HAL open science

Proteins as functional units of biocalcification - An overview.

Paula Ramos-Silva, Frédéric Marin

To cite this version:

Paula Ramos-Silva, Frédéric Marin. Proteins as functional units of biocalcification - An overview.. F. Marin, F. Brümmer, A. Checa, G. Furtos, I.G. Lesci & L. Šiller. Biomineralization: from fundamentals to biomaterials & environmental issues, 672, Trans Tech Publications Ltd, pp.183-190, 2015, Key Engineering Materials, 978-3-03835-591-5. 10.4028/www.scientific.net/KEM.672.183 . hal-01194587

HAL Id: hal-01194587

https://hal.science/hal-01194587v1

Submitted on 18 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Proteins as functional units of biocalcification – an overview

Paula Ramos-Silva[1,2,a*] and Frédéric Marin[2,b]

[1]Instituto Gulbenkian de Ciência, Oeiras, Portugal,
[2]UMR CNRS-uB 6282 Biogéosciences, 6 boulevard Gabriel, Université de Bourgogne, Dijon France
[a]prsilva@igc.gulbenkian.pt, [b]frederic.marin@u-bourgogne.fr
* Corresponding author

Keywords: biomineralization, calcifying matrix, proteins, proteomics, sequences, functional domains

Abstract. High-throughput approaches such as genomics, transcriptomics and proteomics have led to the discovery of a larger set of biomineralization genes than previously foreseen. These gene lists are often difficult to decode in light of the current models of calcification. Here we overview the proteins available in UniProt (Universal Protein Resource), that were identified directly in metazoan calcium carbonate mineralized structures or known to have direct key-functions in calcification processes. Functional annotation of the protein datasets using Gene Ontology reveals that functions like carbohydrate binding, structural and catalytic activities (e.g. hydrolase) are commonly represented across the organic matrices.

1. Introduction

In the metazoan world, many organisms synthesize mineralized structures of calcium carbonate ($CaCO_3$) through a physiological process called biocalcification. A significant number of calcifying species lives in the aquatic environment and belong to distinct groups including sponges, corals, brachiopods, mollusks, crustaceans, echinoderms and chordates. A common feature among these groups of organisms is the production of $CaCO_3$-mineralized structures by a biologically controlled mechanism. Indeed, this process is highly regulated at all stages of biomineral formation [1]. In particular, it requires: (1) the delineation of the mineralizing space by cell membranes or polymers; (2) the formation of an array of macromolecules and its subsequent targeting to the site of mineralization; (3) the pumping of inorganic ionic precursors to set up a saturated environment; and finally (4), the control of certain proteins over crystal nucleation, growth and inhibition, providing a scaffold for mineral deposition. Although biologically controlled mineralization is present in microorganisms such as bacteria (magnetotactic [2]) and unicellular algae (diatoms [3], coccolithophores [4]) or protozoans (ciliates [5]), the process is undeniably more broadly represented in animals. Non-vertebrates in particular can produce a wide-range of external $CaCO_3$-biominerals such as shells (in mollusks and in brachiopods), carapaces (in crustaceans), exoskeletons (in corals and sponges) and calcified tubes (in serpulid annelids), but also internal ones like spicules and spines (in sponges and echinoderms, respectively) and gastroliths (in crustaceans). In vertebrates, $CaCO_3$–mineralized tissues are less common. The main examples refer to structures like fish otoliths [6] and otoconia [7], which primary function is to provide sense of gravity, balance, movement and direction, and eggshells that provide an outer hard protection to avian eggs [8].

It is common to describe biominerals as organo-mineral assemblages where the mineral, the major component, is embedded in a minor organic matrix. This latter is often a complex mixture of macromolecules - proteins, glycoproteins, polysaccharides and lipids - produced by the organisms and directly involved in the regulation of the mineral deposition. In most cases, the macromolecules are occluded within the $CaCO_3$ during its growth, constituting the "organic fraction" of the

biomineral or the more generally called organic matrix (OM). Although there is still much to study about the exact mechanisms by which the components of the OM, in particular the proteins, control the formation of the calcified tissues, some key features of these proteins are recognized as being functionally important. The first common trait is the high content in aspartic acid and, in a lesser degree, glutamic acid. This feature is frequently observed in proteins of mollusk shells and coral exoskeletons [9]. These proteins are recognized to have strong inhibitory function on the crystal growth and they are suspected to control the crystal morphology by selective adsorption on mineral surfaces [9–12], due to their multiple negative charges. Another interesting feature of OM proteins are the post-translational modifications, in particular, glycosylations and phosphorylations, which can greatly contribute to the polyanionicity of such proteins [13–16]. Besides glycoproteins, the OM contains polysaccharides, which are important for structuring [17] the organic-mineral framework but also for interacting directly with the crystal surfaces [13]. Enzymes have also been identified in organic matrices, the best example being the carbonic anhydrase. The function of this family of proteins is to catalyze the reversible hydration of CO_2, forming one bicarbonate ion and one proton, according to the following reaction: $H_2O + CO_2 \rightarrow H_2CO_3 \rightarrow HCO_3^- + H^+$. Since the first report by Miyamoto and co-workers of Nacrein - a modular protein with two carbonic anhydrase domains intercalated by an acidic domain - from the mollusk shell of the pearl oyster *Pinctada fucata* [18], many proteins in shells and other structures were identified by a one-per-one approach using classical molecular biology and biochemistry techniques. Until recently, the existing information was manageable and consisting of some proteins with very specific signatures: secretion signals, low complexity regions, repeats and some predictable domains (acidic, carbonic anhydrase, chitin-binding). The application of high-throughput techniques, genomics, transcriptomics and proteomics has changed the perspective, by revealing a much wider range of proteins that may be taking part in the control of biocalcification [19–22]. This clearly suggests broader functions for the OM proteins such as cross-communication between the mineral-producing tissue and the extracellular calcifying matrix, which go far beyond the catalysis of crystal nucleation and the guiding of mineral growth via binding to specific crystal surfaces. Here we present an overview of the proteins that have been identified in $CaCO_3$-mineralized tissues of some metazoan groups currently available on UniprotKB [23,24]. We characterize this protein dataset from a functional viewpoint by using the Gene Ontology (GO) terms associated to the sequences [25].

2. Functional analysis of $CaCO_3$-biomineralization proteins: data gathering

Keyword searches related to $CaCO_3$-biomineralization processes were performed against the UniProtKB on March 2013 with taxonomic filters for metazoan non-vertebrate aquatic calcifying phyla: Porifera, Cnidaria, Brachiopoda, Mollusca, Echinodermata. This output was manually confirmed in order to have a representative set of reference proteins (Table 1). The most represented group is by far Mollusca due to their economical interest as food source and as 'providers' of cultured pearls. Consequently, these last years have seen an increased application of high-throughput approaches to molluscan tissues in combination with available genomes [26]. Because of the importance of reef-ecosystems in connection to environmental concerns at global scale – such as ocean acidification -, more sequences related to calcification have recently been identified from coral skeletons [19,22]. However the number of sequences available for these organisms is still low. Also for echinoderms the number of confirmed proteins remains limited [27]. Finally, sponges and brachiopods are not popular models in biomineralization studies, exhibiting the smallest publication rate in the field, particularly when regarding molecular characterizations of the organic matrix [28,29]. They consequently remain the groups with the lowest number of identified proteins involved in calcification.

Table 1 – Summary of the proteins identified as directly involved in metazoan $CaCO_3$-biomineralization collected from the UniprotKB (March 2013). The final set was obtained based on keyword search and the following criteria: (1) Experimentally proved to be involved in biomineralization and (2) Inferred by homology (i.e. homologous counterparts in other organisms).

Phylum	# total sequences	Calcified structure	Functions	# biomineralization proteins	# biomineralization proteins with GOs
Mollusca	34370	Shell	Protection	355	178
Echinodermata	28394	Test, Spine, Teeth	Protection	33	27
Cnidaria	16132	Skeleton/Sclerites	Support	81	57
Brachiopoda	135	Shell	Protection	2	0
Porifera	26219	Spicules, Skeleton	Support	9	5

Each set of proteins was characterized for their content in GO terms belonging to the categories - Molecular Function (MF), Cellular Component (CC) and Biological Process (BP). We selected the top 5 GO terms of each group with more than one hit (Fig.1 and Fig. 3). Biological Process (BP) terms were only included in mollusks (Fig. 2) as only this group of aquatic organisms had specific ontologies of the biomineralization process (shell calcification, GO:0031215). GO terms provide good predictions of protein function, and more broadly of the whole organic matrix, since a good share of the proteins listed in Table 1 are associated to ontologies.
Fig. 1 summarizes the top GO terms, in specific categories, associated to shell matrix proteins in mollusks. The main protein functions relate to the catalytic activity of carbonic anhydrases (e.g. nacreins) and hydrolases, many types of binding functions like metal ion binding, carbohydrate and polysaccharide (e.g. perlucin, lustrin A) and enzyme regulatory activity (e.g. the proteins PSPI1 and NSPI1, containing protease inhibitor domains).
As discussed in earlier works [21,22,30], because the biomineral formation takes place extracellularly for each of the considered phylum, they are expected to have extracellular region and proteinaceous extracellular matrix as the main GO terms associated to protein localization.

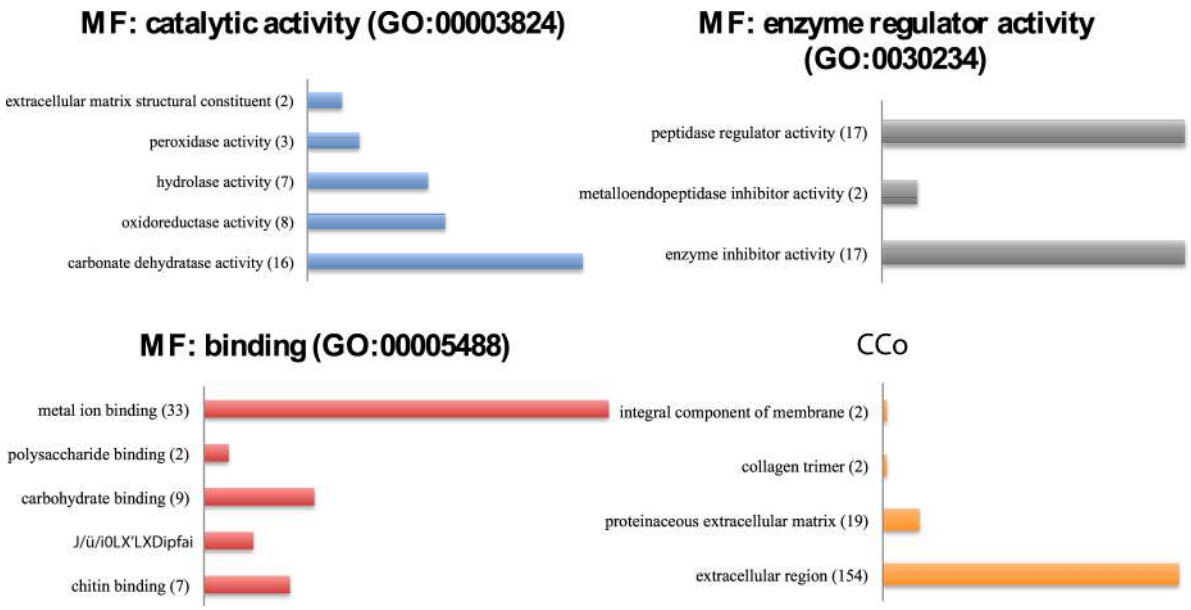

Fig. 1 – Top 5 (> 1 hit) Gene Ontology terms associated to shell matrix proteins (mollusks) within specific GO categories of Molecular Function (MF), and Cellular Component (CC).

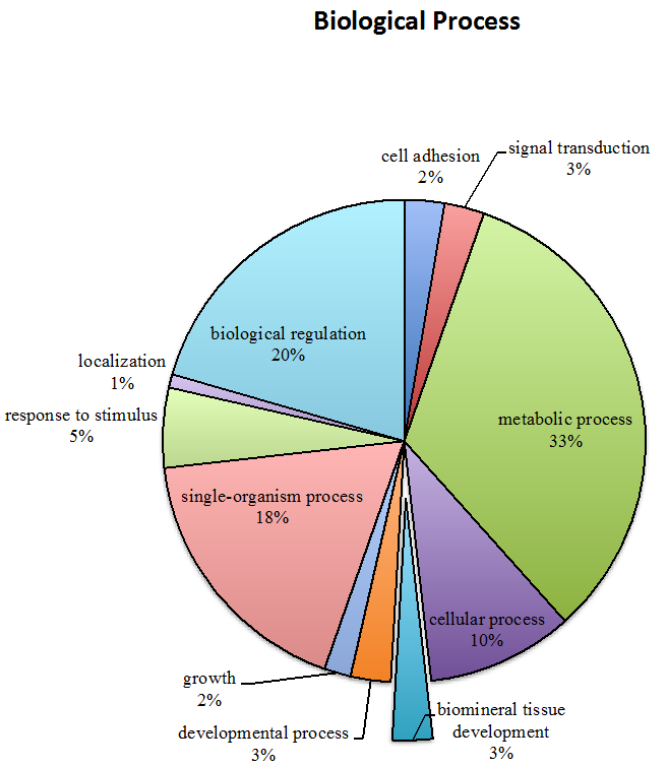

Fig. 2 - Gene Ontology view of the Biological Processes (BP) associated to the shell matrix proteins dataset (mollusks).

In corals and echinoderms, top molecular functions (MF, Fig. 3) are enzymatic and binding/structural. There is equivalent distribution of protein localization regarding the extracellular matrix environment and the cell membrane. This suggests a higher content of transmembrane proteins [22,31] in the OM of these organisms than in the OM of mollusks, for which the proportion of proteins targeted to the extracellular space is exceptionally high. These differences may reflect fundamental differences in the molecular process of biomineral formation in corals and echinoderms, on the one hand, and in molluscs on the other hand: while the first ones may involve both extracellular proteins and proteins with transmembrane domains – a fact that was evidenced for corals [22], extracellular proteins are the dominant constituents of shell matrices in mollusks.

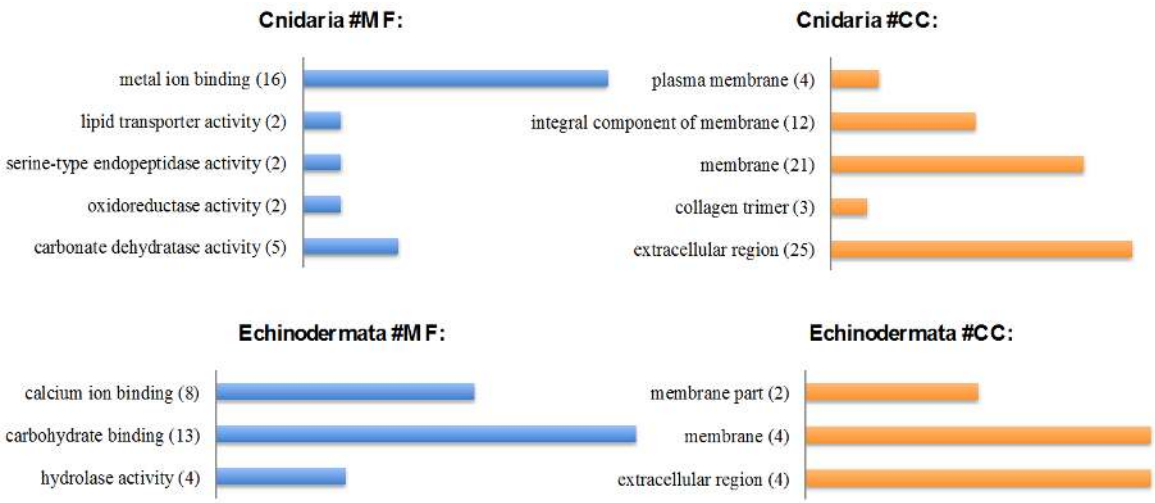

Fig. 3 - Top 5 (> 1 hit) GO terms associated to skeletal matrix proteins (Cnidarians) and test, spine and tooth matrix proteins (Echinoderms). GO categories: Molecular Function (MF) and Cellular Component (CC).

3. Conclusions

The analysis of gene ontology terms associated to biomineralization proteins brings an additional resource of information to characterize the organic matrix from a functional viewpoint. Indeed automatic annotations cover a significant number of the proteins considered in this study - 50% in mollusks, 70% in corals and 83 % in echinoderms. Still it is important to alert those dealing with the discovery of novel proteins to the necessity of providing detailed information on the experimental characterization upon submission of the nucleotide and protein sequences to generalist databases (e.g. DDBJ/EMBL/GenBank, UniProtKB). Only with the help of the bench scientist, it is possible to ensure a faster update process and accurate information on sequence data immediately available to the community. Indeed, there are only a few biomineralization-specific ontologies, which include otolith and eggshell formation in vertebrates (e.g., otolith morphogenesis (GO:0032474), otolith mineralization (GO:0045299), eggshell formation (GO:0030703)), but also shell calcification in mollusks (GO:0031215) and chitin-based cuticle sclerotization in crustaceans (GO:0036340). In addition GO vocabularies specifically developed for $CaCO_3$-biomineralization in other organisms are scarce. The community who studies organic matrix molecules can also contribute to the enrichment of the currently available vocabulary by suggesting new terms via the GO curators. These initiatives will greatly contribute to faster identification of biomineralization proteins *in silico* for non-model organisms and lead to better-automated annotations.
To conclude, the analysis of GO terms contributes to a general overview of the proteins associated to $CaCO_3$-tissues. Fig. 4 summarizes the main types of molecules in terms of the primary structure, localization and possible role in calcification.

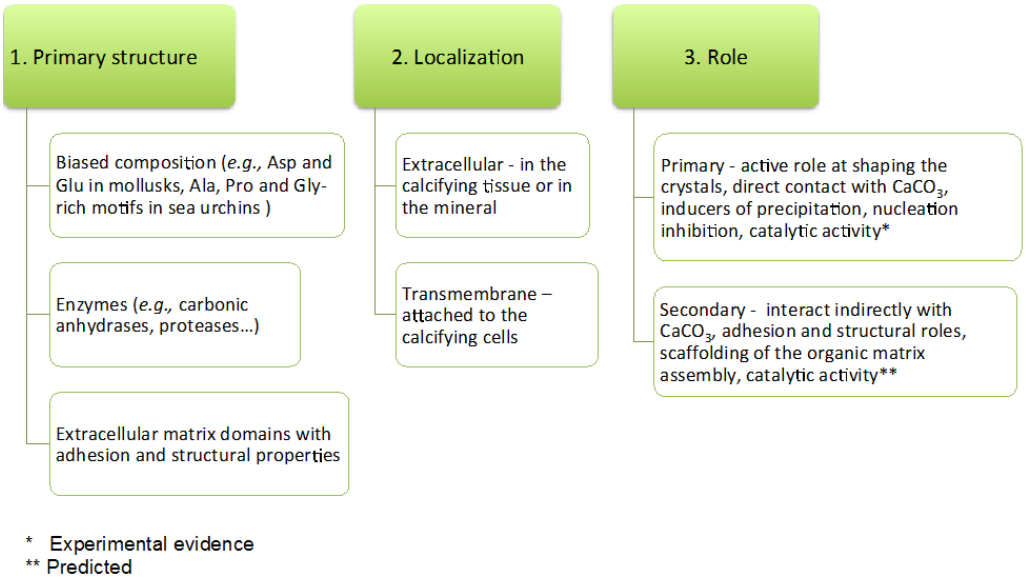

Fig. 4 - Summary of the main types of biomineralization proteins in terms of the primary structure, localization and possible role in calcification.

Acknowledgments

The work of P.R.-S. was supported by the Marie Curie ITN (Initial Training Network), BIOMINTEC piloted by the Johannes Gutenberg University, Mainz, Germany (Pr. H. C. Schröder, see www.biomintec.de) and the BioPreDyn project, coordinated by the Centre for Genomic Regulation (CRG) in Barcelona (see www. biopredyn.eu). Additional support included the COST Action TD0903 ('Biomineralix') during the 2012–2013 period. The work of F. Marin was also supported by an INTERRVIE project from the CNRS-INSU in the same period.

References

[1] H.A. Lowenstam, S. Weiner, On Biomineralization, 1st ed., Oxford University Press, New York, 1989.

[2] D.A. Bazylinski, R.B. Frankel, B.R. Heywood, S. Mann, J.W. King, L. Donaghay, A.K. Hanson, D.A. Bazylinski, R.B. Frankel, B.R. Heywood, S. Mann, J.W. King, P.L. Donaghay, A.K. Hanson, Controlled Biomineralization of Magnetite (Fe3O4) and Greigite (Fe3S4) in a Magnetotactic Bacterium, Appl. Environ. Microbiol. 61 (1995) 3232–3239.

[3] N. Kröger, N. Poulsen, Diatoms-from cell wall biogenesis to nanotechnology, Annu. Rev. Genet. 42 (2008) 83–107.

[4] M.. Marsh, Regulation of CaCO3 formation in coccolithophores, Comp. Biochem. Physiol. Part B (Biochem. Mol. Biol.) 136 (2003) 743–754.

[5] M.-L. Lemloh, F. Marin, F. Herbst, L. Plasseraud, M. Schweikert, J. Baier, J. Bill, F. Brümmer, Genesis of amorphous calcium carbonate containing alveolar plates in the ciliate *Coleps hirtus* (Ciliophora, Prostomatea), J. Struct. Biol. 181 (2013) 155–61.

[6] Y.-J. Kang, A.K. Stevenson, P.M. Yau, R. Kollmar, Sparc protein is required for normal growth of zebrafish otoliths, J. Assoc. Res. Otolaryngol. 9 (2008) 436–51.

[7] Y. Xu, H. Zhang, H. Yang, X. Zhao, S. Lovas, Y.Y.W. Lundberg, Expression, functional, and structural analysis of proteins critical for otoconia development, Dev. Dyn. 239 (2010) 2659–73.

[8] M.T. Hincke, Y. Nys, J. Gautron, K. Mann, A.B. Rodriguez-Navarro, M.D. McKee, The eggshell: structure, composition and mineralization, Front. Biosci. 17 (2012) 1266–80.

[9] L. Addadi, S. Weiner, Interactions between acidic proteins and crystals: stereochemical requirements in biomineralization, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 4110–4.

[10] B.-A. Gotliv, L. Addadi, S. Weiner, Mollusk shell acidic proteins: in search of individual functions, Chembiochem 4 (2003) 522–9.

[11] B.-A. Gotliv, N. Kessler, J.L. Sumerel, D.E. Morse, N. Tuross, L. Addadi, S. Weiner, Asprich: A novel aspartic acid-rich protein family from the prismatic shell matrix of the bivalve *Atrina rigida*, ChembioChem 6 (2005) 304–14.

[12] T. Mass, J.L. Drake, L. Haramaty, J.D. Kim, E. Zelzion, D. Bhattacharya, P.G. Falkowski, Cloning and Characterization of Four Novel Coral Acid-Rich Proteins that Precipitate Carbonates *In Vitro*, Curr. Biol. 23 (2013) 1126–31.

[13] S. Albeck, S. Weiner, L. Addadi, Polysaccharides of Intracrystalline Glycoproteins Modulate Calcite Crystal Growth *In Vitro*, Chem. Eur. J. 2 (1996) 278–84.

[14] K. Mann, A.J. Poustka, M. Mann, Phosphoproteomes of *Strongylocentrotus purpuratus* shell and tooth matrix: identification of a major acidic sea urchin tooth phosphoprotein, phosphodontin, Proteome Sci. 8 (2010) 6.

[15] F. Marin, P. Corstjens, B. de Gaulejac, E. de Vrind-De Jong, P. Westbroek, Mucins and molluscan calcification Molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel *Pinna nobilis* (Bivalvia, pteriomorphia), J. Biol. Chem. 275 (2000) 20667–75.

[16] T. Samata, D. Ikeda, A. Kajikawa, H. Sato, C. Nogawa, D. Yamada, R. Yamazaki, T. Akiyama, A novel phosphorylated glycoprotein in the shell matrix of the oyster *Crassostrea nippona*, FEBS J. 275 (2008) 2977–89.

[17] W.M. Goldberg, Acid polysaccharides in the skeletal matrix and calicoblastic epithelium of the stony coral *Mycetophyllia reesi*, Tissue Cell 33 (2001) 376–87.

[18] H. Miyamoto, T. Miyashita, M. Okushima, S. Nakano, T. Morita, a Matsushiro, A carbonic anhydrase from the nacreous layer in oyster pearls, Proc. Natl. Acad. Sci. USA 93 (1996) 9657–60.

[19] J.L. Drake, T. Mass, L. Haramaty, E. Zelzion, D. Bhattacharya, P.G. Falkowski, Proteomic analysis of skeletal organic matrix from the stony coral *Stylophora pistillata*, Proc. Natl. Acad. Sci. USA 110 (2013) 3788–93.

[20] B. Marie, D.J. Jackson, P. Ramos-Silva, I. Zanella-Cléon, N. Guichard, F. Marin, The shell-forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-specific novelties, FEBS J. 280 (2013) 214–32.

[21] B. Marie, C. Joubert, A. Tayalé, I. Zanella-Cléon, C. Belliard, D. Piquemal, N. Cochennec-Laureau, F. Marin, Y. Gueguen, C. Montagnani, Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell, Proc. Natl. Acad. Sci. USA 109 (2012) 20986–91.

[22] P. Ramos-Silva, J. Kaandorp, L. Huisman, B. Marie, I. Zanella-Cléon, N. Guichard, D.J. Miller, F. Marin, The skeletal proteome of the coral *Acropora millepora*: the evolution of calcification by co-option and domain shuffling, Mol. Biol. Evol. 30 (2013) 2099–112.

[23] The Uniprot Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt), Nucleic Acids Res. 40 (2012) D71–5.

[24] T.U. Consortium, Activities at the Universal Protein Resource (UniProt), Nucleic Acids Res. 42 (2014) D191–8.

[25] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology The Gene Ontology Consortium, Nat. Genet. 25 (2000) 25–9.

[26] G. Zhang, X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P.W.H. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, et al., The oyster genome reveals stress adaptation and complexity of shell formation, Nature 490 (2012) 49–54.

[27] A. Veis, Organic Matrix-related mineralization of sea urchin spicules, spines, test and teeth, Front. Biosci. 16 (2011) 2540.

[14] K. Mann, A.J. Poustka, M. Mann, Phosphoproteomes of Strongylocentrotus purpuratus shell and tooth matrix: identification of a major acidic sea urchin tooth phosphoprotein, phosphodontin, Proteome Sci. 8 (2010) 6.
10.1186/1477-5956-8-6
[15] F. Marin, P. Corstjens, B. de Gaulejac, E. de Vrind-De Jong, P. Westbroek, Mucins and molluscan calcification Molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel Pinna nobilis (Bivalvia, pteriomorphia), J. Biol. Chem. 275 (2000).
10.1074/jbc.m003006200
[16] T. Samata, D. Ikeda, A. Kajikawa, H. Sato, C. Nogawa, D. Yamada, R. Yamazaki, T. Akiyama, A novel phosphorylated glycoprotein in the shell matrix of the oyster Crassostrea nippona, FEBS J. 275 (2008) 2977-89.
10.1111/j.1742-4658.2008.06453.x
[17] W.M. Goldberg, Acid polysaccharides in the skeletal matrix and calicoblastic epithelium of the stony coral Mycetophyllia reesi, Tissue Cell 33 (2001) 376-87.
10.1054/tice.2001.0191
[18] H. Miyamoto, T. Miyashita, M. Okushima, S. Nakano, T. Morita, a Matsushiro, A carbonic anhydrase from the nacreous layer in oyster pearls, Proc. Natl. Acad. Sci. USA 93 (1996) 9657-60.
10.1073/pnas.93.18.9657
[19] J.L. Drake, T. Mass, L. Haramaty, E. Zelzion, D. Bhattacharya, P.G. Falkowski, Proteomic analysis of skeletal organic matrix from the stony coral Stylophora pistillata, Proc. Natl. Acad. Sci. USA 110 (2013) 3788-93.
10.1073/pnas.1301419110
[20] B. Marie, D.J. Jackson, P. Ramos-Silva, I. Zanella-Cléon, N. Guichard, F. Marin, The shellforming proteome of Lottia gigantea reveals both deep conservations and lineage-specific novelties, FEBS J. 280 (2013) 214-32.
10.1111/febs.12062
[21] B. Marie, C. Joubert, A. Tayalé, I. Zanella-Cléon, C. Belliard, D. Piquemal, N. CochennecLaureau, F. Marin, Y. Gueguen, C. Montagnani, Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell, Proc. Natl. Acad. Sci. USA 109 (2012).
10.1073/pnas.1210552109
[22] P. Ramos-Silva, J. Kaandorp, L. Huisman, B. Marie, I. Zanella-Cléon, N. Guichard, D.J. Miller, F. Marin, The skeletal proteome of the coral Acropora millepora: the evolution of calcification by co-option and domain shuffling, Mol. Biol. Evol. 30 (2013).
10.1093/molbev/mst109
[23] The Uniprot Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt), Nucleic Acids Res. 40 (2012) D71-5.
10.1093/nar/gkr981
[24] T.U. Consortium, Activities at the Universal Protein Resource (UniProt), Nucleic Acids Res. 42 (2014) D191-8.
10.1093/nar/gkt1140
[25] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology The Gene Ontology Consortium, Nat. Genet. 25 (2000).
10.1038/75556
[26] G. Zhang, X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que,