The *Magellania venosa* Biomineralizing Proteome: A Window into Brachiopod Shell Evolution

Daniel J. Jackson1,†, Karlheinz Mann2,†, Vreni Häussermann3, Markus B. Schilhabel4, Carsten Lüter5, Erika Grieshaber6, Wolfgang Schmahl6, and Gert Wörheide6,7,*

1Department of Geobiology, Georg-August University of Göttingen, Germany
2Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Munich, Germany
3Escuela de Ciencias del Mar, Valparaíso, Facultad de Recursos Naturales, Universidad Católica de Valparaíso, Chile
4Institute of Clinical Molecular Biology, Christian-Albrechts-Universität Kiel, Germany
5Museum für Naturkunde, Leibniz-Institute for Evolution and Biodiversity Science, Berlin, Germany
6Department of Earth- and Environmental Sciences & GeoBioCenter, Ludwig-Maximilians-Universität München, Germany
7Bavarian State Collections of Palaeontology & Geology, München, Germany

*Corresponding author: E-mail: woerheide@lmu.de.

†These authors contributed equally to this work.

Accepted: April 21, 2015

Data deposition: All transcriptome nucleotide sequences/contigs associated with this manuscript have been deposited at the open data repository of the University of Munich, OpenDataLMU, hosted by the University Library of Munich (http://data.ub.uni-muenchen.de) and permanently accessible at http://dx.doi.org/10.5282/ubm/data.67. Conceptual translations of these contigs representing the protein sequences detected in the shells can be found in the supplementary material, Supplementary Material online. This project has been deposited at the European Nucleotide Archive under the accession http://www.ebi.ac.uk/ena/data/view/PRJEB8161.

Abstract

Brachiopods are a lineage of invertebrates well known for the breadth and depth of their fossil record. Although the quality of this fossil record attracts the attention of paleontologists, geochemists, and paleoclimatologists, modern day brachiopods are also of interest to evolutionary biologists due to their potential to address a variety of questions ranging from developmental biology to biomineralization. The brachiopod shell is a composite material primarily composed of either calcite or calcium phosphate in close association with proteins and polysaccharides which give these composite structures their material properties. The information content of these biomolecules, sequestered within the shell during its construction, has the potential to inform hypotheses focused on describing how brachiopod shell formation evolved. Here, using high throughput proteomic approaches and next generation sequencing, we have surveyed and characterized the first shell-proteome and shell-forming transcriptome of any brachiopod, the South American *Magellania venosa* (Rhynchonelliformea: Terebratulida). We find that the seven most abundant proteins present in the shell are unique to *M. venosa*, but that these proteins display biochemical features found in other metazoan biomineralization proteins. We can also detect some *M. venosa* proteins that display significant sequence similarity to other metazoan biomineralization proteins, suggesting that some elements of the brachiopod shell-forming proteome are deeply evolutionarily conserved. We also employed a variety of preparation methods to isolate shell proteins and find that in comparison to the shells of other spiralian invertebrates (such as mollusks) the shell ultrastructure of *M. venosa* may explain the effects these preparation strategies have on our results.

Key words: Brachiopoda, biomineralization, proteome, transcriptome, evolution, *Magellania*.

Introduction

For much of the Phanerozoic, animal life has constructed an amazing diversity of mineralized structures to fulfill a host of biological functions (reviewed in Simkiss and Wilbur 1989). The products of these biomineralization processes, precisely coordinated by macromolecules such as proteins and carbohydrates, range from individual functional elements, for example, magneto-sensors, teeth, armor (Lowenstam and...
Weiner 1989), to large scale geological structures, for example, coral reefs. Ever since the onset of metazoan biomineralization at the dawn of the Phanerozoic (~550 Ma; Knoll 2003), their biomineralized structures have most commonly been constructed from one of three minerals: Calcium carbonate (CaCO₃), for example, most brachiopod shells, echinoderm spines and tests, molluscan shells, coral skeletons, and sponge spicules; calcium phosphate (CaPO₄), for example, some brachiopod shells and vertebrate bones; or silica (SiO₂), for example, primarily sponge spicules. Brachiopods were among the first animals to biomineralize in the Cambrian (Knoll 2003), and achieved a high diversity in the Paleozoic (4,200 described genera, all of which constructed biomineralized shells; reviewed in Taylor et al. 2010). Although most brachiopod taxa are now extinct, this historic diversity makes them a significant invertebrate index fossil group, providing a continuous deep-time record of a prevalent biocalcification system (e.g., Rudwick 1970). Morphological (Carlson 1995) as well as recent molecular data (Cohen and Weydmann 2005; Sperling et al. 2011; Cohen 2013) suggest that Brachiopoda is monophyletic either to the inclusion of Phoronida (Cohen and Weydmann 2005) or with robustly supported subphyla Linguliformea and Craniformea (classically referred to as “Inarticulata”) as a sister group to the Rhynchonelliformea (classically referred to as “Articulata”) (Sperling et al. 2011; see fig. 1).

The divergence between these sister groups has been estimated by molecular clock analyses at 547 Ma (Sperling et al. 2011), and importantly, this deep divergence is reflected in the different biomineralizing systems that each group employs; although all brachiopod shells are inorganic/organic composites, within the Rhynchonelliformea and Craniformea, calcite (CaCO₃) dominates the organic content on a w/w basis (~98%), whereas Linguliformea shells typically possess calcium phosphate in approximately equal ratio to the organic content (~50%, i.e., “organophosphatic” shells; reviewed in Taylor et al. 2010). Moreover, the “organophosphatic” shells of the Linguliformea display a hierarchically laminated architecture (Williams 1989; Merkel et al. 2007; Schmahl et al. 2008; Merkel et al. 2009), whereas Rhynchonelliformea shells display diverse types of microstructures based on biopolymer reinforced calcite (Goetz et al. 2009, 2011; Schmahl et al. 2012). The dichotomy of shell mineralogies (amongst extinct and extant brachiopods (e.g., Jope 1966a, 1966b). More recent studies by Cusack et al. (1992) reported the N-terminal sequence of a red chromoprotein (ICP-1) from three species of brachiopods, and Gaspard et al. (2008) described the microstructures and general nature of the glycosylated shell-forming proteomes of five rhynchonelliform brachiopods. Geochemical and isotopic techniques have also been applied to brachiopod shells as proxies for climate reconstruction (Curry and Fallick 2002; Brand 2003; Lee et al. 2004; Parkinson et al. 2005; Cusack et al. 2012). Despite the interest from so many disciplines in the organic components of brachiopod shells and how they regulate shell formation, no study has yet conducted a broad scale survey of the proteins that guide brachiopod shell-formation, and no report to our knowledge describes the full-length sequence of any brachiopod biomineralizing protein.

Here, we have surveyed both the shell-forming transcriptome and CaCO₃ shell-proteome of Magellania venosa, a rhynchonelliform brachiopod, the first biomineralizing data sets of their kind for any brachiopod. This represents a major step forward in terms of our molecular understanding of how an extant brachiopod constructs a calcified shell. We have compared these data with the relatively extensive shell-forming proteome data sets now available from several mollusks and more distantly related biocalcifying metazoan taxa (corals and sea urchins), providing further insight into how brachiopod shell-formation has evolved. These data will also be invaluable for future comparative studies aimed at revealing the molecular mechanisms that regulate the deposition of calcitic versus phosphatic biominerals.

**Materials and Methods**

More than 20 specimens of *M. venosa* were collected by SCUBA diving at the jetty in Huinay (42°22'29"S, 72°25'41.58"W) in the Comau Fjord, Southern Chilean fjord region between 2010 and 2013 (6 samples on November 10, 2010; 3 samples on December 11, 2011; 10 samples on January 11, 2013). The shells of whole animals from the first batch were cracked open and preserved in RNAlater overnight at 4°C and then stored at −20°C until shipment to Germany. The samples of the second batch were cleaned of soft tissue and the shells air-dried. The samples of the third batch were preserved in absolute ethanol. Scanning electron micrographs were taken from bleached and cracked shells using a FEI Quanta 200F Field Emission Scanning Electron Microscope (SEM) with an accelerator voltage of 15.3 kV at the Zoologische Staatssammlung München.

**Organic Matrix Preparation**

Shells were cleaned in sodium hypochlorite solution (6–14% active chlorine; Merck, Darmstadt, Germany) for 2 h at room temperature with 5 min ultra sonication and change of
hypochlorite solution every 30 min. Cleaned shells were washed with deionized water and dried. One set of three dorsal (D) and three ventral (V) shells was demineralized for matrix extraction. A second set of shells was crushed using mortar and pestle to obtain a homogeneous powder with no visible shell pieces. This powder was treated with approximately 10 volumes of sodium hypochlorite for another 24 h at 4 °C on a roller mixer. Shell particles were collected by centrifugation for 10 min at 3,000 g, washed three times with deionized water, and air-dried. Demineralization of both types of shell preparation was done in 50% acetic acid (20 ml/g of shell) over night at 4–6 °C, and the resulting suspension was dialyzed (Spectra/Por 6 dialysis membrane, molecular weight cut-off 2,000; Spectrum Europe, Breda, The Netherlands) successively against 3 × 11 of 10% acetic acid and 3 × 11 of 5% acetic acid at 4–6°C and lyophilized. Matrices were analyzed by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) using precast 4–12% Novex Bis-Tris gels in MES buffer using reagents and protocols supplied by the manufacturer (Invitrogen, Carlsbad, CA), except that 1% β-mercaptoethanol was used as a reducing agent in the sample buffer. Samples were suspended in sample buffer (200 μg/30 μl), heated to 70 °C for 10 min and centrifuged for 5 min at 15,800 g in a 5415D Eppendorf centrifuge to remove sample buffer-insoluble material. The molecular weight standard was Novex Sharp prestained (Invitrogen).

Peptide Preparation

Reduction, carbamidomethylation, and enzymatic cleavage of matrix proteins were performed using a modification of the FASP (filter-aided sample preparation) method (Wisniewski et al. 2009) as outlined below. Aliquots of 200 μg of matrix were suspended in 300 μl of 0.1 M Tris, pH 8, containing 6 M guanidine hydrochloride, and 0.01 M dithiothreitol (DTT). The mixture was heated to 56 °C for 60 min, cooled to room temperature, and centrifuged at 15,800 g in an Eppendorf benchtop centrifuge S415D for 15 min. The supernatant was loaded into an Amicon Ultra 0.5 ml 30 K filter device (Millipore; Tullagreen, Ireland). DTT was removed by centrifugation at 15,800 g for 15 min and washing with 2 × 1 vol of the same buffer. Carbamidomethylation was done in the device using 0.1 M Tris buffer, pH 8, containing 6 M guanidine hydrochloride and 0.05 mM iodoacetamide and incubation for 45 min in the dark. Carbamidomethylated proteins were washed with 0.05 M ammonium hydrogen carbonate buffer, pH 8, containing 2 M urea, and centrifuged as before. Trypsin (2 μg, Sequencing grade, modified; Promega, Madison, WI) was added in 40 μl of 0.05 M ammonium hydrogen carbonate buffer containing 2 M urea and the devices were incubated at 37 °C for 16 h. Peptides were collected by centrifugation and the filters were washed twice with 40 μl of 0.05 M ammonium hydrogen carbonate buffer and twice with 1% trifluoroacetic acid in 5% acetonitrile. The acidic peptide solution (pH 1–2) was applied to C18 Stage Tips (Rappsilber et al. 2007) and the eluted peptides were vacuum-dried in an Eppendorf concentrator.

Liquid Chromatography–Mass Spectrometry and Data Analysis

Peptide mixtures were analyzed by on-line nanoflow liquid chromatography (LC) using the EASY-nLC 1000 system (Proxeon Biosystems, Odense, Denmark, now part of Thermo Fisher Scientific) with 20-cm capillary columns of an internal diameter of 75 μm filled with 1.8 μm ReproSili-Pur C18-AQ resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Peptides were eluted with a linear gradient from 5% to 30% buffer B (80% acetonitrile in 0.1% formic acid) in 90 min, 30–60% B in 5 min, and 60–95% B in 5 min at a flow rate of 250 nl/min and a temperature of 40 °C. The eluate was electro-sprayied into an Orbitrap Q Exactive (Thermo Fisher Scientific, Bremen, Germany) using a Proxeon nanoelectrospray ion source. The instrument was operated in an HCD top 10 mode essentially as described (Michalski et al. 2011). The resolution was 70,000 for full scans and 7,500 for fragments (both specified at m/z 400). Ion target values were 1e6 and 5e4 ms, respectively. Exclusion time was 20 s. Sample runs were monitored using the SprayQC quality monitoring system (Scheltema and Mann 2012). Raw files were processed using the Andromeda search engine-based version 1.4.1.12 of MaxQuant (http://www.maxquant.org/, last accessed May 7, 2015) with enabled second peptide, iBAQ, and match between runs (match time window 0.5 min; alignment time window 20 min) options (Cox and Mann 2008; Cox et al. 2011).

In order to generate an M. venosa peptide database against which the LC–mass spectrometry (MS) data could be searched, we sequenced the mantle transcriptome of M. venosa using the Illumina HiSeq platform. Briefly, total RNA from the mantle tissue of an adult individual was isolated using the Qiagen RNeasy Plant Kit according to the manufacturer’s instructions. The resulting RNA was then processed by the sequencing center at the IKMB at the University of Kiel (Germany), where a 230–430 bp paired end insert TrueSeq RNA library was constructed and sequenced for 101 bases from both ends. More than 370 million reads were returned, quality trimmed, and assembled using the CLC Genomics Workbench 6.5. A total of 27,897 contigs with an N50 of 1,803 bp (average 1,319 bp) were assembled and subsequently translated in all six reading frames. These data were combined with the sequences of common contaminants, such as human keratins and mammalian cytoskeletal proteins. Carbamidomethylation was set as fixed modification. Variable modifications were oxidation (M), N-acetyl (protein), pyro-Glu/Gln (N-term), and phospho (STY) (Sharma et al. 2014). One batch of replicates was also
analyzed tentatively adding hydroxyproline as a variable modification, but no additional collagen peptides containing hydroxyproline were detected. The initial mass tolerance was 7 ppm for full scans and 20 ppm for tandem mass spectrometry. Two missed cleavages were allowed and the minimal length required for a peptide was seven amino acids. Maximal false discovery rate for peptide spectral match, proteins, and site was set to 0.01. The minimal score for peptides was 60 and the minimal delta score for modified peptides was 17. Minimal requirements for final acceptance of identifications were the presence of the protein in at least four biological replicates of 12 (V1–6, D1–6) with at least two sequence-unique peptides in two of the four replicate groups (D1–3, V1–3, D4–6, V4–6). Each of the 12 biological replicates was measured with five technical replicates. Identifications with one sequence-unique peptide were exceptionally accepted if the same protein was identified with more peptides in other biological replicates. Identifications with two sequence-unique peptides were accepted after positive validation using the MaxQuant Expert System software (Neuhaus et al. 2012) considering the assignment of major peaks, occurrence of uninterrupted y- or b-ion series of at least four consecutive amino acids, preferred cleavages N-terminal to proline bonds, the possible presence of a2/b2 ion pairs and immonium ions, and assignment of major fragment peaks. The IBQ (intensity-based absolute quantification) (Schwanhausser et al. 2011) option of MaxQuant was used to calculate, based on the sum of peak intensities, the approximate share of each protein in the total proteome, including identifications which were not finally accepted. This enabled us to discern between minor and major proteins.

Sequence similarity searches were performed with FASTA (http://www.ebi.ac.uk/Tools/ss/ss fasta/, last accessed May 7, 2015) (Goujon et al. 2010) against current releases of the Uniprot Knowledgebase. Other bioinformatics tools used were Clustal Omega for sequence alignments (http://www.ebi.ac.uk/Tools/msa/clustalo/, last accessed May 7, 2015) (Sievers et al. 2011), InterPro (http://www.ebi.ac.uk/interpro, last accessed May 7, 2015) (Hunter et al. 2012) for domain predictions, SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/, last accessed May 7, 2015) (Petersen et al. 2011) for signal sequence prediction, and TMHMM (http://www.cbs.dtu.dk/services/TMHMM, last accessed May 7, 2015) (Kristensen et al. 2011) for prediction of transmembrane helices. Amino acid composition and theoretical pl were determined using the ProtParam tool provided by the Expasy server (http://web.expasy.org/protparam/, last accessed May 7, 2015) (Gasteiger et al. 2005). Intrinsically disordered protein structure was predicted using IUPred (http://iupred.enzim.hu/, last accessed May 7, 2015) (Dosztáinyi et al. 2005) and PrDOS (http://prdos.hgc.jp/cgi-bin/top.cgi, last accessed May 7, 2015) (Ishida and Kinoshita 2007). Subcellular location predictions were based on sequence similarities to known proteins, domain predictions, signal sequence predictions, and transmembrane segment predictions. Sequence-based comparisons were performed using Circleto (Darzentas 2010) and Circos (Krzywinski et al. 2009). Tandem repeats were detected using XStream (Newman and Cooper 2007).

Results and Discussion

The Proteome of M. venosa Shells Treated with Hypochlorite for 2 h

*Maggellania venosa* shells treated with sodium hypochlorite solution for 2 h became soft and crumbly. This was most probably due to some extra-crystalline matrix degradation by the 2-h hypochlorite treatment (see next section). Only that part of the shell around the hinge region remained hard and difficult to crush. Yields of organic matrix from three ventral and three dorsal shells treated with hypochlorite for 2 h were 1.9 ± 0.1 mg of matrix/g dorsal shell (0.19%) and 2.7 ± 0.7 mg/g of ventral shell (0.27%). Comparisons of PACE patterns of the organic matrices did not indicate any major differences between the proteomes of ventral and dorsal shells (fig. 2, lanes D2h, and V2h).

Five technical replicates from each of three dorsal and three ventral shells treated with hypochlorite for 2 h were analyzed in two groups (dorsal and ventral). Searching the LC-MS raw files against the *M. venosa* sequence database resulted in 496 identifications from the dorsal set of replicates (supplementary file S1, Supplementary Material online) and 591 from ventral set of replicates (supplementary file S2, Supplementary Material online), with an overlap of approximately 70% of each data set. Differences between dorsal and ventral shells concerned minor proteins only, and may be related to the higher protein content of ventral shells (see above) and to experimental variation. Peptide sequences and relevant parameters are shown in supplementary files S3 and S4, Supplementary Material online. After revising the list of identifications according to the rules described above (see Liquid Chromatography–Mass Spectrometry and Data Analysis) and the grouping together of entries suggested by FASTA searches to belong to different regions of the same protein, we obtained 317 proteins/protein groups (supplementary file S5: table S1, Supplementary Material online). Seventeen proteins were identified in ventral samples only and eight were identified in dorsal samples only. Such minor qualitative differences are likely to be due to experimental variation. The higher number of proteins in ventral samples may be related to the higher protein content of the ventral shells (see above).

In contrast to the recently analyzed shell proteomes of the terrestrial snail *Cepaea nemoralis* (Mann and Jackson 2014) and the marine snail *Lottia gigantea* (Mann and Edsinger 2014), the proteome of *M. venosa* contained several intracellular proteins with high abundance (supplementary file S5, Supplementary Material online). The most abundant of the intracellular proteins was actin. Because of the very high
sequence conservation among actins from different sources, this cytoskeletal protein shared almost all of its peptides with common mammalian contaminants (one sequence-unique peptide of a total of 16). Therefore, it was difficult to estimate the true abundance of *M. venosa* actin. The same was true for several other cytoskeletal proteins such as tubulins. However, because such extraordinarily high abundances were not found in our previous studies of molluscan shell proteomes (Mann et al. 1991), our preferred interpretation is that these peptides in fact originated from common mammalian contaminants (one sequence-unique peptide) and were not the true abundances.

The formation and growth of calcite brachiopod shells in rhynchonelliforms has been described by Williams (1997) as the result of the migration of outer mantle epithelial cells “during their secretion of succeeding biomineral components” (Williams 1997). What has been described as the protein boundary of secondary shell fibers, for example, in *Notoaria nigricans*, may represent organic remnants of this migration process containing, among other components, a considerably large fraction of actin which originally helped the cells to move along their secretional track. Whether this represents extracellular actin as described, for example, in the matrix of mouse smooth muscle cells (Accinni et al. 1983) remains to be investigated. However, the high concentration of actin we detect in *M. venosa* shell material may be the first molecular evidence that Williams’ hypothesis of shell secretion in the Rhynchonelliformea is indeed correct. Interestingly, other metazoans constructing calcified skeletons have also been found to contain actin in their skeletal matrices even after rigorous washing of powdered skeleton, for example, the stony coral *Stylophora pistillata* (Drake et al. 2013a, 2013b; Mass et al. 2014), and the sclerites of the calcitic octocoral *Sinularia* sp. (Rahman et al. 2013), indicating that actin may genuinely have some biomineralization-related function in these organisms.

### The Proteome of *M. venosa* Shell Powder Treated with Hypochlorite for Another 24 h

Despite the apparent accessibility of *M. venosa*’s organic intercrystalline matrix and the danger of losing it during more aggressive washing treatments prior to matrix extraction, the high abundance of intracellular proteins described above leads us to attempt a more rigorous washing procedure. To this end, shells already treated with sodium hypochlorite for 2 h were reduced to a homogeneous powder using mortar and pestle and treated with hypochlorite solution for another 24 h with constant mixing on a roller mixer at 4°C. This treatment is equivalent to half of the incubation time previously used for *Terebratulina septentrionalis* (another brachiopod of the order Terebratulida) shells to destroy the intercrystalline matrix to such an extent that simple shaking was enough to obtain free calcite fibers (Collins et al. 1988). This more aggressive protocol was similar to that used for several other brachiopod shells reported previously (Curry et al. 1991). The yield of organic matrix from this type of preparation was 1.0 ± 0.3 mg/g (0.1%) from both dorsal and ventral shells. Again, PAGE analysis (fig. 2, lanes D24h and V24h) did not reveal any major differences between dorsal and ventral shells. However, there was a distinct difference between the two sodium hypochlorite treatments, likely indicating a loss of extracrystalline matrix and an enrichment of the intracrystalline matrix (fig. 2). Several predominant protein bands disappeared, whereas others, especially three proteins with apparent molecular weights of 110, 50, and 20 kDa were significantly enriched. This three-band pattern was reminiscent to previously published patterns derived from the intracrystalline matrix of the terebratulid brachiopods of the genera *Neothyris* (Curry et al. 1991), although the molecular weights were not the same (47, 16, and 6.5 kDa), and *Calloria* (Collins et al. 1991). The latter showed major protein bands between a molecular weight of 30–110 kDa when reacted with antishell matrix antiserum.

Altogether we identified 243 proteins/protein groups from samples processed with this more aggressive sodium hypochlorite treatment (supplementary file S5, Supplementary Material online). This is in contrast to the 317 proteins/protein groups identified using the milder sodium hypochlorite treatment. Additional data about proteins and peptides including identifications not finally accepted are contained in supplementary files S6–S9, Supplementary Material online. Ten and eight minor proteins were found only in dorsal and ventral samples, respectively. The overlap between samples
treated for 2 or 24 h with sodium hypochlorite was 205 proteins. Thirty-eight minor proteins appeared in the more rigorously cleaned samples only. This was most probably due to the depletion of some major proteins as compared with the less rigorously cleaned samples. Supplementary file S5 (table S1), Supplementary Material online, shows that the number and abundance of common intracellular proteins decrease dramatically with the more aggressive sodium hypochlorite treatment (e.g., β-tubulin, actin). However, several major proteins that could not be matched to known intracellular proteins were significantly reduced or eliminated with the more aggressive sodium hypochlorite treatment. These proteins may be major components of the extracrystalline matrix (discussed below), in agreement with previous reports indicating that prolonged treatment of brachiopod shells with hypochlorite removes the electron microscopically visible extracrystalline matrix and destroys the cohesion between calcified elements of the shell (Collins et al. 1988).

The Major M. venosa Shell Matrix Proteins

The 66 major shell proteins (arbitrarily defined here as proteins occurring with an average percentage of 0.1 in either D1–3 and V1–3 [2 h] or D4–6 and V4–6 [24 h] or both) are presented in table 1. The most abundant protein (R20074237) was a novel protein with no sequence similarity to any sequences in SwissProt, GenBank, GenBank nr, or with any recognizable Pfam domains. This protein possessed a signal sequence (suggesting that it is an extracellular protein) and an unusually high proportion of acidic (Asp—12%) amino acids. In addition, this protein shared no sequence similarity with any of the 765 proteins in our targeted comparisons against mollusk, sea urchin, or coral biomineralizing proteomes (see fig. 3). Although the discovery of such novel proteins that are likely to play prominent roles in brachiopod shell formation is exciting, such results are tempered by the fact that an understanding of the mechanisms of biomineralization formation is exciting, such results are tempered by the fact that an understanding of the mechanisms of biomineralization (discussed below), in agreement with previous reports indicating that prolonged treatment of brachiopod shells with hypochlorite removes the electron microscopically visible extracrystalline matrix and destroys the cohesion between calcified elements of the shell (Collins et al. 1988).

Other proteins relatively abundant in the shell of M. venosa that shared similarity with other metazoan biomineralizing proteins were F20107906 and R10050836 which displayed similarity to oyster, Cepaea nemoralis and Lottia gigantea proteins and possess animal heme-dependent peroxidase domains. Peroxidases are well known to be present in biomineralizing tissues in mollusks (Timmermans 1969; Hohagen and Jackson 2013), and are generally thought to function in the cross-linking of proteins present in molluscan periostracum (Knipprath 1977). Another brachiopod shell-forming protein with similarity to proteins in all three oysters surveyed was R10023620. These proteins all possess a tyrosinase domain, an enzyme that catalyzes the production of melanin and is known to be involved in biomineralization in bivalves (Zhang et al. 2006; Clark et al. 2010; Yu et al. 2014).

Because many biomineralizing proteins are known to possess repetitive low-complexity domains (Jackson et al. 2010; Marie et al. 2013; Le-Roy et al. 2014), we searched for proteins with tandem repeats in the collection of 66 major M. venosa shell proteins. A total of 15 proteins were identified by XStream to possess tandem repeats and low-complexity domains (supplementary file S10, Supplementary Material online). Of these 15, three are apparently unique to M. venosa (i.e., they possess no sequence similarity to any proteins or nucleotide sequences in GenBank’s nr and nt databases) and are putatively full length. One of these proteins (R10063526) has a striking architecture possessing a signal sequence, extraordinarily high proportions of glycine (19.4%) and serine (22.5%) and phosphorylated residues (table 1). Interestingly, Lustrin-A, one of the first molluscan biocalcification proteins to have its primary structure fully
Table 1
Major Proteins of the *Magellania venosa* Shell Matrix

| Contig ID | Protein ID | BLAST Hits | Interpro Domains; Transmembrane Helices; Disordered Domains; Phosphorylation | % of Total Identified Dorsal Proteome 2/24 h | % of Total Identified Ventral Proteome 2/24 h | pI | SS | Anomalous Amino Acid Contents |
|-----------|------------|------------|-----------------------------------------------------------------------------|---------------------------------------------|---------------------------------------------|-----|-----|--------------------------------|
| 1237      | R20074237  | Uncharacterized | 2 TMH; DI                                                                     | 13.8/8.4                                    | 11.6/8.0                                    | 4.34| Y  | Asp 12.1%                        |
| 10941     | F20065642  | Uncharacterized | —                                                                           | 7.8/13.4                                    | 7.9/13.6                                    | 8.88| N  | —                              |
| 13662     | R30081972  | Uncharacterized | —                                                                           | 9.6/8.3                                     | 5.0/8.3                                     | 9.34| N  | —                              |
| 8496      | F20050972  | Uncharacterized | —                                                                           | 7.8/14.4                                    | 6.4/13.2                                    | 5.7 | Y  | —                              |
| 1617      | R20097001  | Uncharacterized | TIMP like OB fold Phosphorylation: S<sub>174</sub> (90%)                     | 6.8/9.8                                     | 6.1/10.4                                    | 9.96| N  | —                              |
| 9860      | F30059157  | Uncharacterized | Di                                                                           | 4.8/0.6                                     | 4.4/0.6                                     | 9.71| N  | Gly 29.3%                       |
| 10182     | R10061090  | Uncharacterized | Di                                                                           | 3.2/4.5                                     | 5.0/4.4                                     | 4.37| Y  | Asp 15.4%                       |
| 14565     | R20087389  | Similar to mesenchyme-specific cell surface glycoprotein/calium and integrin-binding family member 2 | Di; WD40/YVTN repeat like; DI | 4.5/5.3                                     | 3.1/5.2                                     | 9.19| N  | —                              |
| 13982     | R30083352  | Uncharacterized | —                                                                           | 2.9/3.2                                     | 3.5/3.0                                     | 11.38| Y | —                              |
| 481       | R20002885  | Uncharacterized | —                                                                           | 2.801                                       | 3.00.1                                      | 9.88 | N | Gly 17.3%                       |
| 10544     | R20063263  | Uncharacterized | —                                                                           | 2.4/2.7                                     | 2.9/2.5                                     | 6.64 | N | Ser 10.4%                       |
| 21084<sup>a</sup> | R10165002 | Uncharacterized | FAD linked oxidase N term                                                    | 2.1/~0.01                                   | 2.9/0.0                                     | 10.35| Y | —                              |
| 6089<sup>b</sup> | F10036529 | Similar to complement receptor type 1 | Sushi/SCR/CCP (B<sub>2</sub>)                                                | 2.4/5.6                                     | 2.3/5.8                                     | 6.28 | Y | —                              |
| 8840      | F20053036  | Uncharacterized | 2 TMH                                                                        | 2.0/2.2                                     | 2.6/1.9                                     | 9.63 | Y | —                              |
| 25732     | R10154390  | Uncharacterized | —                                                                           | 1.8/1.4                                     | 1.2/1.2                                     | 9.85 | Y | —                              |
| 17985<sup>c</sup> | F20107906 | Similar to peroxidin | Haem peroxidase, TSP1 repeat                                                  | 1.3/0.0                                     | 1.2/0.0                                     | 9.12 | N | —                              |
| 17302     | R20103811  | Uncharacterized | 2 TMH                                                                        | 1.0/0.1                                     | 1.1/0.1                                     | 6.01 | N | Ala 11.8%                       |
| 16240<sup>d</sup> | R10097438 | Uncharacterized | FAD linked oxidase N term, berberine; 3 TMH                                  | 0.5/0.0                                     | 1.3/0.0                                     | 9.64 | N | —                              |
| 8882      | F20053288  | Uncharacterized | 1 TMH                                                                        | 0.8/0.6                                     | 1.0/0.7                                     | 9.99 | Y | Gly 14.5%                       |
| 10588     | R10063526  | Uncharacterized | Di; phosphorylation: S<sub>272</sub> (1%), S<sub>277</sub> (9%), doubly phosphorylated peptide (69%) | 0.6/0.0                                     | 1.2/0.0                                     | 4.48 | Y | Gly 19.4%, Ser 22.6%           |
| 17072     | R30102432  | Uncharacterized | —                                                                           | 0.7/0.6                                     | 0.8/0.5                                     | 9.14 | Y | —                              |
| 1404      | R10008422  | Uncharacterized | Di                                                                           | 0.8/1.1                                     | 0.6/1.3                                     | 9.49 | N | Ser 10.6%                       |
| 3349      | F10020089  | Uncharacterized | Di                                                                           | 0.6/0.0                                     | 0.8/0.0                                     | 4.29 | N | Glu 19.5%, Ser 11.9%           |
| 25890     | F30155337  | Uncharacterized | —                                                                           | 0.6/0.5                                     | 0.8/0.4                                     | 12.18| N | —                              |
| 12037     | F10072217  | Similar to collagen s5 (VII), nontriple helical region | Di | 0.7/0.2                                     | 0.7/0.2                                     | 10.01| Y | Gly 11.5%                       |
| 1290      | R20007740  | Uncharacterized | VWA                                                                          | 0.7/0.0                                     | 0.4/0.0                                     | 10.14| N | —                              |
| 24592     | R20147551  | Uncharacterized | —                                                                           | 0.4/0.0                                     | 0.6/0.0                                     | 9.6  | N | Lys 11.0%                       |
| 13443     | F30098565  | Uncharacterized | 1 TMH                                                                        | 0.5/0.2                                     | 0.5/0.1                                     | 9.59 | Y | —                              |
| 3937      | R10023620  | Similar to CRE-TYR-3 | Tyrosinase                                                                      | 0.3/0.0                                     | 0.6/0.0                                     | 9.06 | N | Cys 9.6%, Gly 10.9%           |
| 3968      | F10023803  | Similar to mesenchyme-specific cell surface glycoprotein | WD40/YVTN repeat like/ cytochrome cd1 (cyt cd1) nitrite reductase like, haem d1; Di | 0.4/1.5                                     | 0.4/1.7                                     | 6.47 | N | —                              |
| 6648      | R30039888  | Similar to lactadherin | Galactose binding like/coagulation factor 5/6.C                              | 0.3/0.8                                     | 0.6/0.7                                     | 9.41 | N | —                              |
| 12090     | F10072535  | Uncharacterized | Di                                                                           | 0.4/1.0                                     | 0.3/1.2                                     | 8.85 | N | Gly 10.2%                       |
| 14281     | F10085681  | Similar to hemagglutinin family protein | Di | 0.5/0.4                                     | 0.3/0.4                                     | 5.12 | N | Ser 11.3%                       |
| 23379     | F20140270  | Uncharacterized | —                                                                           | 0.5/1.9                                     | 0.3/1.9                                     | 10.42| N | Gly 12.8%                       |
| 10375     | R10062248  | Uncharacterized | 3 TMH; Di                                                                     | 0.5/0.7                                     | 0.3/0.8                                     | 10.91| Y | Ala 11.2%                       |
Table 1 Continued

| Contig ID | Protein ID | BLAST Hits | Interpro Domains; Transmembrane Helices; Disordered Domains; Phosphorylation | % of Total Identified Dorsal Proteome 2/24 h | % of Total Identified Ventral Proteome 2/24 h | pl | SS | Anomalous Amino Acid Contents |
|-----------|------------|------------|-----------------------------------------------------------------------------|---------------------------------------------|---------------------------------------------|----|----|-------------------------------|
| 22890     | R30137340  | Uncharacterized | ConA-like lectin | 0.10.0 | 0.60.0 | 9.72 | Y | — |
| 20588     | R20123527  | Uncharacterized | 1 TMH | 0.30.3 | 0.40.3 | 9.47 | N | — |
| 26985     | F30161907  | Uncharacterized | — | 0.30.0 | 0.30.0 | 9 | N | Ser 13.7% |
| 14828*    | R20088967  | Similar to complement receptor type 1 | — | 0.30.8 | 0.20.9 | 6.85 | Y | Gly 10.4% |
| 16143     | F30096855  | Uncharacterized | — | 0.40.4 | 0.30.4 | 9.02 | Y | — |
| 4638      | R30027828  | Uncharacterized | DOMON, SEA | 0.30.3 | 0.10.5 | 4.84 | N | — |
| 8473f     | R10050836  | Uncharacterized | Simi| 0.1/0.01 | 0.20.0 | 9.25 | N | — |
| 10638     | F10063823  | Uncharacterized | Sushi/SCC/CCP (B-| 0.2/1.1 | 0.2/1.1 | 9.39 | N | — |
| 14771     | F20088622  | Uncharacterized | 1 TMH; DI | 0.204 | 0.204 | 10.07 | N | — |
| 17017     | R20102101  | Similar to noelin-2 | Olfactomedin like | 0.10.0 | 0.30.0 | 9.14 | N | — |
| 19193     | R30015158  | Uncharacterized | CUB; 1 TMH | 0.30.3 | 0.20.4 | 8.66 | N | — |
| 19686     | F20118112  | Uncharacterized | — | 0.00.0 | 0.30.0 | 5.26 | N | — |
| 22374     | F30134241  | Uncharacterized | Haem peroxidase 3 | 0.1/0.0 | 0.40.1 | 8.91 | N | Cys 8.2%, Gly 18.9% |
| 22496     | F10134971  | Uncharacterized | — | 0.30.0 | 0.10.0 | 9.3 | N | — |
| 26237     | F30157419  | Uncharacterized | — | 0.20.1 | 0.20.1 | 7.89 | N | — |
| 1954      | F20011720  | Neutral ceramidase B | Neutral/alkaline nonlysosomal ceramidase | 0.20.2 | 0.20.2 | 8.26 | Y | — |
| 24553     | R10147316  | Uncharacterized | TSP1 | 0.30.1 | 0.20.2 | 8 | Y | Cys 10.0%, Gly 11.0% |
| 26581     | F10159481  | Uncharacterized | DI | 0.20.1 | 0.20.1 | 11.96 | Y | — |
| 1953      | R10017117  | Neutral/alkaline nonlysosomal ceramidase | Neutral/alkaline nonlysosomal ceramidase; 1 TMH | 0.10.2 | 0.10.2 | 7.77 | N | — |
| 3405      | F30020427  | Similar to cysteine-rich secretory protein | LCC domain-containing 2 | 0.20.0 | 0.10.0 | 7.92 | N | — |
| 14004     | R10084022  | Uncharacterized | — | 0.10.1 | 0.10.1 | 9.27 | N | — |
| 20189     | F20112330  | Uncharacterized | — | 0.10.0 | 0.10.0 | 9.67 | N | Lys 10.1% |
| 21060     | R20126359  | Similar to MAM domain-containing glycosylphosphatidylinositol anchor protein | ConA like lectin/MAM, peptidase S1 | 0.10.2 | 0.10.1 | 8.38 | N | — |
| 21061     | R30126366  | Similar to atrial natriuretic peptide-converting enzyme | LDLRA 2, SCRC 2, Peptidase S1trypsin | 0.10.2 | 0.10.2 | 8.46 | N | — |
| 14224     | F10085339  | Uncharacterized | EGF; DI | 0.10.1 | 0.00.1 | 7.04 | Y | — |
| 18776     | R30126265  | Uncharacterized | CUB; 1 TMH | 0.10.1 | 0.10.1 | 8.02 | Y | — |
| 21109     | F30126651  | Uncharacterized | — | 0.00.2 | 0.00.2 | 6.69 | N | — |
| 27260     | F20163556  | Similar to SCO-spondin/ hemicentin-1 | TSP-1 | 0.00.3 | 0.00.3 | 8.4 | N | Cys 14.8%, Ser 11.8% |
| 24557     | R30147342  | Uncharacterized | 1 TMH | 0.10.8 | <0.010/0.8 | 5.52 | Y | Cys 11.2%, Gly 12.1% |
| 17584     | R20105503  | Uncharacterized | Trypsin inhibitor like Cys-rich | 0.00.6 | <0.01/1.1 | 10.03 | Y | Cys 7.6%, Gly 16.0% |

*Note.—Major proteins were arbitrarily defined as proteins occurring with an average percentage of ≥0.1 in either D1–3 and V1–3 (2 h) or D4–6 and V4–6 (24 h) or both. Known intracellular proteins or membrane proteins with predominantly intracellular function, some of which showed high percentages in samples treated with hypochlorite for 2 h only, are not included. DI, predicted disordered structure; TMH, predicted transmembrane helices; SS, predicted secretion signal peptide.

aPossibly a fragment of the same protein as peptide R10097436.

bPossibly the same protein as contig 14828 (R20088967), possibly the N-terminal half.

cShares two peptides with R10050836.

dPossibly a fragment of the same protein as peptide R10126502.

ePossibly the same protein as contig 6089 (F10036529)—possibly the C-terminal half.

fShares peptides with F20107906.
elucidated (Shen et al. 1997), possesses an extensive Gly/Ser-rich domain that has been suggested to function as a molecular shock absorber and to inhibit the propagation of fractures through the biomineral (Smith et al. 1999). The phosphorylation sites of R10063526 (fig. 4A) occur in a single peptide on two closely spaced serines and were identified without prior enrichment of phosphopeptides. Comparison of the number of unmodified copies of this peptide and the phosphorylated forms (among approximately 100 identifications) indicated a high site occupancy and therefore suggested some functional or structural importance of these phospho sites, although this possible function may not be related to interaction with the mineral phase because such phosphoproteins usually show multiple phosphorylation sites. This protein disappeared after 24 h of hypochlorite treatment and may therefore be considered as a potential intercrystalline protein (fig. 5). Another phosphorylated major protein that became less abundant after aggressive hypochlorite treatment was F20107906 (fig. 4B), which contained one single phosphoserine with high site occupancy (fig. 4B). This protein, similar to phosphoprotein R10063526, did not return a significant FASTA hit and its sequence was unusually rich in glycine and alanine (table 1).

Conspicuous in their absence from the 66 most abundant proteins we detected in shells of M. venosa are any proteins...
with alpha-carbonic anhydrase (α-CA) domains. As far as we are aware, all full-scale and even most partial proteome surveys of metazoan biominerals have revealed at least one (and usually several) protein with these domains (Jackson et al. 2007; Mann et al. 2008; Marie et al. 2012, 2013; Zhang et al. 2012; Ramos-Silva et al. 2013; Le-Roy et al. 2014; Mann and Edsinger 2014; Mann and Jackson 2014; Voigt et al. 2014). Although we did not detect α-CAs as

**Fig. 4.**—Spectrum of the major phospho-peptides of R10063526 and F30059157. (A) The doubly phosphorylated R10063526 peptide containing phosphorylated S272 and S277 of this entry was identified with a posterior error probability (PEP) of 0.0002 and a mass error of 0.7 ppm. The MaxQuant localization probability was 1 for both sites. Expert System advanced annotations (Neuhauser et al. 2012) were omitted to reduce complexity, except for the last two peaks (m/z 1144.41 and 1045.44). Peak X was not annotated by the expert system but most probably represents [M+H]-2H3PO4-H2O. Y-ions are shown in red, b-ions are shown in blue, b- or y-ions with a loss of water or ammonia are in orange. Asterisks indicate loss of H3PO4. (B) The singly phosphorylated F30059157 peptide (S174) was identified with a mass error of 0.7 ppm, a PEP of 3.4e-67, and a phosphosite localization probability of 1. Y-ions are shown in red, b-ions are shown in blue, b- or y-ions with a loss of water or ammonia are in orange. Asterisks indicate loss of H3PO4. Expert System advanced annotations (Neuhauser et al. 2012) are omitted for clarity of presentation.
members of the major proteinaceous component of *M. venosa*’s shell, we could identify them as minor components (supplementary file S5, Supplementary Material online) and also as sequences present in the mantle transcriptome. We could identify 12 such contigs (140, 2070, 2823, 5097, 5231, 5661, 17483, 17621, 19612, 20103, 20347, and 27050), 6 of which were putatively full length, and 5 of which possessed signal sequences (supplementary file S11, Supplementary Material online). Further analysis of the activity and cellular spatial localization of these transcripts would be informative, especially in a comparative context with other biomineralizing taxa.

As mentioned above, the relative abundances of several major proteins (supplementary file S12, Supplementary Material online) were significantly affected by our extended (24 h cf. 2 h) treatment of *M. venosa* shells with sodium hypochlorite (fig. 5). Although we have no data that can directly explain this phenomenon, one obvious possibility is that these two “populations” of proteins represent inter- and intracrystalline proteins, respectively. An extended sodium hypochlorite treatment might be expected to degrade more accessible intercrystalline proteins resulting in an increase in the relative abundances of intracrystalline proteins. Also of note is the fact that most proteins are found in equal abundance in dorsal and ventral valves (the line of best fit in fig. 5 has a slope of 0.96 and a correlation coefficient of 0.80), and the majority of the peptides we observe does not appear to be significantly affected by the different sodium hypochlorite treatments (the majority of the data is focused on the origin). The functions of those peptides that are apparently affected by our two different hypochlorite treatments are difficult to infer due to a lack of sequence similarity to previously characterized proteins; however, many of these uncharacterized proteins contain anomalous amino acid contents typical of known biomineralizing proteins.

**Relevance of *M. venosa*’s Shell Ultrastructure to Its Reaction to Sodium Hypochlorite Treatment and Its Shell-Forming Proteome**

The shell of *M. venosa* (and of many other brachiopods producing calcitic shells) contains extensive perforations known as
punctae which span the entire thickness of the shell (fig. 6) (Pérez-Huerta et al. 2009 and references cited therein). The punctae of M. venosa are typical of terebratulids, with a relatively narrow opening at the side of the shell facing the mantle epithelium and a funnel-shaped end without ramifications adjacent to the periostracum. Although the function of punctae remains a topic of research, a significant result of their presence is the apparent increased exposure of the intercrystalline organic matrix to molecules in the surrounding medium. Punctae are likely relevant to the in vitro preparation methods we employed in this study, and are possibly relevant to microbial and/or predator mediated degradation of calcitic brachiopod shells by proteases (see discussion below). Softening of calcite brachiopod shells as a consequence of hypochlorite treatment has been previously described and attributed to a loss of organic matrix. Collins et al. (1988) demonstrated that prolonged hypochlorite treatment (48 h) is enough to reduce the strength of calcite brachiopod shells to an extent that they can be easily disintegrated by shaking, a phenomenon in our experience not commonly observed with mollusk shells. A similar softening of brachiopod shells was also observed when the shells of dead Terebratulina retusa animals were simply maintained in sea water; after 218 days these shells became extremely fragile and lost more than 90% of their pressure resistance (Collins 1986; Emig 1990). Such studies clearly demonstrate both the important functional role of the organic extracrystalline matrix (those molecules that surround lamellae, fibers, tablets, and all other biomineral substructures vs. occluded intracrystalline molecules) for maintaining shell strength and integrity, and that the accessibility of this matrix to molecules such as microbial proteases may be the cause of extracrystalline matrix degradation in rotting shells (Collins 1986). Conspicuous in their paucity from the M. venosa shell-forming proteomes are proteins possessing protease inhibitor domains (see table 1 and discussion below). Such proteins feature prominently in several molluscan shell-forming proteomes, such as abalone (Marie et al. 2010), limpet (Marie et al. 2013), and the pulmonate Cepaea nemoralis (Mann and Jackson 2014). The relatively abundant presence of these proteins in such molluscan data sets previously led us to suggest that they may function within molluscan shells to prevent degradation of the organic matrix by microbial or shell-boring predatory action (Mann and Jackson 2014). Although such a hypothesis awaits experimental validation, it is interesting to note that brachiopod shells have been documented to be susceptible to such action (Collins 1986; Collins et al. 1988; Emig 1990).

Our broad survey of the shell-forming proteome of M. venosa reveals a wealth of novel proteins that are likely to be playing critical roles in determining the mineralogy, ultrastructure and material properties of the mature biomineral, and highlights the need for meaningful in vivo functional assays to be developed for nonmodel organisms such as M. venosa. The data that we present here serve as a platform from which more targeted and in-depth analyses can be performed on M. venosa’s biomineralization strategy. Such efforts would clarify the function of the major elements of M. venosa’s shell-forming proteome, and in turn allow for clearer comparisons to be made between the biomineralizing proteomes of more distant metazoan taxa. Insights gained from such work would guide the development of hypotheses related to how the metazoan ability to biomineralize evolved, an ability of great relevance to the rapid radiation of complex multicellular life in the early Cambrian.

**Supplementary Material**

Supplementary files S1–S12 are available at Genome Biology and Evolution online (http://www.gbe.oxfordjournals.org/).

**Acknowledgments**

The authors acknowledge Matthias Mann (MPI of Biochemistry, Martinsried) for generous support, Gaby Sowa
(MP) for preparing the capillary columns, Korbinian Mayr and Igor Paron (both MP) for keeping the mass spectrometers in excellent condition, Simone Schätzle for assistance with RNA extraction, and Roland Melzer for assistance with the SEM work. They also thank to Günter Försterra, Emma Plotnek, Dan Genter, Ulrich Pörtschmann, and Katie McConnell for helping to collect the brachiopods. This work was supported by funding to D.J.J. through the German Excellence Initiative and the Deutsche Forschungsgemeinschaft (DFG) project #JA2108/2-2, as well as funds from the Department of Earth and Environmental Sciences to W.S. and G.W., and from the research department of the Museum für Naturkunde, Berlin to C.L. This is publication nr 121 of Huinay Scientific Field Station. The authors declare no competing interests.

**Literature Cited**

Accinni L, Natali PG, Silvestrini M, Martino CD. 1983. Actin in the extracellular matrix of smooth muscle cells. An immunoelectron microscopic study. Connect Tiss Res. 11:69–78.

Brand U. 2003. Geochemistry of modern brachiopods: applications and implications for oceanography and paleoceanography. Chem Geol. 198:305–334.

Collins MJ. 1986. Post-mortality strength loss in shells of the Recent articulateterebratulina retusa from the west coast of Scotland. Bistratigr Palaeoizkue. 4:209–218.

Collins MJ, et al. 1988. Sero-taxonomy of skeletal macromolecules in living Genome Biol. Evol. 7(5):1349–1362. doi:10.1093/gbe/evv074 Advance Access publication April 24, 2015 1361

Collins MJ, Muyzer G, Curry GB, Sandberg P, Westbroek P. 1991. Terebratulina retusa from the west coast of Scotland. Acta Zool Fenn. 182:133–143.

Curry GB, Fallick AE. 2002. Use of stable oxygen isotope determinations from brachiopod shells in palaeoenvironmental reconstruction. Palaeogeogr Palaeoclimat Palaeoecol. 182:133–143.

Cusack M, Curry G, Clegh H, Abbott G. 1992. An intracrystalline chromoprotein from red brachiopod shells: implications for the process of biomineralization. Comp Biochem Physiol B 102:93–95.

Cusack M, Huerta AP, Huerta AP. 2012. Brachiopods recording seawater temperature, a matter of class or maturation? Chem Geol. 334: 139–143.

Darzentas N. 2010. Circos: visualizing sequence similarity with Circos. Bioinformatics 26:2620–2621.

Dosztányi Z, Csizmok V, Tompa P, Simon I. 2005. iUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21:3433–3434.

Drake JL, et al. 2013a. Peptidomics analysis of several subtaxa of brachiopods (Brachiopoda: Phoronata) and that genetic divergence of metazoan phyla began long before the early Cambrian. Proc Natl Acad Sci U S A. 110:3788–3793.

Drake JL, et al. 2013b. Reply to Ramos-Silva et al.: Regarding coral skeletal proteome. Proc Natl Acad Sci U S A. 110:E2147–E2148.

Empic CC. 1990. Examples of post-mortality alteration in Recent brachiopod shells and (paleo)ecological consequences. Mar Biol. 104:233–238.

Ettenssohn CA. 2014. Horizontal transfer of the msp130 gene supported the evolution of metazoan biomineralization. Evol Dev. 16:139–148.

Gaspard D, et al. 2008. Shell matrices of Recent rhynchonelliform brachiopods: microstructures and glycosylation studies. Earth Environ Sci Trans R Soc Edinb. 98:415–424.

Gasteiger E, et al. 2005. Protein identification and analysis tools on the ExPaSy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa (NJ): Humana Press. p. 571–607.

Goetz AJ, et al. 2009. Calcite morphology, texture and hardness in the distinct layers of rhynchonelliform brachiopod shells. Eur J Min. 21:303–315.

Goetz AJ, et al. 2011. Interdigitating biocalcitic dendrites form a 3-D jigsaw structure in brachiopod shells. Acta Biomater. 7:2237–2243.

Goujon M, et al. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 38:W695–W699.

Hohagen J, Jackson DJ. 2013. An ancient process in a modern mollusc: early development of the shell in Lymnaea stagnalis. BMC Dev Biol. 13:27.

Hunter S, et al. 2012. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res. 40: D306–D312.

Ishida T, Kinoshita K. 2007. PrDOS: prediction of disordered protein regions from amino acid sequence. Nucleic Acids Res. 35:W460–W464.

Jackson DJ, et al. 2010. Parallel evolution of nacre building gene sets in molluscs. Mol Biol Evol. 27:591–608.

Jackson DJ, Macis L, Retnier J, Degnan BM, Wörheide G. 2007. Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis. Science 316:1893–1895.

Jope M. 1966a. The protein of brachiopod shell—I. Amino acid composition and implied protein taxonomy. Comp Biochem Physiol. 20: 593–600.

Jope M. 1966b. The protein of brachiopod shell—II. Shell protein from fossil articulate shells. Comp Biochem Physiol. 20: 601–605.

Knirps E. 1977. Zur onogenese des schalenfeldes von Lymnaea stagnalis. Dev Genes Evol. 181:11–30.

Knoll AH. 2003. Biomineralization and evolutionary history. Rev Mineralogy Geochem. 54(1):329–356.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol Evol. 27:195–202.

Krywinski M, et al. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Le-Roy N, Jackson DJ, Marie B, Ramos-Silva P, Marin F. 2014. The evolution of metazoan α-carbonic anhydrases and their roles in calcium carbonate biomineralization. Front Zool. 11:75.

Lee X, et al. 2004. Ontogenetic trace element distribution in brachiopod shells: an indicator of original seawater chemistry. Chem Geol. 209: 49–65.
Lowenstam HA, Weiner S. 1989. On biomineralization. New York: Oxford University Press.

Mann K, Edsinger E. 2014. The Lottia gigantea shell matrix proteome: re-analysis including MaxQuant iBAQ quantitation and phosphoproteome analysis. Proteome Sci. 12:28.

Mann K, Jackson DJ. 2014. Characterization of the pigmented shell-forming proteome of the common groove snail Cepaea nemoralis. BMC Genomics 15:249.

Mann K, Poustka AJ, Mann M. 2008. The sea urchin (Strongylocentrotus purpuratus) test and spine proteomes. Proteome Sci. 6:22.

Marie B, et al. 2010. Proteomic analysis of the calcified organic matrix of the tropical abalone Holotis asinina shell. Prot Sci. 8:54.

Marie B, et al. 2012. Different secretory repertoire controls 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, et al. 2013. The shell-forming proteome of Lottia gigantea reveals both deep conservations and lineage specific novelties. FEBS J. 280:214–232.

Mass T, Drake JL, Peters EC, Wang W, Falkowski PG. 2014. Immunolocalization of skeletal matrix proteins in tissue and mineral of the coral Stylophora pistillata. Proc Natl Acad Sci U S A. 111:12728–12733.

Merkel C, et al. 2007. Micromechanical properties and structural characterization of modern inarticulated brachiopod shells. J Geophys Res Biogeosci. 112:1–12.

Merkel C, et al. 2009. Mechanical properties of modern calcite- (Mergerlia truncata) and phosphate-shelled brachiopods (Discradisca stella and Lingula anatina) determined by nonindentation. J Struct Biol. 168:396–408.

Michalski A, et al. 2011. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. Mol Cell Proteomics. 10:1–11.

Neuhouser N, Michalski A, Cox J, Mann M. 2012. Expert system for computer-assisted annotation of MS/MS spectra. Mol Cell Proteomics. 11:1500–1509.

Newman AM, Cooper JB. 2007. XSTREAM: a practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. BMC Bioinformatics 8:382.

Parkinson D, Curry GB, Cusack M, Fallick AE. 2005. Shell structure, patterns and trends of oxygen and carbon stable isotopes in modern brachiopod shells. Chem Geol. 219:193–235.

Pérez-Huerta A, et al. 2009. Brachiopod punctae: a complexity in shell biomineralisation. J Struct Biol. 167:62–67.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignAP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 8(10):785–786.

Rahman MA, Shinjo R, Oomori T, Worheide G. 2013. Analysis of the proteinaceous components of the organic matrix of calcitic sclerites from the soft coral Sinularia sp. PLoS One 8:e58781.

Ramos-Silva P, et al. 2013. The skeletal proteome of the coral Acropora millepora: the evolution of calcification by co-option and domain shuffling. Mol Biol Evol. 30:2099–2112.

Ramos-Silva P, et al. 2014. The skeleton of the staghorn coral Acropora millepora: molecular and structural characterization. PLoS One 9:e97454.

Rappold J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc. 2:1896–1906.

Rudwick MIS. 1970. Living and fossil brachiopods. London: Hutchinson.

Scheltema RA, Mann M. 2012. SprayQc: a real-time LC-MS/MS quality monitoring system to maximize uptime using off the shelf components. J Proteome Res. 11:3458–3466.

Schmahl WW, et al. 2008. Hierarchical fibre composite structure and micromechanical properties of phosphatic and calcitic brachiopod shell biomaterials—an overview. Mineral Mag. 72:541–562.

Schmahl WW, et al. 2012. Hierarchical structure of marine shell biomaterials: biomechanical functionalization of calcite by brachiopods. Z Kristallogr Cryst Mater. 227:793–804.

Schwanhauser B, et al. 2011. Global quantification of mammalian gene expression control. Nature 473:337–342.

Sharma K, et al. 2014. Ultra deep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep. 8:1583–1594.

Shen X, Belcher AMB, Hansma PK, Stucky GD, Morse DEM. 1997. Molecular cloning and characterization of Lustrin A, a matrix protein from shell and pearl nacre of Haliotis fucunensis. J Biol Chem. 272:32472–32481.

Sievers F, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7:539.

Simkiss K, Wilbur KM. 1989. Biomineralization, cell biology and mineral deposition. San Diego (CA): Academic Press.

Smith BL, et al. 1999. Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. Nature 399:761–763.

Sperling EA, Pisani D, Peterson KJ. 2011. Molecular paleobiological insights into the origin of the Brachiopoda. Evol Dev. 13:290–303.

Taylor PD, Vinn O, Wilson MA. 2010. Evolution of biomineralization in lophophorates. Spec Pap Palaeontol. 84:317–333.

Timmermans LPM. 1969. Studies on shell formation in molluscs. Neth J Zool. 19:417–523.

Voigt O, Adamski M, Sluzek K, Adamska M. 2014. Calcareous sponge genomes reveal complex evolution of alpha-carbonic anhydrases and two key biomineralization enzymes. BMC Evol Biol. 14:230.

Williams A. 1989. Biomineralization in the lophophorates. Skeletal biomineralization: patterns, processes and evolutionary trends. In: Carter JG, editor. Skeletal biomineralization: patterns, processes and evolutionary trends, Vol. 1. New York: Van Nostrand Reinhold. p. 67–82.

Williams A. 1997. Shell structure. In: Kaesler RL, editor. Treatise on invertebrate paleontology, Part H: Brachiopoda, revised. Vol. 1. Boulder (CO)/Lawrence (KS): The Geological Society of America/University of Kansas. p. 267–320.

Wisniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. Nat Methods. 6:359–362.

Yu X, et al. 2014. Molecular cloning and differential expression in tissues of a tyrosinase gene in the Pacific oyster Crassostrea gigas. Mol Biol Rep. 41:5403–5411.

Zhang C, Xie L, Huang J, Chen L, Zhang R. 2006. A novel putative tyrosinase involved in periostracum formation from the pearl oyster (Pinctada fucata). Biochim Biophys Res Commun. 342:632–639.

Zhang G, et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49–54.

Associate editor: Dan Graur

Jackson et al.