciences, UMR CNRS 6282, University of Burgundy-Franche-Comté, 6 Boulevard Gabriel, 21000 Dijon, France.

3. Institute of Molecular Chemistry, ICMUB UMR CNRS 6302, University of Burgundy-Franche-Comté, 9 Avenue Alain Savary, 21000 Dijon, France.

4. Laboratoire Chimie de la Matière Condensée de Paris, UMR, CNRS 7574, Sorbonne Université, Place Jussieu 4, 75252 Paris, France.

5. IPANEMA, CNRS, ministère de la Culture, UVSQ, USR3461, Université Paris-Saclay, F-91192 Gif-sur-Yvette, France.

6. Synchrotron SOLEIL, L'Orme des Merisiers, 91192 Gif sur Yvette Cedex, France

7. 3P5 Proteomic Platform, University of Paris, Cochin Institute, INSERM, U1016, CNRS, UMR8104, F-75014 Paris, France

Corresponding Authors:
Jorune Sakalauskaite, jorune.sakalauskaite@unito.it, Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin, Italy & UMR CNRS 6282 Biogéosciences, University of Burgundy-Franche-Comté, 6 Boulevard Gabriel, 21000 Dijon, France
Frédéric Marin, frederic.marin@u-bourgogne.fr, UMR CNRS 6282 Biogéosciences, University of Burgundy-Franche-Comté, 6 Boulevard Gabriel, 21000 Dijon, France

Keywords

Biomineralization, Shell biochemistry, evolution, proteomics, liquid chromatography-tandem mass spectrometry

Abstract

Molluscs, the largest marine phylum, display extraordinary shell diversity and sophisticated biomineral architectures. However, mineral-associated biomolecules involved in biomineralization are still poorly characterized.

We report the first comprehensive structural and biomolecular study of *Spondylus gaederopus,* a pectinoid bivalve with a peculiar shell texture. Used since prehistoric times, this is the best-known shell of Europe's cultural heritage. We find that *Spondylus* microstructure is very poor in mineral-bound organics, which are mostly intercrystalline and concentrated at the interface between structural layers.

Using high-resolution liquid chromatography tandem mass spectrometry (LC-MS/MS) we characterized several shell protein fractions, isolated following different bleaching treatments. Several peptides were identified as well as six shell proteins, which display features and domains typically found in biomineralized tissues, including the prevalence of intrinsically disordered regions. It is very likely that these sequences only partially represent the full proteome of Spondylus, considering the lack of genomics data for this genus and the fact that most of the reconstructed peptides do not match with any known shell proteins, representing consequently lineage-specific sequences.

This work sheds light onto the shell matrix involved in the biomineralization in spondylids. Our proteomics data suggest that *Spondylus* has evolved a shell-forming toolkit, distinct from that of other better studied pectinoids - fine-tuned to produce shell structures with high mechanical properties, while limited in organic content. This study therefore represents an important milestone for future studies on biomineralized skeletons and provides the first reference dataset for forthcoming molecular studies of *Spondylus* archaeological artifacts.

Introduction

Biomineralization is a cellular and molecular process by which living systems precipitate mineral salts, predominantly calcium carbonate. In metazoans, this process emerged during the Precambrian-Cambrian transition in multiple lineages. Among these, molluscs represent one of the most diversified phyla, both from the morphological and ecological point of view. Their ability to biomineralize an exoskeleton, the shell, represented an innovative strategy in terms of support of soft body parts and protection against predation and desiccation in addition to constituting an abundant reserve of calcium ions (Kocot et al., 2016). This certainly explains their evolutionary success in marine, freshwater and terrestrial environments.

The biomineralization of the shell is controlled by a specialized organ, the mantle, a thin polarized epithelium that extrudes inorganic ions (mostly calcium and bicarbonate) via ion pumps and channels, and that, at the same time, secretes a macromolecular matrix (Marin et al., 2012). Both inorganic and organic components assemble in well defined microstructures at the nano- and microscales (Marin et al., 2013) resulting in a diversity of micro-architectures

– the most common being the crossed-lamellar, foliated, prismatic and nacreous (mother-of-pearl) ones (Carter, 1990; Taylor et al., 1973, 1969).
The precise molecular mechanisms controlling shell biomineralization are still poorly understood. However, extensive research on the physico-chemical properties of calcium-based biominerals (Lowenstam, 1981) and the shell biochemistry of several model organisms (Addadi and Weiner, 1997; Evans, 2008; Kocot et al., 2016; Marin et al., 2013), have shown that the secreted organic matrix, (*i.e.* a set of proteins, glycoproteins, polysaccharides and lipids), plays a crucial role in the mediation of crystal nucleation and growth (Addadi et al., 2006; Falini et al., 1996; Jackson et al., 2010; Marin et al., 2013).
Very important components are the shell matrix proteins (referred to as SMPs), which get occluded inside the shell mineral during calcification. An increasing number of SMPs have been identified in the past decade using proteomics, transcriptomics and, to a lesser extent, genomics, which allowed researchers to identify and compare the shell proteomes - abbreviated as 'shellomes'- of different groups. Although most of these studies focused on mother-of-pearl (Marie et al., 2017, 2013a; Ramos-Silva et al., 2012; Yarra et al., 2016), they shed light on peculiar characteristics of shell proteins (Arivalagan et al., 2017; Sleight et al., 2016). One interesting feature is the widespread occurrence of low complexity domains/ repetitive low complexity domains (LCDs/RLCDs) (Kocot et al., 2016; Marin et al., 2016, 2014) in nacre proteins (Jackson et al., 2010; McDougall et al., 2013; Sakalauskaite et al., 2019) and also in other microstructures (Marie et al., 2013a). This advocates for a key-role of these domains in binding mineral surfaces and in providing molecular flexibility (Sarashina and Endo, 2006). However, most importantly, 'shellomics' has also evidenced the huge diversity of the SMP repertoires between shells that have similar microstructures or species that are phylogenetically close (Jackson et al., 2010; Marie et al., 2009a, 2009b). This suggests that , beside a set of shared conserved functions in shell repertoires, the recruitment of numerous lineage-dependent molecular functions may have happened in parallel. In other words, independent evolution of SMPs may have occurred multiple times in different lineages.

Whilst there are numerous shell proteome studies of nacreous molluscs (Gao et al., 2015; Jackson et al., 2010; Marie et al., 2017, 2012, 2011, 2009a; Marin et al., 2005), only few studies have been performed on other shell microstructures, such as foliated (Yarra et al., 2016) and cross-lamellar (Herlitze et al., 2018; Osuna-Mascaró et al., 2014). This can be explained by the remarkable mechanical properties of nacre, which make it a material of choice for biomimetics (Espinosa et al., 2009; Finnemore et al., 2012; Luz and Mano, 2009), and by the commercial importance of nacreous molluscs, especially the pearl oyster and the edible mussel (Gosling, 2003; Watabe et al., 2011). In addition, nacre is considered to be an ancestral microstructural type and has been studied from an evolutionary point of view (Jackson et al., 2010; Marie et al., 2009a). On the contrary, the molecular information on the SMPs of crossed-lamellar shells is scarce (Agbaje et al., 2019; Herlitze et al., 2018), despite that it is the most common microstructure in molluscs (Carter, 1990; Wilmot et al., 1992) and possesses high fracture toughness (Kamat et al., 2000).

The thorny oyster *Spondylus gaederopus* Linnaeus, 1758 represents a typical example of this bias in "shellomic" studies. It is a member of a small family of pteriomorphid bivalves, the Spondylidae. *S. gaederopus* is endemic to the Mediterranean sea, with a fixosessil life in water depths up to 50 m. The origin and evolution of spondylids is unclear(Yonge, 1973). Based on morphology and microstructure, the Spondylidae family was thought to be derived from a Pectinid ancestor in the middle Jurassic, possibly from the genus *Spondylopecten (Logan,*

1974; Waller, 2006). This evolutionary process may have been driven by the increase of shell-crushing predators, which likely led to the development of more complex skeletal frameworks with improved density and strength (Knoll, 2003). Conversely, genetics studies suggest that Spondylidae is a sister group of the Pectinidae (Barucca et al., 2004; Matsumoto, 2003; Matsumoto and Hayami, 2000), albeit this issue has not been addressed thoroughly, since many of the molecular phylogenetic analyses use *Spondylus* as an outgroup to reconstruct the Pectinid family (Barucca et al., 2004; Puslednik and Serb, 2008). Considering that Pectinidae shells are mostly foliated calcitic, whereas *Spondylus* is mostly crossed-lamellar aragonitic, the question of evolution is puzzling and requires a careful reexamination.

Finally, *Spondylus* is of particular importance in the field of archaeology and has been regarded as an iconic shell in prehistoric times: its presence is well documented in the archaeological record and can be tracked all over Europe during the Neolithic period (Borrello and Micheli, 2011, 2004; Ifantidis and Nikolaidou, 2011; Moshkovitz, 1971; Shackleton and Elderfield, 1990; Windler, 2017), as well as in Pre-Columbian South America (Paulsen, 1974; Pillsbury, 1996). The possession of *Spondylus* was probably an indicator of wealth or high social status, and therefore reified important ritual/ symbolic messages.

In this context, our study has a dual goal:

1) characterizing the shell microstructure of *Spondylus gaederopus* and the molecular signature of the associated organic matrix, including an in-depth proteomics study;
2) establishing a biomolecular dataset for the *Spondylus* shell matrix, which will constitute a reference in future studies to identify the biological origin of fragments of prehistoric shell ornaments, where morphological identification is unattainable (Demarchi et al., 2014; Sakalauskaite et al., 2019), taking into account the diagenetic processes.

Results

The microstructure of *S. gaederopus* shell

The shell of *S. gaederopus* displays a very distinctive appearance: characteristic hollow spines, a deep purple upper valve and a completely white lower valve (Fig. 1). When observed in cross-section (Fig. 2), it displays a complex combination of microstructures that includes at least 3 different layers: prismatic (P), crossed-lamellar (CL) and foliated (F). At the basis of the shell spines, the microstructure is complex and irregular, but becomes crossed-foliated (calcitic CF) at later stages of spine growth (Fig. 2a, b). The outermost layer of the shell is foliated calcitic (F), similarly to what can be observed in pectinid and ostreid shells (Fig. 2c). The underlying layer is aragonitic crossed-lamellar (CL) and exhibits the typical oblique patterns that characterize this microstructure (Fig. 2d, e). This layer is traversed by a thin prismatic (aragonitic) myostracal layer. The inner and uppermost CL layers, when viewed in cross-section, have different lamellar orientations. The crossed-lamellar layer dominates the hinge area and is very thick at the central part of the shell, but at the margin, the proportion changes, and the foliated becomes the dominant microstructure. The innermost layer is composed of thin aragonitic prisms, about 15-20 µm in diameter, that exhibit a fine layering perpendicular to the elongation axis.

We note that while the changes between the different underlying aragonitic microstructures are gradual, the transition between the aragonitic crossed-lamellar and the foliated calcitic

structures is abrupt and well marked (Fig. 2c). This microstructural and mineralogical transition zone was also studied by UV photoluminescence imaging. The analysis was carried out at the DISCO Beamline (SOLEIL synchrotron, Paris) using an excitation wavelength of 275 nm. The spatial distribution of shell organics was studied using a series of emission filters to detect the luminescence patterns. Fluorescence signal was detected in several different emission regions (327-353; 370-400; 420-480; 499-529). It washowever, the strongest in the [456-481 nm] and [535-607 nm] regions (Fig. 3). An intense luminescenting layer was observed at the interface between the aragonitic crossed-lamellar and the foliated calcitic microstructures, likely arising due to the presence of organics (however, the exact nature of fluorescent components is unknown). This indicates the existence of an intercrystalline organic layer (OL) in the transitional zone between the two layers of different polymorphs.

Figure 1: *Spondylus gaederopus* Linnaeus, 1758 shell, analysed in this study (Greece, Saronikos. Dived at 15 m., taken live in 2010).

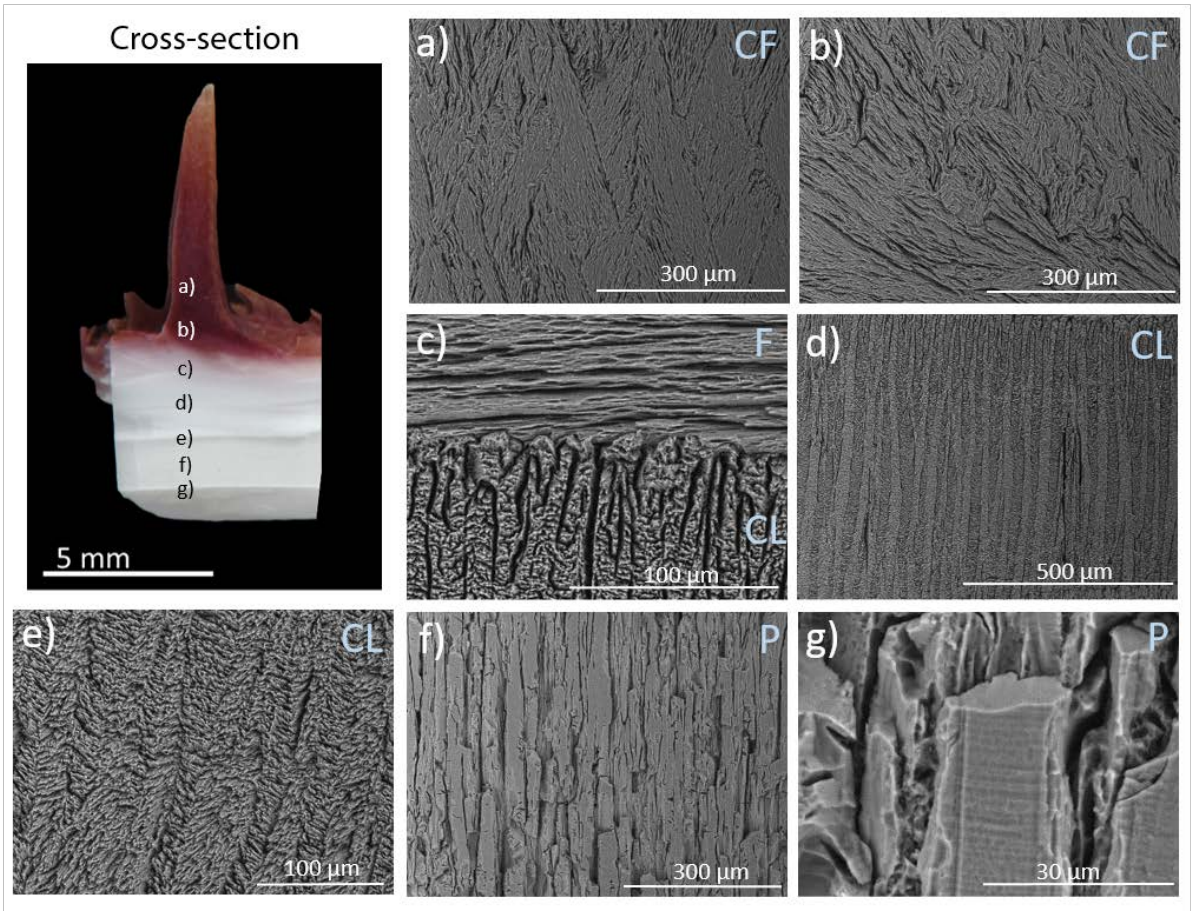

Figure 2: SEM micrographs of the microstructures observed along a polished transversal section of *Spondylus gaederopus* shell (upper valve): a) calcitic crossed-foliated (CF) structure of the spine; b) irregular calcitic CF microstructure observed at the point of spine formation; c) transition between the calcitic foliated (F) and the aragonitic crossed-lamellar (CL) structures; d) aragonitic crossed-lamellar (CL) structure in the upper part of the shell; e) aragonitic crossed-lamellar (CL) structure in the lower part of the section; f,g) aragonitic prismatic (P) layer in the bottom part of the shell and myostracum. Microstructural abbreviations: F - calcitic foliated; P - aragonitic prismatic; CL - aragonitic crossed-lamellar, CF - calcitic crossed-foliated.

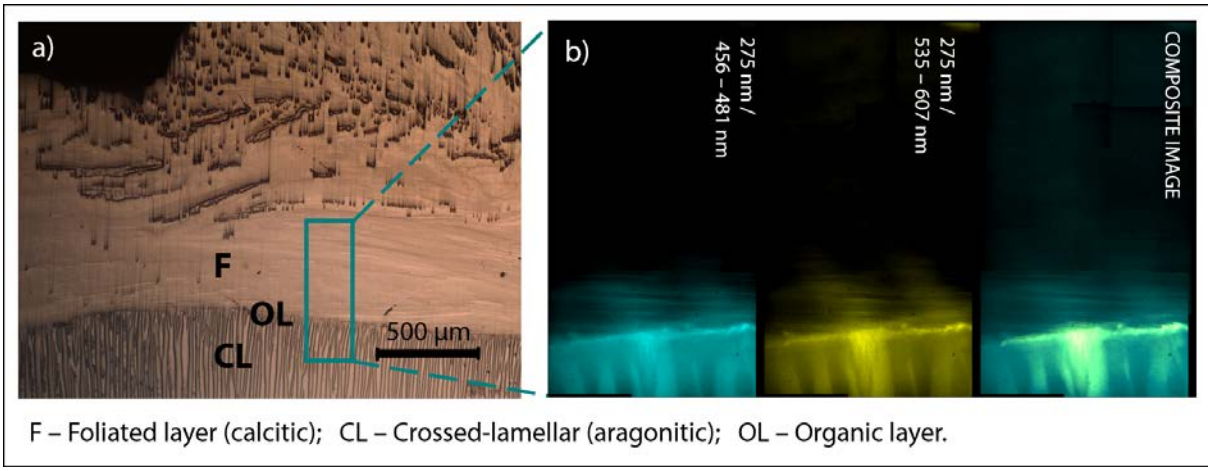

Figure 3. Micro-imaging luminescence of the organo-mineral interface between the foliated calcitic and aragonitic crossed-lamellar layers in the upper valve of *Spondylus* shell, obtained using a DUV photoemission beamline of 275 nm: a) area of interest, picture taken with optical microscope; b) false color luminescence images obtained using 456-481 nm and 535-607 nm filters for the detection. Marked layers: F - foliated, OL - organic layer, CL - crossed-lamellar.

Extraction of the shell organic matrix

The organic matrices were extracted according to standard protocols (Marie et al., 2013b) from finely ground shell powders, which were prepared in different ways. First of all, separated valves were cleaned to remove any surface contamination (mechanically and by bleaching in NaOCl for 2 hrs, see Material and Methods for more details). Following this, we split the coarsely crushed powders into subsamples from the upper (UV) red and lower (LV) white valves. To these, two different bleaching treatments(varying in stringency) were applied to both subsamples:

a) “Mild: approach: coarsely crushed shell powders were bleached for 4 hrs (hereafter “2BL extracts”); these were further ground in fine powders (S1, S2, S3, S4).
b) “Strong” approach: obtained 2BL fine powders were bleached for extra 14 hrs (hereafter “ 3BL extracts”) (S5, S6,S7, S8).

In total, eight shell matrices were obtained - four acid soluble (ASM) and four acid insoluble (AIM) organic fractions (Fig. 4; see also SI.1). After the 2BL treatment, the obtained organics represent ~1.5 - 2 wt‰ (note: per mille) of the total shell powder; more than two thirds of this matrix is the AIM. The 14-hrs additional bleaching step resulted in a severe reduction in the amount of ASM (from ~0.5 wt‰ to ~0.04 wt‰) and in an even more drastic loss of AIM (from ~1.6 wt‰ to ~0.005 wt‰). We note that for both bleaching treatments, the upper valve yielded a slightly higher amount of matrix.

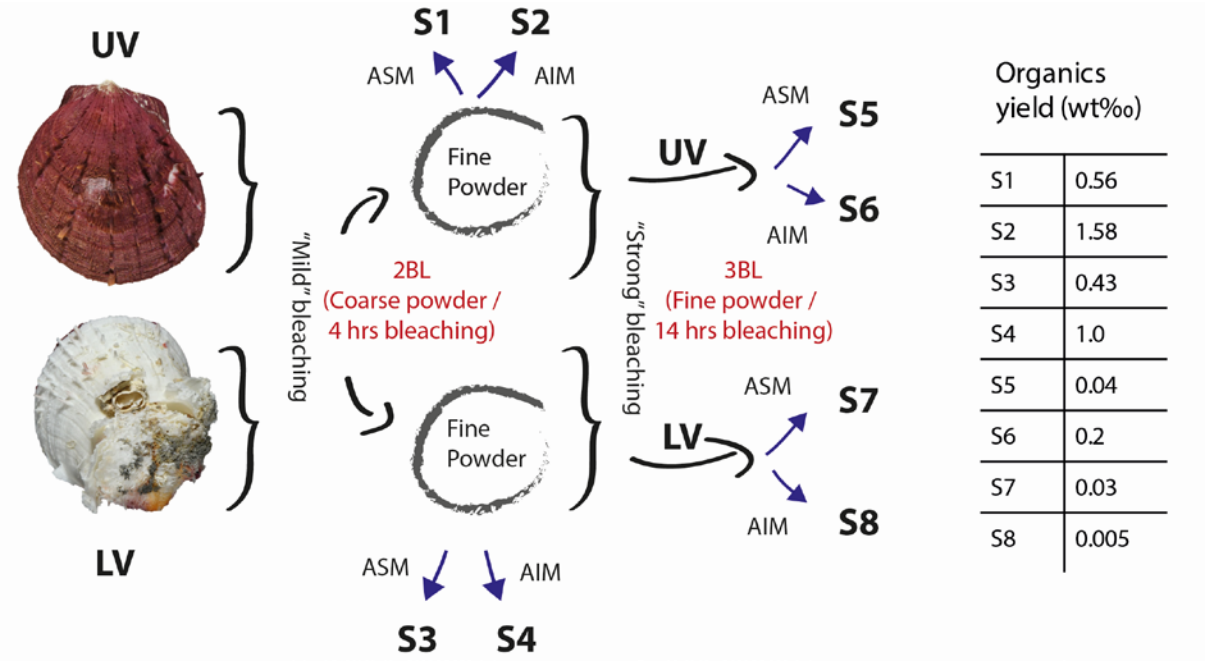

Figure 4: The scheme represents different approaches (bleaching pretreatments) by which shell organic matrix extracts were obtained (S1 - S8) and their subsequent yields. After initial bleach to remove superficial contaminants, all shell powders were bleached twice (2BL) or thrice (3BL). Bleaching procedure: 2BL - coarse powder bleached for 4 hrs; 3BL - fine sieved powder bleached for additional 14 hrs. UV - upper valve, LV - lower valve. ASM - acid soluble matrix, AIM - acid insoluble matrix.

FT-IR analysis of organic matrices

FTIR spectroscopy was used to characterize qualitatively the presence of functional groups in the shell organic matrix extracts. Due to low quantities of the 3BL matrices, the characterization was applied solely on the four 2BL extracts (Fig. 5). Distinctly, all samples exhibit absorption bands characteristic of proteins: the broad band between 3300 and 3260 cm^{-1} corresponding to amide A stretchings ($\boldsymbol{\nu}$N–H), the amide I band ($\boldsymbol{\nu}$C=O) at around 1640-1650 cm^{-1} and the amide II bands ($\boldsymbol{\nu}$C–N) found between 1510-1529 cm^{-1}. In addition, the four spectra present also $\boldsymbol{\nu}$(C–O) absorption bands specific to carbohydrates, at around 1070 cm^{-1}, and another at around 1405-1460 cm^{-1} which can be related to carboxylate groups [$\boldsymbol{\nu}_s(COO^-)$]; the latter appears clear in the two ASMs, while in the AIMs it is shifted towards higher wavenumbers and overlaps with another strong band at 1446-1460 cm^{-1}, only detected in the AIMs. A weak absorption band found in the four extracts at around 1227-1240 cm^{-1} can be assigned to the $\boldsymbol{\nu}$S=O vibration and indicates the presence of sulphate groups (Takeuchi et al., 2018). Moreover, we note that the two ASM spectra (S1, S3) are almost completely superimposable. They mainly differ from the two AIMs (S2, S4) by the amplitude ratios between the amide I and amide II bands. The two AIMs spectra are also very similar, in particular by exhibiting two close bands at 700-712 cm^{-1} associated with a more intense band at 854 cm^{-1}. These bands, including the intense peak at 1446-1460 cm^{-1}, slightly shifted from 1477 cm^{-1}, are characteristic of the aragonite fingerprint (internal vibration modes of CO_3^{2-}), suggesting that some mineral crystallites remain undissolved, in spite of the duration of the decalcification (overnight) and of the excess of acetic acid used for this step. In addition, the

AIMs are characterized by the presence of $\boldsymbol{\nu}$(C–H) stretching vibrations at around 2920-2852 cm^{-1} whereas in ASMs this signal is weaker and shifted to 3072-2943 cm^{-1}.

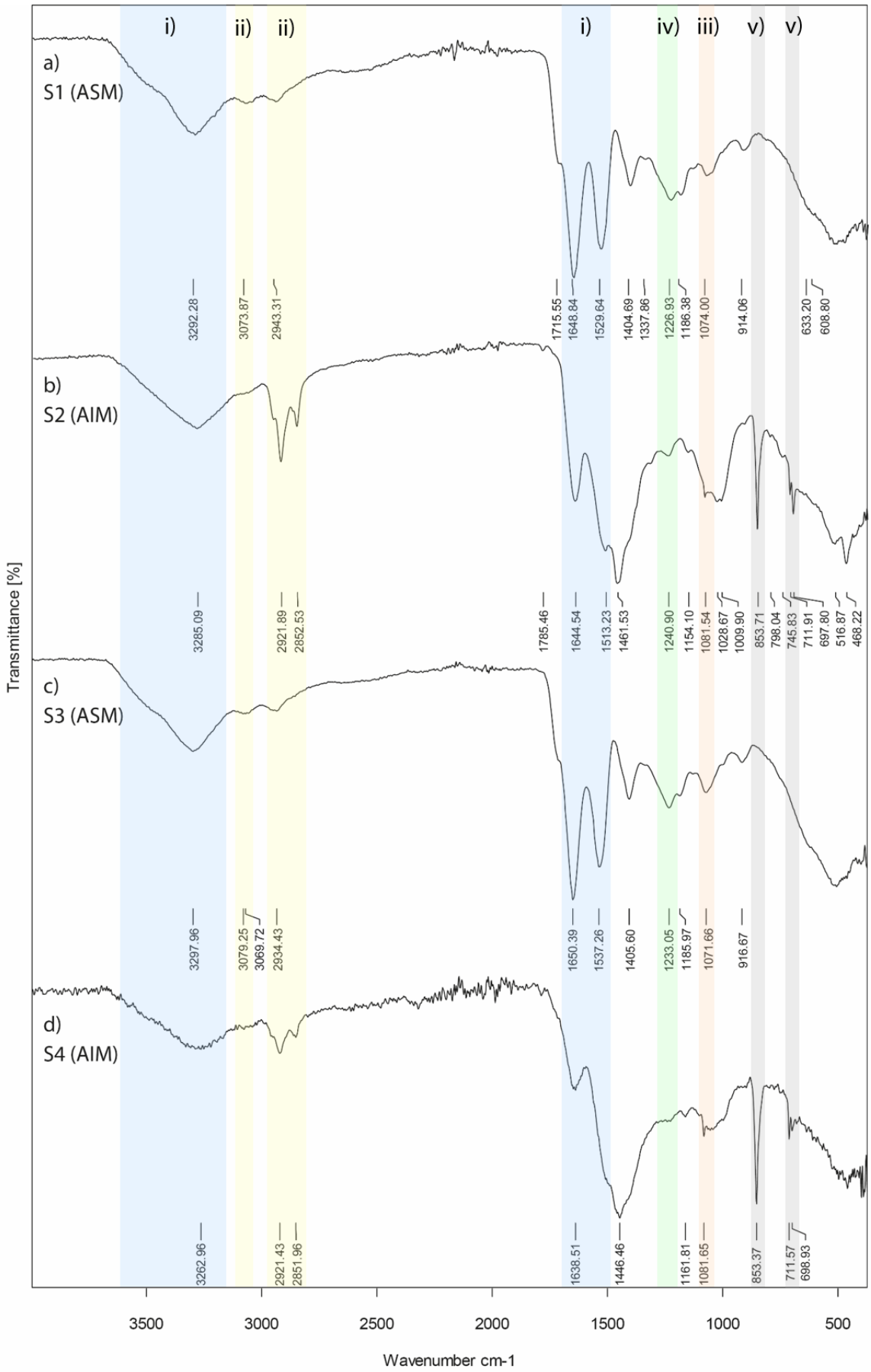
i)
ii)
ii)
i)
iv)
iii)
v)
v)
a)
S1 (ASM)
3292.28
3073.87
2943.31
1715.55
1648.84
1529.64
1404.69
1337.86
1226.93
1186.38
1074.00
914.06
633.20
608.80
b)
S2 (AIM)
3285.09
2921.89
2852.53
1785.46
1644.54
1513.23
1461.53
1240.90
1154.10
1081.54
1028.67
1009.90
853.71
798.04
745.83
711.91
697.80
516.87
468.22
c)
S3 (ASM)
3297.96
3079.25
3069.72
2934.43
1650.39
1537.26
1405.60
1233.05
1185.97
1071.66
916.67
d)
S4 (AIM)
3262.96
2921.43
2851.96
1638.51
1446.46
1161.81
1081.65
853.37
711.57
698.93
Transmittance [%]
3500
3000
2500
2000
1500
1000
500
Wavenumber cm-1

Figure 5: FT-IR (ATR) spectra of shell organic matrices collected in the range of 4000 - 400 cm^{-1}: a, b) S1, S2 - ASM and AIM extracts from the shell upper valve; c, d) S3, S4 - ASM and AIM extracts from the shell lower valve. Peaks discussed in the text are marked in colour which correspond to: i) proteins; ii) organics rich in hydrocarbon groups; iii) glycosides; iv) sulphates; v) aragonite. ASM - acid soluble matrix, AIM - acid insoluble matrix.

SDS-PAGE

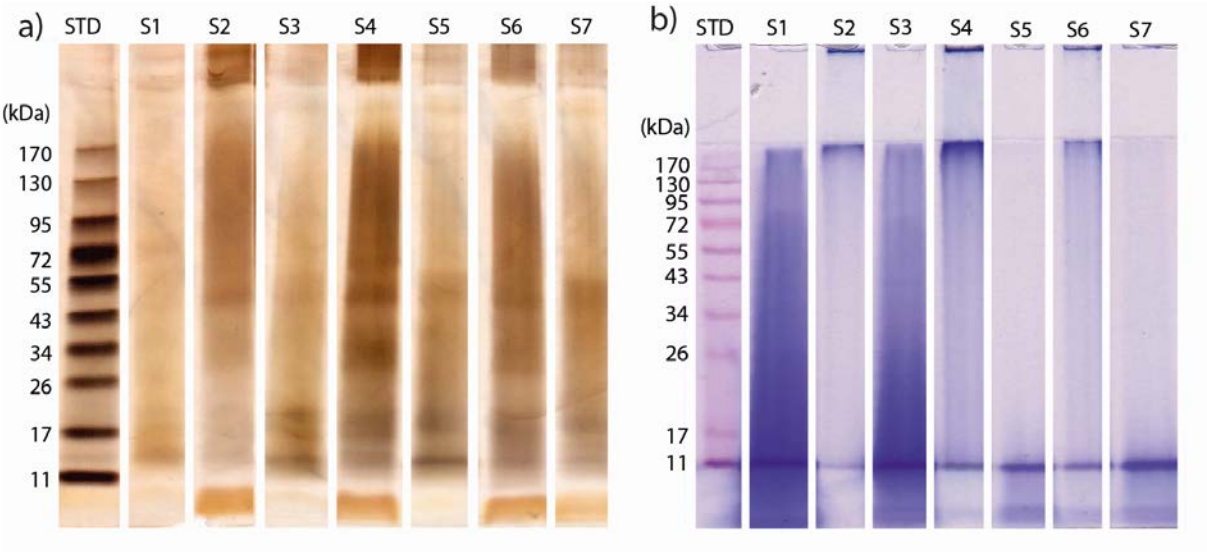

Figure 6: SDS-PAGE gels of *Spondylus* shell organic matrices: a) stained using silver staining; b) stained using Stains-all staining. Analysed samples are as indicated in Figure 3; STD - standard molecular weight markers.

All organic shell extracts were analysed by SDS-PAGE, and stained with silver nitrate and Stains-all (Fig. 6), except sample S8, which yielded a very low quantity and was entirely dedicated to proteomics. Note that the AIM extracts correspond to the Laemmli-soluble AIMs (LS-AIMs).

The silver stained gel (Fig. 6a) shows that all samples have a peculiar migration pattern, characterized by a smear from high to low molecular weights. There are no significant differences between the migration profiles from the upper (S1/2/5/6) and lower valves (S3/4/7/8); the 2BL and 3BL cleaning approaches (S1-4 and S5-7 respectively) produced similar patterns. In spite of the smear, several discrete diffuse bands can be observed in the intermediate molecular weight range: i) a set of three diffuse bands around 20, 17 and 13 kDa in all of the samples; ii) bands in the 70 - 30 kDa range are shared by most of the samples. In the low molecular weight range, a thick band (<11 kDa) is strongly stained brown in only half of the samples (S2/4/6/7) which migrated faster than the migration front, suggesting the occurrence of very low molecular weight compounds. At last, around 11 kDa, a relatively sharp negatively-stained band is present in all samples.

Stains-all (Fig. 6b) allows the detection of putative Ca^{2+} binding proteins (metachromatic blue), but is also efficient in staining acidic proteins; in summary it stains sialoglycoproteins in blue, Ca^{2+} binding proteins in deep blue to violet, proteins in red, and lipids in yellow-orange. All our preparations stained violet at different degrees of intensity, in contrast to the standards, which stained pink. Considering the two cleaning approaches (2BL *vs* 3BL), Stains-all shows that

the signal is far less intense in the 3BL samples (S5-7). Contrary to silver staining, Stains-all does not reveal the blurred bands in the 70 - 11 kDa range. However, it stains intensely violet in all lanes the sharp band at 11 kDa, which is precisely the one negatively stained with silver. We also find low molecular weight components (<11 kDa stained band) in all ASMs (S1/3/5/7), while very high molecular weight (>170 kDa) components are present in the AIMs (S2/4/6).

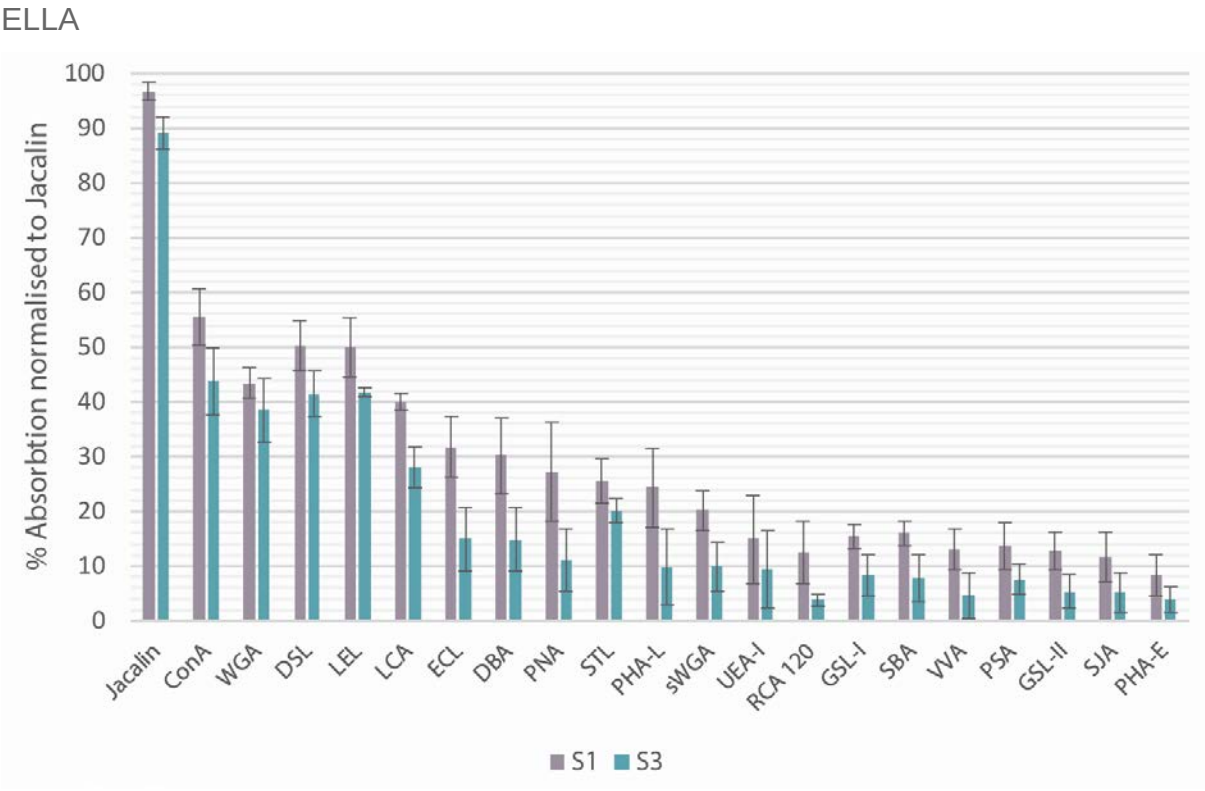

Figure 7: Enzyme-linked lectin assay (ELLA) on two ASMs extracted from the upper (S1) and lower (S3) valves. ELLA was performed with 21 lectins, absorbance values at 405 nm were normalized to the highest value (Jacalin) corresponding to 100% reactivity (n = 4, means ± S.D).

21 lectins were tested for their affinity to saccharidic residues present in soluble shell matrices. The analysis was performed on two ASM samples extracted after 2BL treatments from the upper (S1) and lower (S3) valves and the resulting intensity of cross-reactivity is expressed as a percentage, the most reactive lectin giving 100% (Fig. 7). Jacalin was found to be the most reactive lectin with both extracts. This lectin binds preferentially D-galactose and oligosaccharides terminating with this sugar. It also targets the galactose-β(1-3)N-acetylgalactosamine dimer, and consequently, is specific to O-linked glycoproteins. A group of four lectins including concavalin A (conA), wheat germ agglutinin (WGA), *Datura stramonium* lectin (jimsonweed, DSL) and tomato lectin (*Lycopersicon*, LEL) gave reactivities between 60 and 40%. ConA is a mannose-binding lectin of N-linked glycopeptides, while WGA, DSL and LEL bind monomers to oligomers of N-acetylglucosamine, and thus, are usually considered as chitin-binding lectins. A second group of 6 lectins, including DBA, LCA, ECL, PHA-L, PNA and STL gave moderate reactivities, between 40 and 20%. DBA binds to N-acetyl-D-galactosamine, LCA, to complex α-linked mannose-containing sequences, ECL, to N-acetyllactosamine (dimers of β-galactose and N-acetylglucosamine), PHA-L, to N-linked

oligosaccharides containing galactose, mannose and N-acetyl-glucosamine, PNA, to lactose (dimer of glucose and galactose), and finally, STL, to oligomers of N-acetylglucosamine (chitin-binding). The other ten lectins were considered as unreactive with the extracts (signals below 20%).
The two soluble matrices exhibited very similar reactivity profiles, in particular with the six highest-reacting lectins and the non-reacting lectins. In the moderate reactivity range (20 to 40%), 4 lectins, namelyECL, DBA, PNA and PHA-I, gave slightly different responses. In general, we noticed that in all of the cases S1 gave higher intensity than S3.

In vitro $CaCO_3$ crystallization assay

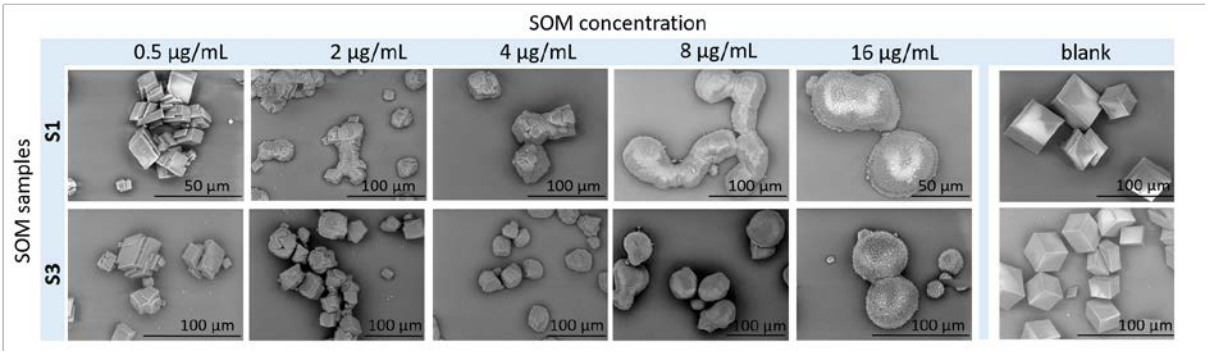

Figure 8: SEM images of the *in vitro* $CaCO_3$ crystallization assay in the presence of *Spondylus* ASMs (fractions S1 and S3) tested at different concentrations: the effect of ASMs was tested at increasing concentrations of 0.5, 2, 4, 8 and 16 $\mu g \cdot mL^{-1}$.

The crystallization assay was used to check in which manner *Spondylus* ASMs (samples S1 and S3, obtained from the upper and lower valves of the shell respectively) interact with the formation of calcite crystals grown *in vitro*. Only the ASM fractions were tested due to the insolubility of their AIM counterparts; the results were obtained at concentrations of 0.5, 2, 4, 8 and 16 $\mu g \cdot mL^{-1}$ of each extract (Fig. 8).
In blank conditions (no matrix), calcite crystals were typically rhombohedral with flat surfaces. The first signs of crystal shape modification could be observed already at low concentrations of ASMs (0.5 $\mu g \cdot mL^{-1}$), where aggregates were obtained. With both ASM extracts, the effects became more pronounced at 2 and 4 $\mu g \cdot mL^{-1}$ concentrations, at which crystals developed rounded edges. At higher concentrations (8 $\mu g \cdot mL^{-1}$ and above), crystal morphologies were completely altered. At 16 $\mu g \cdot mL^{-1}$ drop-shaped forms were produced, from the dense packing of numerous microcrystals. Overall, both ASMs exert similar effects on the crystallization of $CaCO_3$ and this effect is well pronounced already at low matrix concentrations.

In situ lectin-gold localization of sugar moieties

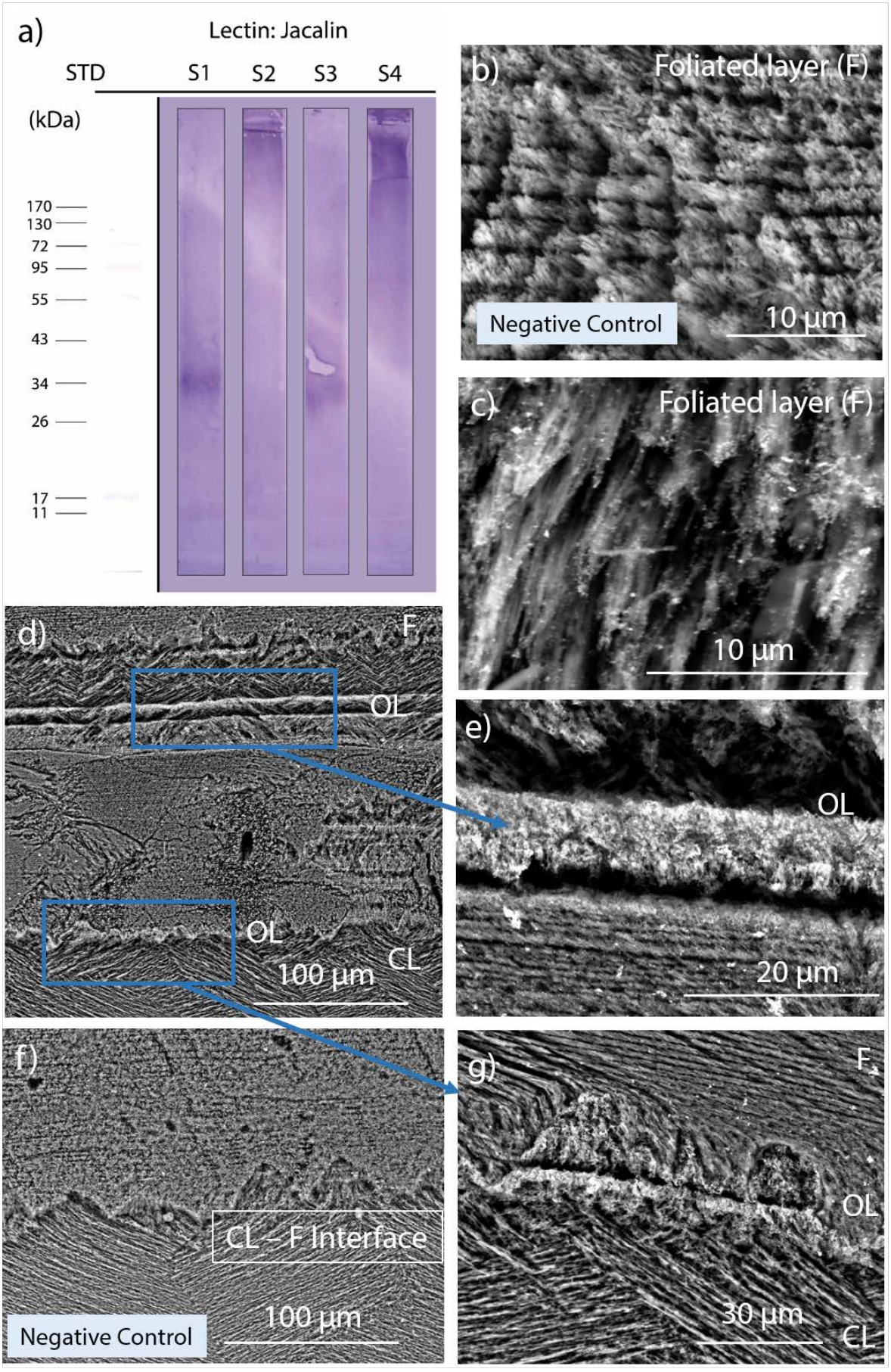
a)
Lectin: Jacalin
STD
S1
S2
S3
S4
(kDa)
170
130
72
95
55
43
34
26
17
11
b)
Foliated layer (F)
Negative Control
10 µm
c)
Foliated layer (F)
10 µm
d)
F
OL
OL
CL
100 µm
e)
OL
20 µm
f)
CL – F Interface
Negative Control
100 µm
g)
F
OL
30 µm
CL

Figure 9: Lectin-gold *in situ* localization assay on *Spondylus gaederopus* shell sections. a) Western blot on *Spondylus* shell organic matrices (S1-4); the membrane was incubated with lectin Jacalin. b-g) *In situ* lectin-gold localization of the sugar moieties to which Jacalin binds in different zones. Images b) and f) corresponds to negative controls. F - foliated calcite layer; CL - aragonitic crossed-lamellar; OL - organic layer.

Lectin-gold *in situ* tests were carried out to directly localize glycans that bind Jacalin along the shell structure. This lectin previously showed the strongest signal in ELLA assay and indicated the presence of galactose/N-acetylgalactosamine residues and/or oligosaccharides terminating with D-galactose. *In situ* localization experiments were performed on small fragments that were sampled in the region close to the margin, where foliated and crossed-lamellar layers are in contact. Jacalin was also tested on Western blots to check its binding specificity to shell matrix components (Fig. 9a). We found that it reacted mostly with a thick discrete band located around 30-34 kDa in the two tested ASMs and it gave a strong, although polydisperse, signal at very high molecular weights in the two tested LS-AIMs.
The experiments showed that Jacalin binds to the sugars groups (possibly glycoproteins) present only in the upper, foliated calcite layer (Fig. 9c) and not in the aragonitic cross-lamellar structure. A very intense signal is observed at the interface between the crossed-lamellar (CL) and foliated (F) microstructures, (Fig. 9g) and also where irregular microstructure transits to well ordered CL layer (Fig. 9d, e). This suggests that such zones concentrate most of the glycoside-rich organic matrix (marking the OL layer). This matrix is also enriched in proteins, as the same area gave a strong signal when stained with the antibody K5090 (elicited against the soluble matrix of the nacreous layer of *Pinna nobilis*) by using immunogold technique (for further details on this see the supplementary information SI.2).

ShellOmics

We characterized the proteome of *Spondylus* shell matrix, by analysing the eight organic extracts (Table 2). The tryptic digests (see Materials and Methods) were analysed by liquid chromatography coupled to a tandem mass spectrometer (LC-MS/MS). The raw mass spectrometry data were analysed using the software program PEAKS Studio 8.5 (Ma et al., 2003), searching against a protein database obtained from NCBI ("NCBI," n.d.) and restricting the taxonomy to Mollusca. We have also separately carried out data analysis using MASCOT search engine against NCBI nr and Swissprot databases of "other metazoa". However, this latter search has resulted in very few findings (only several peptides were identified), therefore, we have focused our analyses using the results obtained via PEAKS search. The number of identified proteins (with a minimum of 2 unique peptides) was higher in the AIM fractions than in the ASM (up to 21 vs 14 sequences respectively). This was also consistent with higher numbers of supporting peptides (up to 67 in AIM v*s* 18 in ASM) and the *de novo* tags (up to ~9100 *vs* 5600) (Table 1).
Considering sequences identified with at least one unique peptide, a total of 46 proteins were found (SI Table 1). Many of these, while poorly covered, were detected in several of the eight extracts and some of them may potentially be relevant for calcification, *e.g.*, alkaline phosphatase, mucin-5AC-like, insoluble matrix protein, extensin-like, ADAM family mig-17, fibronectin type III domain-containing protein 2-like, glycine-rich cell wall structural protein-like, testis-expressed protein 9-like, 15 kDa selenoprotein-like, brevican core protein, and biotin-protein ligase-like.

However, we discuss in detail only the identifications supported by a minimum of two unique peptides, in at least one of the matrices, hence restricting the pool to 6 proteins (Table 2). No significant differences were observed between the lower and the upper valves of the shell in terms of proteome composition. However, the duration of the bleaching treatment appears to have an effect on the quality of the identification: counterintuitively, we obtained more supporting peptides and higher coverages from the 3BL extracts than from the 2BL.

Table 1: Summary of proteomics results: number of identified proteins, total number of supporting peptides, total number of *de novo* only tags present in each sample. The thresholds for peptide and protein identification were set as follows: protein false discovery rate (FDR) = 0.5%, protein score 10lgP > 20, unique peptides 2, *de novo* sequences scores (ALC%) > 50.

	2BL				3BL			
	Upper valve		Lower valve		Upper valve		Lower valve	
Data	ASM S1	AIM S2	ASM S3	AIM S4	ASM S5	AIM S6	ASM S7	AIM S8
Identified proteins / protein families:	2	16	5	6	4	12	14	21
Supporting peptides	4	51	9	22	7	29	18	67
De novo only tags	5864	9186	6651	8189	5610	8260	5340	8046

Table 2: Skeletal matrix proteins identified in *Spondylus* shell. The presented sequences are identified with a minimum of 2 unique peptides (numbers in bold) in at least one of the eight matrices (S1-S8). Cell colour corresponds to the sequence coverage in the sample. The numbers indicate the total number of supporting peptides (Pep), unique peptides (Un) and the identified *de novo* tags (Dn.t) which share at least 6 amino acid residues (see legend for details).

Samples				Protein Hits	Name	Uncharacterized protein LOC110461617 [Mizuhopecten yessoensis]	Uncharacterized protein LOC110452388 [Mizuhopecten yessoensis]	Laccase [Mizuhopecten yessoensis] & Laccase isoforms	G-protein coupled receptor 112 / Uncharacterized protein LOC110443881 [Mizuhopecten yessoensis]	BMSP [Mytilus galloprovincialis]	Acidic mammalian chitinase, chitinase-3-like protein & isoforms [Mizuhopecten yessoensis]
					Protein Characteristics	BMSP-like protein; pI=8.25; Domains: 4 von Willebrand A, 3 CB, RLCD	pI=11.33	Enzyme L-ascorbate oxidase superfamily; pI=6.5; Domains: CuRO_1-3	pI=4.36; Domains: 7tmB2_Adhesion, GPS, MSCRAMM_SdrD/C, CLECT.	pI=8.92; Domains: 4 von Willebrand A, 1 CB, RLCD	Enzyme chitinase; pI=9.03; Domains: GH18_chitolectin, CBM_14, ChtBD2.
2BL	Upper valve	ASM	S1		Pep. (Un)		1 (1)	1 (1)			
					Dn.t		10	2			
		AIM	S2		Pep. (Un)	**22 (22)**	1 (1)	1 (1)	1 (1)	**2 (2)**	**2 (2)**
					Dn.t	140	8	4	22	36	4
	Lower valve	ASM	S3		Pep. (Un)	1 (1)	**3 (3)**				
					Dn.t	31	5				
		AIM	S4		Pep. (Un)	**12 (12)**			1 (1)		
					Dn.t	204			12		
3BL	Upper valve	ASM	S5		Pep. (Un)	1 (1)	1 (1)				
					Dn.t	35	19				
		AIM	S6		Pep. (Un)	**13 (13)**		1 (1)	1 (1)	**2 (2)**	
					Dn.t	180		7	13	41	
	Lower valve	ASM	S7		Pep. (Un)	**3 (3)**	**2 (2)**	1 (1)			
					Dn.t	35	6	4			
		AIM	S8		Pep. (Un)	**30 (29)**	1 (1)	**3 (3)**	**3 (3)**	**3 (2)**	
					Dn.t	234	7	9	13	53	
					Protein NCBI accesion number*	gi|1207960950|ref|XP_021370834.1|	gi|1207934318|ref|XP_021356561.1| ; gi|1207934321|ref|XP_021356563.1|	gi|1205898913|gb|OWF48064.1| ; gi|1207937549|ref|XP_021358299.1 & of other isoforms	gi|1205812537|gb|OWF34702.1| ; gi|1207988369|ref|XP_021343995.1|	gi|347800228|dbj|BAK86420.1|	gi|1205814209|gb|OWF36273.1| ; gi|1207982551|ref|XP_021340846.1|

LEGEND
*Some of the proteins have different accesion numbers in the NCBI database
In bold - proteins that were identified with more than 2 peptides
Pep. (Un) - All supporting peptides (Un - unique peptides)
Dn.t - de novo tags sharing 6 amino acid residues

COVERAGE			
0 - 1 %	2 - 4 %	5 - 8 %	≥9 %

Protein identifications

Five out of six shell proteins identified in *Spondylus* were originally identified from the japanese scallop *Mizuhopecten yessoensis* (family Pectinidae), for which the full genome has been published in 2017 (Wang et al., 2017).

- Uncharacterized protein LOC110461617 (gi|1207960950|ref|XP_021370834.1) was found in all samples except S1, and with 9% coverage in sample S8 (30 supporting peptides and 234 identified *de novo* tags sharing at least 6 AA residues). It is a high molecular weight (289.152 kDa) and basic (theoretical pI 8.25) shell protein, enriched in glycine residues (20.1%), with a notable presence of serine (7.5%), glutamine (7.1%), threonine (6.2%), alanine (6%) and asparagine (6%). The protein exhibits four von Willebrand factor type A domains (VWA) and three chitin-binding domains. In sample S8, these domains are partially covered (Fig. 10). The protein shows the potential occurrence of glycosylation, phosphorylation and N-myristoylation sites (however only a few of the first two modifications have been actually identified in the peptides by the PEAKS software, using the Spider search); there are also predicted NHL repeats (commonly found in a variety of enzymes). Intrinsically disordered regions (IDRs) dominate the C-terminal half of the protein sequence (predicted by ANCHOR2 and IUPred2), but short disordered motifs are also located between the protein family (PFAM) domains (Fig. 10). Overall, an NCBI BLASTp search revealed that this sequence is similar to that of PIF protein (with a query coverage of 68%) and collagen alpha-6(VI) chain (both from *Mizuhopecten yessoensis*), collagen alpha-4(VI) chain (*Crassostrea gigas*) and BMSP (from *Pinctada fucata* and *Mytilus galloprovincialis*), all showing E-values of 0. The homologous BMSP protein from *Mytilus galloprovincialis* (gi|347800228|dbj|BAK86420.1|), a basic calcium-binding protein occurring in nacreous shells (Suzuki et al., 2011), which similarly has four VWA and two chitin-binding domains interspersed by IDRs (SI Fig. 4), was also identified in three of our *Spondylus* AIM extracts (samples S2/6/8).
- Uncharacterized protein LOC110452388 (gi|1207934318|ref|XP_021356561.1|; gi|1207934321|ref|XP_021356563.1|) is a small (280 AA, 30.58 kDa), very basic protein (theoretical pI 11.33) rich in proline (11.4%), arginine (9.6%), glycine (9.3%), serine (8.2%), threonine (8.2%) and asparagine (7.9%). This protein is predicted to have several glycosylation, phosphorylation and N-myristoylation sites (predicted with Interpro (Mitchell et al., 2019)), but none of these modifications have been identified in our sequences by the PEAKS software. Most of the protein is predicted to be intrinsically disordered, with an ordered structure limited solely to the C-terminus (SI Fig. 4). A BLASTp search indicates that it is homologous to other shell proteins such as ubiquilin-like isoform X3 and heterogeneous nuclear ribonucleoprotein 87F-like isoform X2 proteins (both from *Crassostrea virginica*, E-values 2e-05 and 3e-05 respectively), although the maximum query coverage is only around 30%.
- Laccase (gi|1205898913|gb|OWF48064.1|) and laccase-7-like isoforms X1, X2 (all from *Mizuhopecten yessoensis*) were identified in five of the samples (S1/2/6/7/8), albeit with a very weak coverage (1-3%). It is a slightly acidic protein (theoretical pI 6.5) with a molecular weight of 104.16 kDa and is enriched in serine (9.3%), glycine (8.1%), proline (7.4%) and histidine (7.3%). This enzyme belongs to L-ascorbate oxidase superfamily (E-value 7.44e-51) and exhibits three cupredoxin domains (Cu-oxidase_3, Cu-oxidase and Cu-oxidase_2). This protein possesses intrinsically disordered regions, located between the first two cupredoxin domains and at the C-

terminal part of the sequence (SI Fig. 4). It is homologous to other shell proteins including L-ascorbate oxidase [*Crassostrea gigas*], oxidoreductase OpS5-like [*Crassostrea virginica*] and other shell laccase-like proteins such as laccase-2-like [*Crassostrea virginica*], laccase-1-like [*Pomacea canaliculata*] or laccase-4-like from the brachiopod *Lingula anatina*.

- G-protein coupled receptor 112 from *Mizuhopecten yessoensis* (gi|1205812537|gb|OWF34702.1|) was identified only in the AIM matrices (samples S2/4/6/8). It is a very acidic (theoretical pI 4.36) high molecular weight (302.80 kDa) protein, with a composition dominated by serine (11.5%), glutamic (10.9%) and aspartic (8.4%) acids. Threonine (7.5%), leucine (7%), asparagine (6.7%) and isoleucine (6%) are also abundant. It belongs to the GPCR, family 2, secretin-like protein family, which is homologous to the C-type lectin-like superfamily. It possesses immunoglobulin-like domains, GPS motifs and GPCR domains. Most of its sequence is intrinsically disordered, except at the C-terminus (SI Fig. 4). The sequence shows similarities to several other uncharacterized shell proteins and adhesion G-protein coupled receptor G proteins from shells, such as *Crassostrea virginica*, *Pomacea canaliculata*, or the brachiopod *Lingula anatina*.
- Acidic mammalian chitinase (gi|1205814209|gb|OWF36273.1|) and its isoforms from *Mizuhopecten yessoensis* were identified in sample S2 only. This is a basic protein (pI 9.03) with a molecular weight of 78.79 kDa. It has chitinase II and chitin-binding domains as well as predicted intrinsically disordered region in the last third part of the sequence (SI Fig. 4) and shows similarity to the glycoside hydrolase superfamily proteins. The protein is homologous to many other molluscan chitinases, among them those from *Hyriopsis cumingii* (the triangle shell mussel), *Crassostrea gigas* (the Pacific edible oyster) and *Octopus vulgaris* (the common octopus). Chitinases are enzymes that cleave glycosidic bonds of chitin, and are consequently involved in the remodelling of chitinous scaffolds.

Uncharacterized protein LOC110461617 [Mizuhopecten yessoensis].

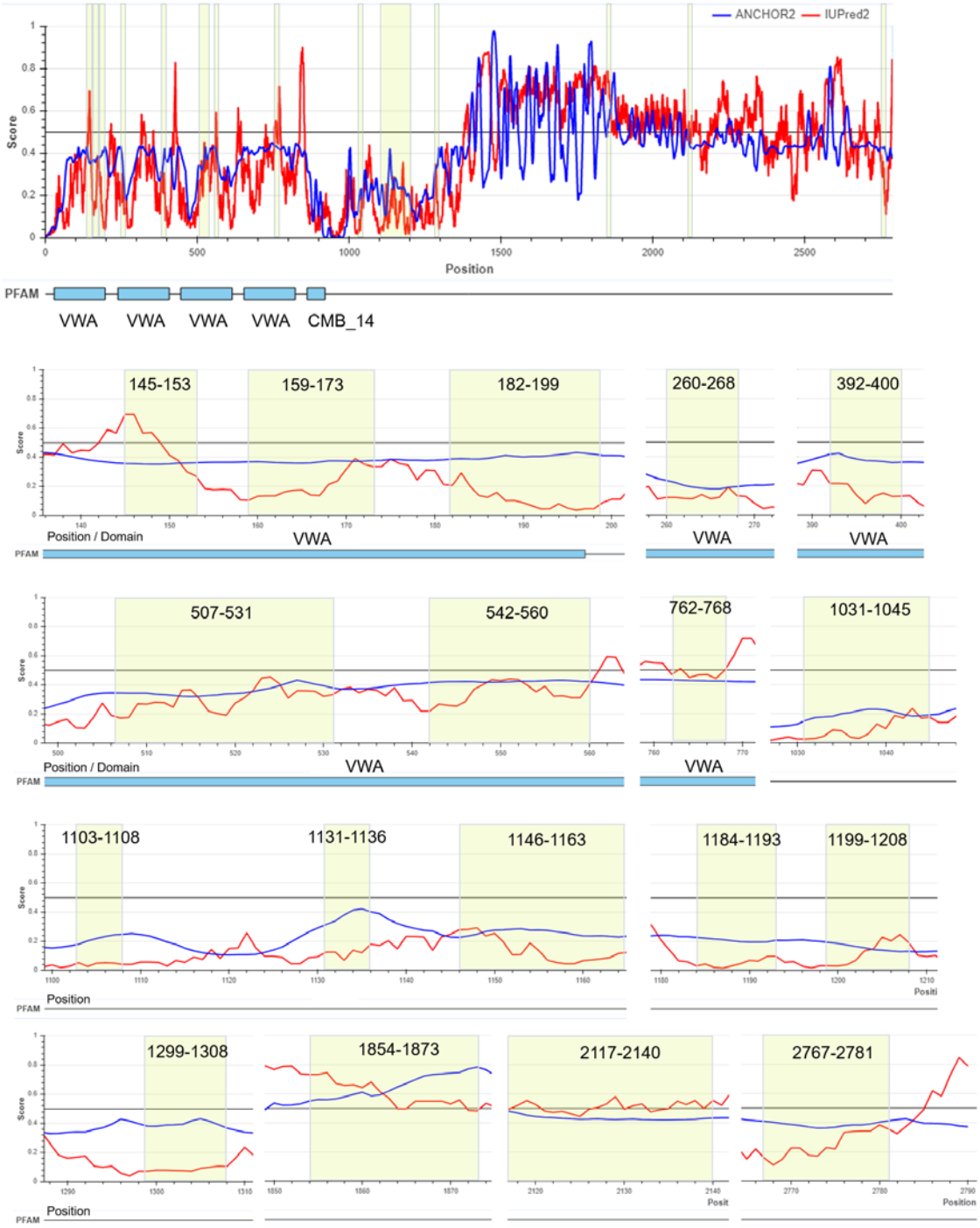

Figure 10: Prediction of intrinsically disordered regions (IDRs) for Uncharacterized protein LOC110461617 [*Mizuhopecten yessoensis*] and the coverage of this protein in sample S8. Structure prediction was obtained using the IUPred2A software (IUPred2 and ANCHOR2 tools). The graph shows the disorder tendency of the residues, where a higher score (y axis) corresponds to a higher probability of disorder (values from 0 to 1). The top graph shows this tendency throughout the full sequence; the regions where covering peptides have been identified (in sample S8) are marked in yellow and detailed in the figures below. Where present, protein domain families (PFAM) are indicated.

Peptidome analysis

Given the low number of proteins identified in the *Spondylus* shell organic matrix using standard database searches, we implemented a different characterisation strategy, focusing on the analysis of the several thousands peptide sequences which can be reconstructed by assisted *de novo* algorithms (using PEAKS Studio 8.5). First, all of the raw MS files were searched against a database of known common laboratory contaminants (cRAP). In all of the samples, between 51 to 226 peptide sequences were identified belonging to contamination (corresponding to human keratin, trypsin, wool keratin, beta-casein) and such contaminant peptides were discarded. We then produced a list of genuine *de novo* peptides (ALC > 80 %) (SI Table2) and searched these "peptidomes" for any pattern emerging in the different extracts (Fig. 11). In agreement with previous findings we observed that the AIMs contain more peptides than the ASMs (~1930 *vs* ~1150). A number of these sequences display amino acid patterns that are typically observed in other shell proteins, such as poly-Gly (predominantly in AIMs), poly-Ser, poly-Gln (predominantly in ASMs) and acidic motifs (EE, DD), but also unusual motifs, like poly-Leu and poly-Pro.

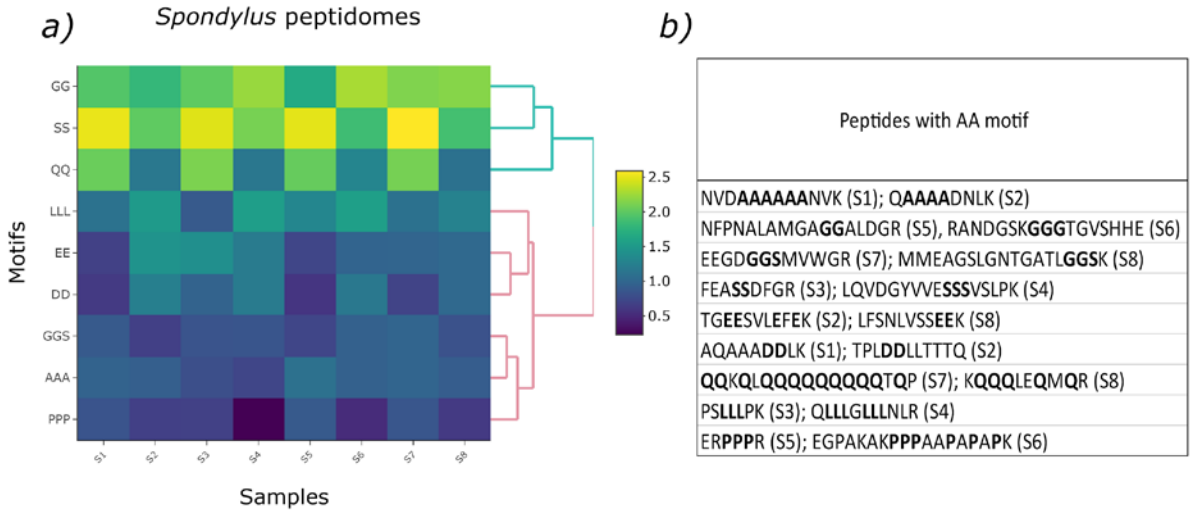

Figure 11: *Spondylus* shell peptidomes (samples S1-S8) characterized by occuring motifs: a) heatmap that represents the occurrence of amino acid motifs (vertical axis) found in peptides of samples S1-S8 (horizontal axes); b) example peptides with these motifs. The full table with peptidome sets can be found in SI Table 2.

Discussion

This study presents a comprehensive biomolecular and micro-structural investigation of the Mediterranean thorny oyster *Spondylus gaederopus*. The bivalve, which underwent a relatively recent evolution, shows divergent molecular features compared to other well studied species. Our discussion examines different aspects, including microstructure, matrix biochemistry and proteomics.

Microstructure

Spondylus, and the typical representatives of the Pectinidae family, *e.g. Pecten,* belong to the same superfamily (Pectinoidea), however their microstructures are different; the only shared characteristic being their upper foliated calcitic layers. Compared to other pectinids (Carter, 1990) *S. gaederopus* shells are certainly more complex in their texture combination, especially with regard to the middle and inner layers, where at least three different microstructures are observed (Logan, 1974): aragonitic prismatic, crossed lamellar (with also different lamellae orientations) and calcitic foliated. Furthermore, the thickness ratio of these three microstructures changes along the growth line, an effect that has been discovered to be temperature-dependent rather than taxon-specific(Nishida et al., 2015, 2012). We also observed that the change between different aragonitic microstructures is very gradual, while the switch between the layers of two different polymorphs (cross-lamellar aragonite and foliated calcite) is abrupt and well marked, confirming earlier descriptions (Wilmot et al., 1992). The long hollow spines, a morphological shell character of spondylids and which function both as an inducement for the settlement of epibionts on the shell and as antipredator-defense (Feifarek, 1987; Logan, 1974), also display a complex microstructure, which is rather disordered at their basis, morphing into crossed-foliated. This microstructural feature has been evidenced in gastropod shells (Fuchigami and Sasaki, 2005; Taylor and Reid, 1990), but, interestingly, has not been reported in bivalves so far.

Biochemistry

The amount of extracted organic fraction in *S. gaederopus* shell is very low. It decreases even further with progressive bleaching times (in 3BL samples, only 0.04-0.25 wt‰ of organics, noted as “intracrystalline”, is obtained).

One may argue that the shell cleaning with NaOCl is the first responsible for a low amount of organic matrix. This is true after the strong 3BL step, but not after 2BL, which remains mild: a similar treatment performed a few years ago on nacre coarse powder clearly showed that the extracted matrix was still abundant, even after an overnight incubation in NaOCl (Parker et al. 2015).

We also note that the isolation of an “operational” intracrystalline fraction of organics is usually carried out using concentrated NaOCl (12% w/v) and exposure times of 48 hours or more (Penkman et al., 2008). However, when considering the 2BL extracts, the reactivity of Jacalin (the most reactive lectin with the sugar moieties present in the shell organics) is comparable

to that observed on 48 hrs bleached extracts (see supplementary information, SI.5); this indicates that the 3BL step isolates a fraction that is in all likelihood “intracrystalline”.

The finding that *Spondylus* matrices are rather organic-poor is consistent with the data obtained on other cross-lamellar (Agbaje et al., 2019; Demarchi et al., 2015; Marie et al., 2013a; Osuna-Mascaró et al., 2014) and foliated (Pierini et al., 2016) mollusc shells; this contrasts to what is observed in organic-rich nacreous and prismatic microstructures (Marie et al., 2009b; Marin et al., 2013; Sakalauskaite et al., 2019). As spondylids represent a bivalve family that emerged rather recently, in the Middle Jurassic (Waller, 2006), our observations corroborate the pioneering idea of Palmer that, in molluscs, the “cost” for making a shell is mostly driven by the “cost” to synthesize the organic matrix (Palmer, 1992); we can argue that evolution was directed to strike a balance between the mechanical properties of a biomineral and the actual “cost” required to produce it. For example, nacre is very “expensive”, while crossed-lamellar microstructures, albeit exhibiting only slightly inferior mechanical properties (Kamat et al., 2000), are “cheaper” to produce. In the course of the Phanerozoic, mollusc evolution has favoured the latter solution; the Spondylidae may illustrate the emergence of very tough microstructures that do not require a lot of mineral-bound shell matrix for their synthesis.

The biochemical composition of *Spondylus* shell matrix is typical of mollusc shells, *i.e.* a mixture of proteins and saccharide moieties that display different solubilities in acidic conditions (hence, the distinction between AIM and ASM). In our study, no compelling differences were observed between the shell matrices of the upper red and lower white valves. On gels, the ASM set of macromolecules, with the exception of few, show mostly non-discrete band profiles - an observation that is similar to most, if not all, shell matrices (Marin et al., 2007). While typical calcium-binding proteins (like calmodulin, which exhibits the EF-hand motif) stain blue with Stains-all, *Spondylus* ASMs stain purple. This feature had already been observed for other acidic shell matrices, such as that extracted from the calcitic prisms of the fan mussel *Pinna nobilis* (Marin et al., 2005). We have also observed that *Spondylus* ASMs (which represent the sum of inter- and intracrystalline matrices) induced strong modifications of the crystal formation during the *in vitro* $CaCO_3$ crystallization; the effect was much more prominent compared to ASMs extracted from other aragonite shells, *e.g.* the land snail *Helix aspersa maxima* (Pavat et al., 2012). This result is interesting because *in vitro* assays typically show that acidic macromolecules (proteins and glycoproteins) induce the strongest effect and are usually considered to be intracrystalline (Albeck et al., 1996, 1993). Indeed, glycoproteins are identified in our *Spondylus* organic matrices, yet, more interestingly the proteomics data reveals that basic proteins are also abundant in the intracrystalline fraction, a finding that expands the knowledge on mineral-bound intracrystalline organics (Demarchi et al., 2016).

Skeletal matrices have their own saccharide signature (Kanold et al., 2015) and the matrix carbohydrates are suspected to perform several functions related to biomineralization (Albeck et al., 1996; Marie et al., 2007; Samata et al., 2008; Sarashina and Endo, 2001). However, they are largely ignored in biomineralization studies, whether they are present as free polysaccharides or glycoproteins. In our case, the lectin assay of the ASMs of *S. gaederopus* underlines the diversity of the sugar residues and of the glycosylation pattern:

1) The preponderance of O-linked glycoproteins and/ or oligosaccharides with Gal, Gal-GalNAc dimers (as observed by the strong affinity to Jacalin);
2) The presence of N-linked mannose containing oligopeptides;

3) The presence of N-acetylglucosamine (or oligomers of this sugar residue), suggesting that chitin is most likely part of the shell organic matrix. As chitin is usually insoluble, it may be either that the glycans in the ASMs are constituted of oligomeric forms, or that the ASM contains a soluble form of chitin, partially cleaved by the bleaching treatment prior to the extraction.

It is interesting to note that Jacalin on Western blots shows different binding profiles in the ASM and LS-AIM matrices; this implies that either the glycan targets are different in their monosaccharide sequences, or that they are identical but differently crossed-linked, *i.e.* the soluble monomeric form being present in the ASM while the crossed-linked one in the LS-AIM. Finally, *in situ* gold-mapping shows that the glycans targeted with Jacalin are not evenly distributed throughout the shell cross-section, but are specific to the upper foliated calcitic layer and are particularly concentrated at the junction with the underlying aragonitic crossed-lamellar layer (Fig. 9), as seen also by luminescence imaging. The data suggest that this organic layer plays an important role in this structural and mineralogical transition - switching from foliated calcite to crossed-lamellar aragonite, while maintaining these two microstructures glued together (Wilmot et al., 1992).

Proteomics

Shotgun proteomics was employed in order to characterise the shell proteome composition. It is however important to stress that the results obtained will be forcibly biased, due to several limitations, including the intrinsic peculiarities of shell matrix proteins and the paucity of -omics reference data for shell matrices, particularly evident in poorly studied organisms such as spondylids.

1) Shell matrix proteins are frequently characterized by the presence of repeated low complexity domains (LCDs or RLCDs, frequently found as long poly-Ala/Asp/Ser/Gly blocks (Kocot et al., 2016; Marie et al., 2017; Sakalauskaite et al., 2019), and post translational modifications such as phosphorylation (Du et al., 2018) and glycosylation (Boskey and Villarreal-Ramirez, 2016; Mann and Edsinger, 2014), making the protein sequences difficult to cleave using standard proteases and thus reducing the chance of identification.
2) The identification is database-dependent. Mollusc shell proteomes are likely to bevery different across taxa and microstructures (Jackson et al., 2010; Marie et al., 2013a) but only a small part of these “shellomes” (mainly those of commercially-relevant species) have been studied so far in a phylum that comprises between 85000 and 100000 shell-forming species (Feng et al., 2017; Marie et al., 2012, 2011; Yarra et al., 2016). Crossed-lamellar microstructures are particularly common among bivalves, gastropods and scaphopods and yet, they are the most understudied.

In our work - the first on spondylids - we had to rely on the omics data available for closely related species (in particular pectinoid shells). However, it is likely that spondylids “shellomes” are much more diversified than presumed, given the complex evolution of Pectinoidea (Smedley et al., 2019).

Comparison between shell extracts

Despite the low organics yield obtained from the bleached shell powders, high-resolution tandem mass spectrometry was successful in obtaining several thousands of product ion spectra for each sample; their sequences were reconstructed using the software PEAKS Studio 8.5.

Spondylus ASMs and AIMs exhibit similar sets of proteins, but more sequences were identified in the AIMs. Interestingly, this is particularly evident for the intracrystalline AIM matrix (the 3BL sample) indicating that by exposing shell to longer bleaching times and removing a large part of the intercrystalline organics, we can facilitate their detection.

The use of PEAKS software resulted in the identification of between 5000-9000 *de novo* sequences for each sample (Table 1). Several of these reconstructed peptide sequences show interesting features that are commonly observed in other shell proteomes (Fig. 11). These include glycine and serine repeats as well as acidic blocks (such as EE, DD) in the AIM extracts and polyQs in ASMs. Another motif, commonly encountered in nacreous shell proteins, poly-Ala blocks, was rare in *Spondylus* proteins.

While, at first sight, *Spondylus* peptidomes resemble “typical” shell protein sequences, the majority of *de novo* sequences could not be assigned to any known molluscan proteins. This clearly indicates that the six identified proteins represent only the partial picture of the full Spondylus shell associated proteome and the true number of proteins is likely much higher. Five out of the six identified sequences belong to the pectinoid *Mizuhopecten yessoensis*, the japanese scallop, which is the phylogenetically closest relative with a full genome analysed (Wang et al., 2017). The coverages were low for all of the identified sequences (1-9%), but most of these proteins possess features typical of shell biomineralization, such as WVA and chitin-binding domains present in Uncharacterized protein LOC110461617 [*Mizuhopecten yessoensis*] and BSMP. Both of these proteins are homologues to PIF, suggesting that these sequences found in *Spondylus* are important in calcium binding activity and possibly play a role in the biomineralization of the aragonitic structures (Arivalagan et al., 2017; Feng et al., 2017; Suzuki et al., 2011).

The two identified enzymes identified have functions related to shell formation: i) chitinase, a chitin binding protein that catalyses the cleavage of chitin, could be involved in the (re)modeling of the chitinous scaffold (Yonezawa et al., 2016); ii) laccase [*Mizuhopecten yessoensis*], a copper oxidase enzyme, could be involved in cross-linking of the matrix during biomineralization. In the prism matrix of the pearl oyster, an enzyme with similar function was identified (Marie et al., 2012), while laccase-like genes have been shown to be involved in shell immunity, biomineralization and pigmentation (Mao et al., 2018; Yue et al., 2019).

We note that all six identified shell matrix proteins exhibit intrinsically disordered regions, IDRs (Fig. 10, Fig. SI.4). An important characteristic of IDRs lies in their flexibility, which enables matrix assembly and protein-mineral interaction (Boskey and Villarreal-Ramirez, 2016; Kalmar et al., 2012). This view is in agreement with the preponderance of these regions in many skeletal matrix proteins (Evans, 2012). The uncharacterized protein LOC110461617 identified in *Spondylus* matrices shows the presence of the IDRs located in between the functional domains (the WVAs in the N-terminal part of the sequence) and dominate almost half of the C-terminal part of the sequence (Fig. 10). Interestingly, the highest coverage of this protein in our samples coincides with the conserved regions, while IDRs have very low and sporadic coverage; since IDRs are believed to evolve fast (Kocot et al., 2016) and to be lineage-specific, this may explain why these domains are poorly covered, when using the sequence data from

other genera as reference. These sequence gaps will be certainly filled by the acquisition of a transcriptome of actively calcifying S. gaederopus specimens.

Conclusion

Our paper describes the first molecular characterization of the shell of *S. gaederopus*, a representative of spondylid family, with a particularly complex microstructure. We found that *Spondylus* possesses a remarkably small amount of mineral-bound skeletal matrix. This latter is rich in glycosidic moieties and is mainly intercrystalline, in particular concentrated at the transition between different microstructural layers. We carried out proteomic analysis, which enabled to partly characterize the shell protein repertoire of *Spondylus.* Even though almost all of the identified sequences were most closely matched to pectinoid shell proteins (belonging to *Mizuhopecten yessoensis*), the low numbers of identified proteins and their scant coverages suggest that the S. gaederopus shellome is quite different from that of Pecten. The lack of genomics data for spondylids implies that most of the Spondylus shell proteins remain largely unknown. Our key finding therefore is that *S. gaederopus* (and this might be extended to the whole spondylid family) is likely to have adopted distinctive molecular strategies to biomineralize its shell. In the future, this may result in a reassessment of the phylogenetic relationships of Spondylidae within the pteriomorphid bivalves clade.
This unique molecular fingerprint, albeit puzzling from an evolutionary point of view, is rich in perspectives for “palaeoshellomics”, *i.e.* the study of ancient shell proteins (Sakalauskaite et al., 2019). This is particularly interesting for biomolecular archaeologists aiming to obtain taxonomic identification of fragmented archaeological shell objects, where biological origin cannot be assigned via classical means (microstructural observations). Our findings suggest that the molecular identification of precious *Spondylus* shell fragments among many other shells in the palaeontological/ archaeological records, is now made possible.

Materials and Methods

Sampling and preparation

Spondylus gaederopus Linnaeus, 1758 (Mollusca; Bivalvia; Pteriomorphia; Pectinida; Pectinoidea; Spondylidae; *Spondylus*) is endemic to the Mediterranean sea and is not listed in the International Union for Conservation of Nature (IUCN’s) list of threatened species. Adult shells (around 100 mm size) were purchased from Conchology, Inc (“Conchology, Inc,” n.d.). The specimens were collected live at 15 m depth (by diving, as listed by the vendors), in Saronikos, Greece, in 2010. In the laboratory, the shell was cleaned using abrasive tools and immersed in diluted NaOCl (~1-1.5% active chlorine) for 2 hrs (we call this step as 1BL - 1st bleach treatment). The upper and lower valves were rinsed with water, air-dried and crushed into ~2 mm fragments with a Jaw-crusher (Retsch BB200) separately, obtaining two sample subsets of the lower valve (LV) and the upper valve (UV). The obtained fragments were bleached in ~1-1.5% NaOCl for 4 hrs (2BL), rinsed with water (5x), ethanol (1x) and air-dried. Coarsely crushed samples were powdered with a mortar grinder (Frisch Pulverisette 2) to particle size < 200 µm (sieving). Half of both subsamples (UV and LV) were set aside (2BL samples) and the second batch was bleached for the third time in ~1-1.5% NaOCl for 14 hrs (3BL), thoroughly rinsed with water (5x) and air-dried. In total, 4 powdered samples were

obtained: 2 batches of different cleaning approach (2Bl and 3 BL) each with 2 subsamples - upper and lower valves (UV and LV).

Extraction of shell organic matrix (SOM)

The extraction was carried out using a well-established protocol developed in our lab (Ramos-Silva et al., 2012; Takeuchi et al., 2018). In short, powdered samples were decalcified using cold diluted acetic acid (10% v/v, 100 µL every 5 seconds) overnight and then centrifuged (3893 G, 30 min.) to separate the acid soluble (ASM) and acid insoluble matrices (AIM). The AIM matrix was rinsed (5x)with ultra pure water and freeze-dried. The ASM was ultrafiltered using 10 kDa cut-off membrane (Millipore, Ref. PLGC07610), dialysed against 1L of MiliQ water, with 6 changes over 2 days, and freeze-dried.

Biochemical characterization

FT-IR spectroscopy

FT-IR analysis in Attenuated Total Reflectance (ATR) mode was used to check the overall composition of ASM and AIM matrices of the 2BL samples. The analysis was performed with a Bruker Vector 22 FT-IR spectrometer (BrukerOptics Sarl, France, Marne la Vallée) fitted with a GoldenGate ATR device (SpecacLtd, Orpington, UK) in the 4000–400 cm^{-1} range (twelve scans with a spectral resolution of 4 cm^{-1}). Spectra analysis and assignment of absorption bands were performed by comparison with previously published data (Kanold et al., 2015; Takeuchi et al., 2018).

SDS-PAGE

ASMs were fully dissolved in ultrapure water to a final concentration of 4 µg/mL and 4x Laemmli buffer was added (ratio 3:1). AIMs were resuspended and treated similarly. Denaturation was carried out at 99 °C for 5 min, then samples were cooled on ice. Because of limited AIM solubility, only the Laemmli-soluble supernatants, referred to as LS-AIM, were analysed by SDS-PAGE. Proteins were run on precast (4-15 %) gradient (Bio-Rad) and 12% gels, then stained with silver nitrate and Stains-all (Campbell et al., 1983; Takeuchi et al., 2018) respectively.

Western-blot

Samples S1 and S2 were run on a 12% SDS-PAGE gel as reported above, then electro-transferred on a PVDF Immobilon P membrane (Millipore) (Matsudaira, 1987). To detect the glycosylated complex, the membrane was incubated in Carbo-Free blocking solution (Vector Laboratories, Peterborough, UK, ref. SP-5040) then in a solution that contained the biotinylated lectin Jacalin (Vector Labs. Ref. B-1155, diluted in TBS/Tween20, ratio 1:100) and finally, after rinsing, in a solution containing avidin-AP conjugate (Sigma A7294, dilution 1/70000). Thorough rinsing steps with TBS/Tween20 were performed. The membrane was stained with NBT/BCIP (Sigma B5655).

Enzyme-Linked Lectin Assay (ELLA)

The enzyme-linked immuno assay was carried out on two ASM fractions following previously described procedures (Kanold et al., 2015; Takeuchi et al., 2018). In short, ASM aqueous solutions (100 µL per well, approx. ~200 ng/ well) were deposited in 96-well microplates (MaxiSorp, Nunc/Thermo Scientific, Nunc A/S, Roskilde, Denmark), incubated for 90 min at

37 °C, washed with TBS/Tween20 and blocked with Carbon-Free solution. Three sets of 7 biotinylated lectins were tested (Vector Laboratories, Peterborough, UK, Ref. BK-1000, BK-2000, BK-3000; diluted in TBS/Tween-20 with ratio 1:200 for kits 1,2 and 1:100 for kit 3); lectins were incubated for 90 min at 37 °C, washed, and a solution of alkaline phosphatase-conjugated avidin (A7294 Sigma, 1:70000) was added for another 90 min at 37 °C. Microplates were washed and incubated with the ELISA (aqueous diethanolamine solution, pH 9.8) substrate solution, containing phosphatase substrate (p-nitrophenylphosphate) at 37 °C. Optical density was read using a BioRad Model 680 microplate reader at 405 nm; background values (blank tests) were subtracted and the values were converted to percentage of reactivity, with the highest OD values at 100%. A total of 4 replicates were performed.

In vitro $CaCO_3$ crystallization

The capacity of ASM to interact *in vitro* with $CaCO_3$ ASM was tested in the diffusion assay, as described previously (Kanold et al., 2015; Takeuchi et al., 2018). 200 µL of 10 mM $CaCl_2$ solutions containing ASMs at increasing concentrations (0.25 – 16 µg/ml) were placed in 16-well culture slides (Lab-Tek, Nunc/ Thermo Scientific, Rochester, NY, USA). The plastic cover was pierced beforehand to allow the reaction with ammonium bicarbonate vapours, placed on top of the slide and sealed with parafilm. It was placed in a desiccator under vacuum which contained ammonium bicarbonate crystals and incubated at 4 °C for 72 hrs. Afterwards, the solution was carefully removed with a blunt-end needle and the slide was air-dried. The 16-well culture slide was disassembled and $CaCO_3$ crystals were observed directly (without any carbon sputtering) using a tabletop scanning electron microscope (TM 1000, Hitachi).

In situ localization

Lectin-gold assay

Using a diamond saw, a transverse section, perpendicular to the growth line, was cut from the upper valve of the shell (area closer to the margin). The section was polished using a fine-grain sandpaper, followed by an alumina suspension (0.05 µm) and cut into small square fragments. They were cleaned using sonication, then bleached for 10 min with NaOCl (~1-1.5 % active chlorine) and rinsed with ultrapure water. Fragments were etched with 1% EDTA (v/v) in the ultrasound bath for 5 min and rinsed again with ultrapure water. The rest of the procedure follows that reported in (Takeuchi et al., 2018): incubation in Carbo-Free blocking solution (1 hour); overnight incubation with Jacalin solution (prepared in TBS/Tween-20, diluted 1:10; plus NaN_3, 1/10000 to prevent bacterial growth); rinsing step with TBS/Tween-20 (6x times, 10 min each); incubation with goat anti-biotin antibody conjugated to ultra-small gold particles (0.8 nm) (GABio, Ultra Small, Ref. 800.088; Aurion, Wareningen, The Netherlands) for 2 hrs; after rinsing step and gentle drying, incubation in the silver enhancement solution (British Biocell International, Ref. SEKL15) for about 10-15 min; rinsing in milli-Q water and drying at 37 °C. All incubations were performed at room temperature with gentle shaking. Samples were observed with SEM (Hitachi, TM1000) under back-scattered electron mode

To preclude false positive signals, 3 negative controls were performed in parallel using the same conditions but omitting one or two of the following steps: 1) incubation with Jacalin (negative control 1); 2) incubation with gold-coupled anti-biotin (negative control 2); 3) both incubations with Jacalin and the gold-coupled anti-biotin (negative control 3). All other steps

were done similarly. Total of 4 replicas of the experiments were carried out to validate the observations.

Luminescence imaging by DUV synchrotron photoemission beamline

The Deep-Ultraviolet (DUV-visible) photoluminescence imaging was applied to trace the distribution of organics in the biomineral of a selected area. Fluorescence maps were recorded on the DISCO beamline (Jamme et al., 2013) using the setup TELEMOS build around an Olympus IX71 inverted microscope stand and a Peltier iDus CCD detector with 26x26 µm pixel size. A small fragment was cut from the upper valve of the shell, that also included a formed spine, and was mirror-polished. The area was scanned using a x40 objective and the luminescence maps were reconstructed and corrected using an in-house script in Matlab, developed at DISCO beamline by one of the authors (F.J.).

Proteomic analysis

For proteomic analyses, digestion with trypsin of each of the eight shell organic extracts were carried out in-gel after a short migration (around one centimeter in the separating gel) in an 8% acrylamide gel stained with Coomassie blue. In brief, after the gel runs, samples were destained twice with a mixture of 100 mM ammonium bicarbonate (ABC) and 50% (v/v) acetonitrile (ACN) for 30 min at 22 °C and then dehydrated using 100% ACN for 15 min, before being reduced with 25 mM ABC containing 10 mM DTT for 1 h at 56 °C and alkylated with 55 mM iodoacetamide in 25 mM ABC for 30 min in the dark at 22 °C. Gel pieces were washed twice with 25 mM ABC and dehydrated (twice, 20 min) with 100% ACN. Gel cubes were incubated with sequencing grade modified trypsin (Promega, USA; 12.5 ng/µl in 40 mM ABC with 10 % ACN, pH 8.0) overnight at 37 °C. After digestion, peptides were extracted twice with a mixture of 50% ACN – 5% formic acid (FA) and then with 100% ACN. Extracts were dried using a vacuum centrifuge Concentrator plus. Resulting peptides were resuspended in 10 µL of 10% ACN - 0.1 % trifluoroacetic acid (TFA) and subjected to analysis by LC-MS/MS.

Mass spectrometry (MS) was performed using an Ultimate 3000 Rapid Separation Liquid Chromatographic (RSLC) system (Thermo Fisher Scientific) online with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific). 1µL of peptides was loaded, concentrated and washed on a C_{18} reverse phase pre-column (3 µm particle size, 100 Å pore size, 75 µm i.d., 2 cm length, Thermo Fisher Scientific). The loading buffer contained 98% H_2O, 2% ACN and 0.1% TFA. Peptides were then separated on a C_{18} reverse phase resin (2 µm particle size, 100 Å pore size, 75 µm i.d., 25 cm length, Thermo Fisher Scientific) with a 1 hour gradient from 99% solvent A (0.1% FA and 100% H_2O) to 90% solvent B (80% ACN, 0.085% FA and 20% H_2O).

The mass spectrometer acquired data throughout the elution process and operated in a data dependent scheme with full MS scans acquired with the Orbitrap, followed by MS/MS HCD fragmentations acquired with the Ion Trap on the most abundant ions detected. Mass spectrometer settings were: full MS (AGC: 2x10e5, resolution: 6x10e4, m/z range 350-1500, maximum ion injection time: 60 ms) and MS/MS (HCD collision energy: 30%, AGC: 2x10e4, resolution:3x10e4, maximum injection time: 100 ms, isolation windows: 1.6m/z Da, dynamic exclusion time setting: 30 s). The fragmentation was permitted for precursors with a charge state of 2, 3, 4 and up.

Bioinformatic search

Product ion spectra were analysed using PEAKS Studio (v. 8.5, Bioinformatics Solutions Inc. (BSI) (Ma et al., 2003)) and searched against a "Mollusca protein" database, downloaded from NCBI ("NCBI," n.d.) on 19/12/2018, restricting the taxonomy to Mollusca, including the search of common laboratory contaminants (cRAP; common Repository of Adventitious Proteins: http://www.thegpm.org/crap/). Search parameters were defined assuming no enzyme digestion (in order to also detect naturally cleaved peptides, if any), fragment ion mass tolerance of 0.05 Da and a parent ion tolerance of 10 ppm. Results obtained by SPIDER searches (*i.e.* including all possible modifications) were used for peptide identification and protein characterization, choosing the following threshold values for acceptance of high-quality peptides: false discovery rate (FDR) threshold 0.5%, protein scores −10lgP ≥ 20, unique peptides ≥ 2, *de novo* sequences scores (ALC %) ≥ 50. An additional full list of protein sequences identified with just one unique peptide were also exported. The peptide sequences identified in shell proteins were individually checked using the Blastp tool (https://blast.ncbi.nlm.nih.gov/), and any sequences that were homologous to common laboratory contaminants were excluded from any further analysis.

Protein profiles

Protein physical and chemical parameters (MW, theoretical PI, amino acid composition) were assessed using ProtParam tool on the EXPASY server. Protein classification, prediction of domains and motifs was carried using Interpro database (Mitchell et al., 2019) and Motif scan tool on the EXPASY server. Intrinsically disordered regions in proteins were predicted using IUPred2A (Mészáros et al., 2018) and the bLASTp tool was used to identify homologous proteins.

Peptidome analysis

For the shell peptidome analysis, the reconstructed *de novo* tags in PEAKS were searched against a database of common laboratory contaminants (cRAP, search parameters defined above). The results obtained by SPIDER search were used by selecting these threshold values: false discovery rate (FDR) threshold 0.5%, protein scores −10lgP ≥ 20, unique peptides ≥ 0. Only the very high quality *de novo* sequences were taken into account, by selecting the local confidence score ALC % ≥ 80; exported sequences were grouped in Supplementary Information Table 2.

Data deposition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al., 2013) partner repository with the dataset identifier PXD016760.

Acknowledgements

JS, FM and BD would like to thank Emmanuel Fara (University of Burgundy-Franche-Comté) and Matthew Collins (Universities of Copenhagen and Cambridge) for valuable insights and discussions.

JS, FM and BD are supported by the PHC Galilée programme, Italo-French University (UIF/UFI) (project G18-464 / 39612SB) and JS acknowledges the support of the Campus France fund obtained from French embassy in Italy (BGF) and program Eiffel.

BD is funded by the “Giovani Ricercatori - Rita Levi Montalcini” Programme (MIUR; Ministero dell'Istruzione dell'Università e della Ricerca).
The work performed at UMR CNRS 6282 Biogéosciences (JS, FM), Dijon was financed via the annual recurrent budget of this research unit allocated to JS and FM, and via extra CNRS funding (INTERRVIE).
The work performed at SOLEIL synchrotron (JS, FM, MA, MT, JP, FJ) was financed via proposal no. 20180213 at DISCO IMAGING beamline.

Author contribution

F.M, J.S, B.D: Conceptualization. J.S, F.M, L.P, J.T, M.A, F.J, C.B: Investigation, methodology. M.A, F.J, M.T, J.P: Resources. J.S., F.M., B.D, L.P: Data curation. J.S, F.M and B.D: Writing - original draft. F.M, B.D: Project administration, Funding acquisition.

Competing interests statement

The authors declare that no competing interests exist.

References

Addadi, L., Joester, D., Nudelman, F., Weiner, S., 2006. Mollusk shell formation: a source of new concepts for understanding biomineralization processes. Chemistry 12, 980–987.
Addadi, L., Weiner, S., 1997. A pavement of pearl. Nature 389, 912–913.
Agbaje, O.B.A., Thomas, D.E., Dominguez, J.G., Mclnerney, B.V., Kosnik, M.A., Jacob, D.E., 2019. Biomacromolecules in bivalve shells with crossed lamellar architecture. J. Mater. Sci. 54, 4952–4969.
Albeck, S., Addadi, I., Weiner, S., 1996. Regulation of calcite crystal morphology by intracrystalline acidic proteins and glycoproteins. Connect. Tissue Res. 35, 365–370.
Albeck, S., Aizenberg, J., Addadi, L., Weiner, S., 1993. Interactions of various skeletal intracrystalline components with calcite crystals. J. Am. Chem. Soc. 115, 11691–11697.
Arivalagan, J., Yarra, T., Marie, B., Sleight, V.A., Duvernois-Berthet, E., Clark, M.S., Marie, A., Berland, S., 2017. Insights from the Shell Proteome: Biomineralization to Adaptation. Mol. Biol. Evol. 34, 66–77.
Barucca, M., Olmo, E., Schiaparelli, S., Canapa, A., 2004. Molecular phylogeny of the family Pectinidae (Mollusca: Bivalvia) based on mitochondrial 16S and 12S rRNA genes. Mol. Phylogenet. Evol. 31, 89–95.
Borrello, M.A., Micheli, R., 2011. *Spondylus gaederopus* in Prehistoric Italy: Jewels from Neolithic and Copper Age Sites, in: Ifantidis, F., Nikolaïdou, M. (Eds.), *Spondylus* in Prehistory: New Data and Approaches. Contributions to the Archaeology of Shell Technologies. Archaeopress, pp. 25–37.
Borrello, M.A., Micheli, R., 2004. *Spondylus gaederopus*, gioiello dell’Europa preistorica. Preistoria Alpina 40, 71–82.
Boskey, A.L., Villarreal-Ramirez, E., 2016. Intrinsically disordered proteins and biomineralization. Matrix Biol. 52-54, 43–59.
Campbell, K.P., MacLennan, D.H., Jorgensen, A.O., 1983. Staining of the Ca2+-binding proteins, calsequestrin, calmodulin, troponin C, and S-100, with the cationic carbocyanine dye “Stains-all.” J. Biol. Chem. 258, 11267–11273.
Carter, J.G. (Ed.), 1990. Skeletal Biomineralization: Pattern, Processes, and Evolutionary Trends. Amer Geophysical Union.
Conchology, Inc [WWW Document], n.d. URL https://www.conchology.be/?t=1 (accessed 11.22.19).

Demarchi, B., Clements, E., Coltorti, M., van de Locht, R., Kröger, R., Penkman, K., Rose, J., 2015. Testing the effect of bleaching on the bivalve *Glycymeris*: A case study of amino acid geochronology on key Mediterranean raised beach deposits. Quat. Geochronol. 25, 49–65.

Demarchi, B., Hall, S., Roncal-Herrero, T., Freeman, C.L., Woolley, J., Crisp, M.K., Wilson, J., Fotakis, A., Fischer, R., Kessler, B.M., Rakownikow Jersie-Christensen, R., Olsen, J.V., Haile, J., Thomas, J., Marean, C.W., Parkington, J., Presslee, S., Lee-Thorp, J., Ditchfield, P., Hamilton, J.F., Ward, M.W., Wang, C.M., Shaw, M.D., Harrison, T., Domínguez-Rodrigo, M., MacPhee, R.D.E., Kwekason, A., Ecker, M., Kolska Horwitz, L., Chazan, M., Kröger, R., Thomas-Oates, J., Harding, J.H., Cappellini, E., Penkman, K., Collins, M.J., 2016. Protein sequences bound to mineral surfaces persist into deep time. Elife 5. https://doi.org/10.7554/eLife.17092

Demarchi, B., O'Connor, S., de Lima Ponzoni, A., de Almeida Rocha Ponzoni, R., Sheridan, A., Penkman, K., Hancock, Y., Wilson, J., 2014. An integrated approach to the taxonomic identification of prehistoric shell ornaments. PLoS One 9, e99839.

Du, J., Xu, G., Liu, C., Zhang, R., 2018. The role of phosphorylation and dephosphorylation of shell matrix proteins in shell formation: an *in vivo* and *in vitro* study. CrystEngComm 20, 3905–3916.

Espinosa, H.D., Rim, J.E., Barthelat, F., Buehler, M.J., 2009. Merger of structure and material in nacre and bone – Perspectives on *de novo* biomimetic materials. Prog. Mater Sci. 54, 1059–1100.

Evans, J.S., 2012. Aragonite-associated biomineralization proteins are disordered and contain interactive motifs. Bioinformatics 28, 3182–3185.

Evans, J.S., 2008. "Tuning in" to mollusk shell nacre- and prismatic-associated protein terminal sequences. Implications for biomineralization and the construction of high performance inorganic-organic composites. Chem. Rev. 108, 4455–4462.

Falini, G., Albeck, S., Weiner, S., Addadi, L., Others, 1996. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. Science-AAAS-Weekly Paper Edition 271, 67–69.

Feifarek, B.P., 1987. Spines and epibionts as antipredator defenses in the thorny oyster *Spondylus americanus* Hermann. J. Exp. Mar. Bio. Ecol. 105, 39–56.

Feng, D., Li, Q., Yu, H., Kong, L., Du, S., 2017. Identification of conserved proteins from diverse shell matrix proteome in *Crassostrea gigas*: characterization of genetic bases regulating shell formation. Sci. Rep. 7, 45754.

Finnemore, A., Cunha, P., Shean, T., Vignolini, S., Guldin, S., Oyen, M., Steiner, U., 2012. Biomimetic layer-by-layer assembly of artificial nacre. Nat. Commun. 3, 966.

Fuchigami, T., Sasaki, T., 2005. The shell structure of the Recent Patellogastropoda (Mollusca: Gastropoda). Paleontol. Res. 9, 143–168.

Gao, P., Liao, Z., Wang, X.-X., Bao, L.-F., Fan, M.-H., Li, X.-M., Wu, C.-W., Xia, S.-W., 2015. Layer-by-Layer Proteomic Analysis of *Mytilus galloprovincialis* Shell. PLoS One 10, e0133913.

Gosling, E., 2003. Bivalve Molluscs - Biology, Ecology and Culture, Fishing New Books. Blackwell Science, Oxford, UK.

Herlitze, I., Marie, B., Marin, F., Jackson, D.J., 2018. Molecular modularity and asymmetry of the molluscan mantle revealed by a gene expression atlas. Gigascience 7. https://doi.org/10.1093/gigascience/giy056

Ifantidis, F., Nikolaidou, M., 2011. Spondylus in Prehistory: New Data and Approaches: Contributions to the Archaeology of Shell Technologies.

Jackson, D.J., McDougall, C., Woodcroft, B., Moase, P., Rose, R.A., Kube, M., Reinhardt, R., Rokhsar, D.S., Montagnani, C., Joubert, C., Piquemal, D., Degnan, B.M., 2010. Parallel evolution of nacre building gene sets in molluscs. Mol. Biol. Evol. 27, 591–608.

Jamme, F., Kascakova, S., Villette, S., Allouche, F., Pallu, S., Rouam, V., Réfrégiers, M., 2013. Deep UV autofluorescence microscopy for cell biology and tissue histology. Biol. Cell 105, 277–288.

Kalmar, L., Homola, D., Varga, G., Tompa, P., 2012. Structural disorder in proteins brings

order to crystal growth in biomineralization. Bone 51, 528–534.
Kamat, S., Su, X., Ballarini, R., Heuer, A.H., 2000. Structural basis for the fracture toughness of the shell of the conch *Strombus gigas*. Nature 405, 1036–1040.
Kanold, J.M., Guichard, N., Immel, F., Plasseraud, L., Corneillat, M., Alcaraz, G., Brümmer, F., Marin, F., 2015. Spine and test skeletal matrices of the Mediterranean sea urchin *Arbacia lixula*--a comparative characterization of their sugar signature. FEBS J. 282, 1891–1905.
Knoll, A.H., 2003. Biomineralization and Evolutionary History. Reviews in Mineralogy and Geochemistry 54, 329–356.
Kocot, K.M., Aguilera, F., McDougall, C., Jackson, D.J., Degnan, B.M., 2016. Sea shell diversity and rapidly evolving secretomes: insights into the evolution of biomineralization. Front. Zool. 13, 23.
Logan, A., 1974. Morphology and Life Habits of the Recent Cementing Bivalve *Spondylus Americanus* Hermann from the Bermuda Platform. Bull. Mar. Sci. 24, 568–594.
Lowenstam, H.A., 1981. Minerals formed by organisms. Science 211, 1126–1131.
Luz, G.M., Mano, J.F., 2009. Biomimetic design of materials and biomaterials inspired by the structure of nacre. Philos. Trans. A Math. Phys. Eng. Sci. 367, 1587–1605.
Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., Lajoie, G., 2003. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun. Mass Spectrom. 17, 2337–2342.
Mann, K., Edsinger, E., 2014. The *Lottia gigantea* shell matrix proteome: re-analysis including MaxQuant iBAQ quantitation and phosphoproteome analysis. Proteome Sci. 12, 28.
Mao, J., Zhang, W., Zhang, X., Tian, Y., Wang, X., Hao, Z., Chang, Y., 2018. Transcriptional changes in the Japanese scallop (*Mizuhopecten yessoensis*) shellinfested by Polydora provide insights into the molecular mechanism of shell formation and immunomodulation. Sci. Rep. 8, 17664.
Marie, B., Arivalagan, J., Mathéron, L., Bolbach, G., Berland, S., Marie, A., Marin, F., 2017. Deep conservation of bivalve nacre proteins highlighted by shell matrix proteomics of the *Unionoida Elliptio complanata* and *Villosa lienosa*. J. R. Soc. Interface 14. https://doi.org/10.1098/rsif.2016.0846
Marie, B., Jackson, D.J., Ramos-Silva, P., Zanella-Cléon, I., Guichard, N., Marin, F., 2013a. The shell-forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-specific novelties. FEBS J. 280, 214–232.
Marie, B., Joubert, C., Tayalé, A., Zanella-Cléon, I., Belliard, C., Piquemal, D., Cochennec-Laureau, N., Marin, F., Gueguen, Y., Montagnani, C., 2012. Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. Proc. Natl. Acad. Sci. U. S. A. 109, 20986–20991.
Marie, B., Le Roy, N., Marie, A., Dubost, L., Marin, F., 2009a. Nacre Evolution: A Proteomic Approach. MRS Online Proceeding Library Archive 1187. https://doi.org/10.1557/PROC-1187-KK01-03
Marie, B., Le Roy, N., Zanella-Cléon, I., Becchi, M., Marin, F., 2011. Molecular evolution of mollusc shell proteins: insights from proteomic analysis of the edible mussel *Mytilus*. J. Mol. Evol. 72, 531–546.
Marie, B., Luquet, G., Pais De Barros, J.-P., Guichard, N., Morel, S., Alcaraz, G., Bollache, L., Marin, F., 2007. The shell matrix of the freshwater mussel *Unio pictorum* (Paleoheterodonta, Unionoida). Involvement of acidic polysaccharides from glycoproteins in nacre mineralization. FEBS J. 274, 2933–2945.
Marie, B., Marin, F., Marie, A., Bédouet, L., Dubost, L., Alcaraz, G., Milet, C., Luquet, G., 2009b. Evolution of nacre: biochemistry and proteomics of the shell organic matrix of the cephalopod *Nautilus macromphalus*. Chembiochem 10, 1495–1506.
Marie, B., Ramos-Silva, P., Marin, F., Marie, A., 2013b. Proteomics of CaCO3 biomineral-associated proteins: how to properly address their analysis. Proteomics 13, 3109–3116.
Marin, F., Amons, R., Guichard, N., Stigter, M., Hecker, A., Luquet, G., Layrolle, P., Alcaraz, G., Riondet, C., Westbroek, P., 2005. Caspartin and calprismin, two proteins of the shell

calcitic prisms of the Mediterranean fan mussel *Pinna nobilis*. J. Biol. Chem. 280, 33895–33908.
Marin, F., Bundeleva, I., Takeuchi, T., Immel, F., Medakovic, D., 2016. Organic matrices in metazoan calcium carbonate skeletons: Composition, functions, evolution. J. Struct. Biol. 196, 98–106.
Marin, F., Le Roy, N., Marie, B., 2012. The formation and mineralization of mollusk shell. Front. Biosci. 4, 1099–1125.
Marin, F., Le Roy, N., Marie, B., Ramos-Silva, P., Wolf, S., Benhamada, S., Guichard, N., Immel, F., 2014. Synthesis of calcium carbonate biological materials: how many proteins are needed? Key Eng. Mater. 614.
Marin, F., Luquet, G., Marie, B., Medakovic, D., 2007. Molluscan Shell Proteins: Primary Structure, Origin, and Evolution, in: Current Topics in Developmental Biology. Academic Press, pp. 209–276.
Marin, F., Marie, B., Hamada, S.B., Ramos-Silva, P., Le Roy, N., Guichard, N., Wolf, S.E., Montagnani, C., Joubert, C., Piquemal, D., Saulnier, D., Gueguen, Y., 2013. “Shellome”: Proteins involved in mollusk shell biomineralization-diversity, functions, in: Watabe, S., Maeyama, K., Nagasawa, H. (Eds.), . Terrapub Tokyo, p. 149:166.
Matsudaira, P., 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. 262, 10035–10038.
Matsumoto, M., 2003. Phylogenetic analysis of the subclass Pteriomorphia (Bivalvia) from mtDNA COI sequences. Mol. Phylogenet. Evol. 27, 429–440.
Matsumoto, M., Hayami, I., 2000. Phylogenetic analysis of the family Pectinidae (Bivalvia) based on mitochondrial cytochrome c oxidase subunit I. J. Molluscan Stud. 66, 477–488.
McDougall, C., Aguilera, F., Degnan, B.M., 2013. Rapid evolution of pearl oyster shell matrix proteins with repetitive, low-complexity domains. J. R. Soc. Interface 10, 20130041.
Mészáros, B., Erdos, G., Dosztányi, Z., 2018. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 46, W329–W337.
Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., Gough, J., Haft, D.R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A.P., Paysan-Lafosse, T., Pesseat, S., Potter, S.C., Qureshi, M.A., Rawlings, N.D., Redaschi, N., Richardson, L.J., Rivoire, C., Salazar, G.A., Sangrador-Vegas, A., Sigrist, C.J.A., Sillitoe, I., Sutton, G.G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Yong, S.-Y., Finn, R.D., 2019. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. 47, D351–D360.
Moshkovitz, S., 1971. Neolithic *Spondylus*. Nature.
NCBI [WWW Document], n.d. URL https://www.ncbi.nlm.nih.gov/protein/?term=%22molluscs%22%5Bporgn%3A__txid6447%5D (accessed 12.19.18).
Nishida, K., Ishimura, T., Suzuki, A., Sasaki, T., 2012. Seasonal changes in the shell microstructure of the bloody clam, *Scapharca broughtonii* (Mollusca: Bivalvia: Arcidae). Palaeogeogr. Palaeoclimatol. Palaeoecol. 363-364, 99–108.
Nishida, K., Suzuki, A., Isono, R., Hayashi, M., Watanabe, Y., Yamamoto, Y., Irie, T., Nojiri, Y., Mori, C., Sato, M., Sato, K., Sasaki, T., 2015. Thermal dependency of shell growth, microstructure, and stable isotopes in laboratory-reared *Scapharca broughtonii* (Mollusca: Bivalvia): Thermal control on shell microstructure. Geochem. Geophys. Geosyst. 16, 2395–2408.
Osuna-Mascaró, A., Cruz-Bustos, T., Benhamada, S., Guichard, N., Marie, B., Plasseraud, L., Corneillat, M., Alcaraz, G., Checa, A., Marin, F., 2014. The shell organic matrix of the crossed lamellar queen conch shell (*Strombus gigas*). Comp. Biochem. Physiol. B Biochem. Mol. Biol. 168, 76–85.
Palmer, A.R., 1992. Calcification in marine molluscs: how costly is it? Proc. Natl. Acad. Sci.

U. S. A. 89, 1379–1382.
Parker, A., Immel, F., Guichard, N., Broussard, C., Marin, F., 2015. Thermal stability of nacre proteins of the Polynesian oyster: a proteomic study. Key Engineer. Mater. 672, 222-231.
Paulsen, A.C., 1974. The Thorny Oyster and the Voice of God: *Spondylus* and *Strombus* in Andean Prehistory. Am. Antiq. 39, 597–607.
Pavat, C., Zanella-Cléon, I., Becchi, M., Medakovic, D., Luquet, G., Guichard, N., Alcaraz, G., Dommergues, J.-L., Serpentini, A., Lebel, J.-M., Marin, F., 2012. The shell matrix of the pulmonate land snail *Helix aspersa maxima*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 161, 303–314.
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, 2–25.
Pierini, F., Demarchi, B., Turner, J., Penkman, K., 2016. Pecten as a new substrate for IcPD dating: the Quaternary raised beaches in the Gulf of Corinth, Greece. Quat. Geochronol. 31, 40–52.
Pillsbury, J., 1996. The Thorny Oyster and the Origins of Empire: Implications of Recently Uncovered *Spondylus* Imagery from Chan Chan, Peru. Latin American Antiquity 7, 313–340.
Puslednik, L., Serb, J.M., 2008. Molecular phylogenetics of the Pectinidae (Mollusca: Bivalvia) and effect of increased taxon sampling and outgroup selection on tree topology. Mol. Phylogenet. Evol. 48, 1178–1188.
Ramos-Silva, P., Benhamada, S., Le Roy, N., Marie, B., Guichard, N., Zanella-Cléon, I., Plasseraud, L., Corneillat, M., Alcaraz, G., Kaandorp, J., Marin, F., 2012. Novel molluskan biomineralization proteins retrieved from proteomics: a case study with Upsalin. Chembiochem 13, 1067–1078.
Sakalauskaite, J., Andersen, S.H., Biagi, P., Borrello, M.A., Cocquerez, T., Colonese, A.C., Dal Bello, F., Girod, A., Heumüller, M., Koon, H., Mandili, G., Medana, C., Penkman, K.E., Plasseraud, L., Schlichtherle, H., Taylor, S., Tokarski, C., Thomas, J., Wilson, J., Marin, F., Demarchi, B., 2019. “Palaeoshellomics” reveals the use of freshwater mother-of-pearl in prehistory. Elife 8. https://doi.org/10.7554/eLife.45644
Samata, T., Ikeda, D., Kajikawa, A., Sato, H., Nogawa, C., Yamada, D., Yamazaki, R., Akiyama, T., 2008. A novel phosphorylated glycoprotein in the shell matrix of the oyster *Crassostrea nippona*: A novel acidic glycoprotein from the oyster shells. FEBS J. 275, 2977–2989.
Sarashina, I., Endo, K., 2006. Skeletal matrix proteins of invertebrate animals: Comparative analysis of their amino acid sequences. Paleontol. Res. 10, 311–336.
Sarashina, I., Endo, K., 2001. The complete primary structure of molluscan shell protein 1 (MSP-1), an acidic glycoprotein in the shell matrix of the scallop *Patinopecten yessoensis*. Mar. Biotechnol. 3, 362–369.
Shackleton, J., Elderfield, H., 1990. Strontium isotope dating of the source of Neolithic European *Spondylus* shell artefacts. Antiquity 64, 312–315.
Sleight, V.A., Marie, B., Jackson, D.J., Dyrynda, E.A., Marie, A., Clark, M.S., 2016. An Antarctic molluscan biomineralisation tool-kit. Sci. Rep. 6, 36978.
Smedley, G.D., Audino, J.A., Grula, C., Porath-Krause, A., Pairett, A.N., Alejandrino, A., Lacey, L., Masters, F., Duncan, P.F., Strong, E.E., Serb, J.M., 2019. Molecular phylogeny of the Pectinoidea (Bivalvia) indicates Propeamussiidae to be a non-monophyletic family with one clade sister to the scallops (Pectinidae). Mol. Phylogenet. Evol. 137, 293–299.
Suzuki, M., Iwashima, A., Tsutsui, N., Ohira, T., Kogure, T., Nagasawa, H., 2011. Identification and characterisation of a calcium carbonate-binding protein, blue mussel shell protein (BMSP), from the nacreous layer. Chembiochem 12, 2478–2487.
Takeuchi, T., Plasseraud, L., Ziegler-Devin, I., Brosse, N., Shinzato, C., Satoh, N., Marin, F., 2018. Biochemical characterization of the skeletal matrix of the massive coral, *Porites australiensis* - The saccharide moieties and their localization. J. Struct. Biol. 203, 219–

229.
Taylor, J.D., Kennedy, W.J., Hall, A., 1973. The shell structure and mineralogy of the Bivalvia. II. *Lucinacea-Clavagellacea*, Conclusions, 9. The British Museum (Natural History), London.
Taylor, J.D., Kennedy, W.J., Hall, A., 1969. The shell structure and mineralogy of Bivalvia: introduction, *Nuculacea-Trigonacea*, Bulletin of the British Museum (Natural History), Zoology, Supplement. The British Museum (Natural History), London.
Taylor, J.D., Reid, D.G., 1990. Shell microstructure and mineralogy of the Littorinidae: ecological and evolutionary significance. Hydrobiologia 193, 199–215.
Vizcaíno, J.A., Côté, R.G., Csordas, A., Dianes, J.A., Fabregat, A., Foster, J.M., Griss, J., Alpi, E., Birim, M., Contell, J., O'Kelly, G., Schoenegger, A., Ovelleiro, D., Pérez-Riverol, Y., Reisinger, F., Ríos, D., Wang, R., Hermjakob, H., 2013. The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res. 41, D1063–D1069.
Waller, T.R., 2006. Phylogeny of families in the Pectinoidea (Mollusca: Bivalvia): importance of the fossil record. Zool. J. Linn. Soc. 148, 313–342.
Wang, S., Zhang, J., Jiao, W., Li, J., Xun, X., Sun, Y., Guo, X., Huan, P., Dong, B., Zhang, L., Hu, X., Sun, X., Wang, J., Zhao, C., Wang, Y., Wang, D., Huang, X., Wang, R., Lv, J., Li, Y., Zhang, Z., Liu, B., Lu, W., Hui, Y., Liang, J., Zhou, Z., Hou, R., Li, X., Liu, Y., Li, H., Ning, X., Lin, Y., Zhao, L., Xing, Q., Dou, J., Li, Y., Mao, J., Guo, H., Dou, H., Li, T., Mu, C., Jiang, W., Fu, Q., Fu, X., Miao, Y., Liu, J., Yu, Q., Li, R., Liao, H., Li, X., Kong, Y., Jiang, Z., Chourrout, D., Li, R., Bao, Z., 2017. Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat Ecol Evol 1, 120.
Watabe, S., Maeyama, K., Nagasawa, H. (Eds.), 2011. Recent Advances in Pearl Research—Proceedings of the International Symposium on Pearl Research 2011. Terrapub, Tokyo, Japan.
Wilmot N. V., Barber D. J., Taylor J. D., Graham A. L., 1992. Electron microscopy of molluscan crossed-lamellar microstructure. Philos. Trans. R. Soc. Lond. B Biol. Sci. 337, 21–35.
Windler, A., 2017. From the Aegean Sea to the Parisian Basin: *Spondylus* shell exchange in Europe during the process of Neolithisation, in: Eisenach, P., Stöllner, T., Windler, A. (Eds.), The RITaK Conferences 2013–2014. Bochum, pp. 95–110.
Yarra, T., Gharbi, K., Blaxter, M., Peck, L.S., Clark, M.S., 2016. Characterization of the mantle transcriptome in bivalves: *Pecten maximus*, *Mytilus edulis* and *Crassostrea gigas*. Mar. Genomics 27, 9–15.
Yonezawa, M., Sakuda, S., Yoshimura, E., Suzuki, M., 2016. Molecular cloning and functional analysis of chitinases in the fresh water snail, *Lymnaea stagnalis*. J. Struct. Biol. 196, 107–118.
Yonge, M., 1973. Functional morphology with particular reference to hinge and ligament in *Spondylus* and *Plicatula* and a discussion on relations within the superfamily Pertinacea (Mollusca: Bivalvia). Philos. Trans. R. Soc. Lond. B Biol. Sci. 267, 173–208.
Yue, X., Zhang, S., Yu, J., Liu, B., 2019. Identification of a laccase gene involved in shell periostracal tanning of the clam *Meretrix petechialis*. Aquat. Biol. 28, 55–65.