Marine biology and biomineralization suffer from a lack of access to cutting-edge Omics technologies

Kalimuthu Kalishwaralal[a], Sam Aldrin Chandran[b], Jaison Arivalagan[c*]

[a] Division of Cancer Research, Rajiv Gandhi Centre for Biotechnology (RGCB), Thiruvananthapuram, India.
[b] School of Chemical and Biotechnology, SASTRA Deemed to be University, Thanjavur, India
[c] Discovery Life Sciences, Malden, MA, USA

1. Introduction

Marine biology is a vast field that focuses its research on organisms ranging from bacteria to the biggest animal on earth. To date, various biomolecules have been studied to characterize marine organisms, their diseases, adaptations, natural products, etc. In this current article, we will focus mainly on the research on biomineralization in marine organisms such as mollusks that could be representative of all the other fields in marine biology (Fig. 1).

Biomineralization is a process by which organisms produce inorganic minerals (biominerals) under the control of organic macromolecules. Biomineralized structures offer excellent protection to the organisms from the external environment and predators. Although the biomineralized structures are made of 95% inorganic minerals, the 5% of organic molecules make it 3000 times fracture resistant than naturally occurring inorganic crystals (Currey, 1977). The organic molecules are secreted by specialized cells lining the biomineralized structures. Studying these organic molecules, including proteins, is critical in conservation and biomimetic material.

Proteomics is the study that identifies and quantifies all the proteins in the sample. In mass spectrometry-based proteomics, proteins are identified based on their mass and charge. Proteomics on the mollusk shell became popular in the past few decades due to the composite structures formed by fabricating proteins, polysaccharides, and calcium carbonate crystals (Arivalagan et al, 2017). However, most studies still use primitive biochemical studies and mass spectrometers, resulting in incomplete data. The current article aims to unravel the recent trends in proteomics and how this could improve the in-depth proteomics analysis.
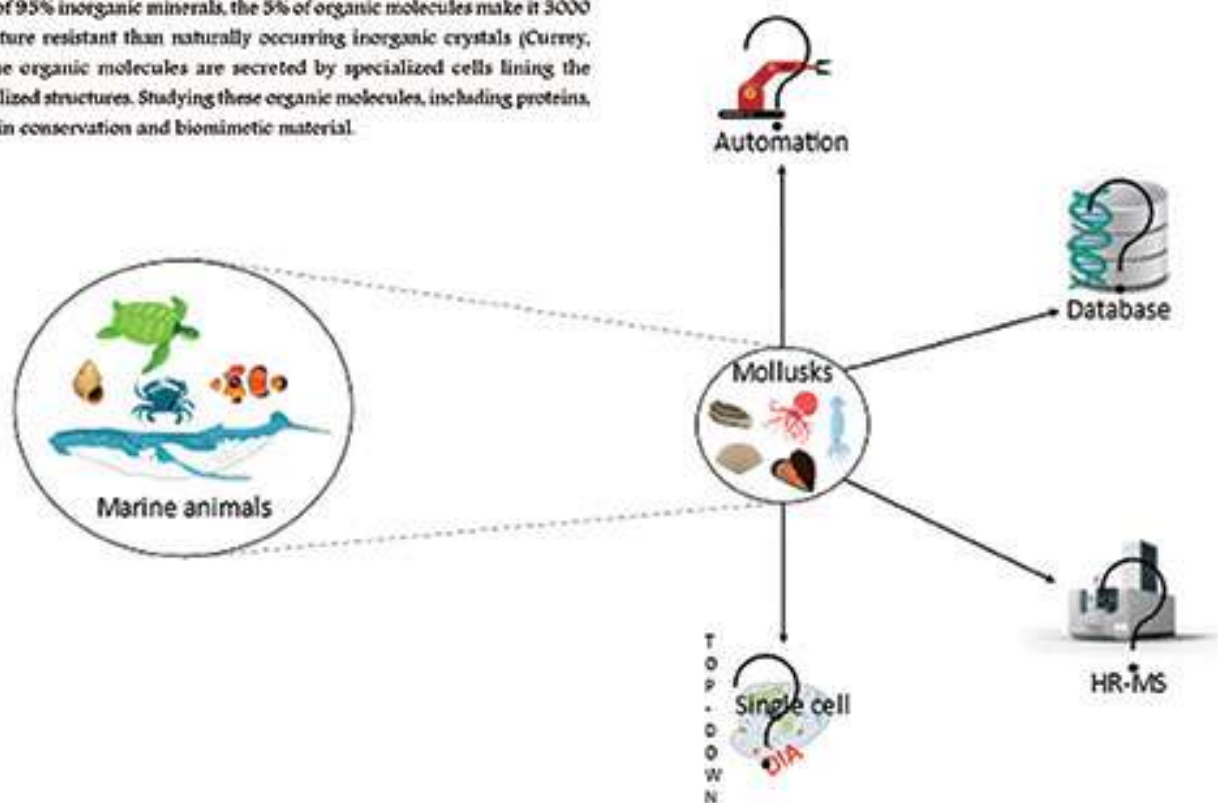

Fig. 1. The cutting-edge technologies and other areas where proteomics is lagging in marine biology and biomineralization

2. Non-model organisms

Most of the proteomics works use the bottom-up approach, where the proteins are digested into peptides and analyzed. The peptides are then matched to proteins using computational tools that rely on proteomic or genomic databases. It is often easier to work with organisms such as yeast, Thale cress, fruit flies, and humans whose genome was completely sequenced (Heck & Neely, 2020). However, non-model organisms were either partially sequenced or not sequenced at all. Post-2015, the new era of genomics started, resulting in an explosion of genomic data, thanks to the continuous drop in the cost of data generation and assembly (Wetterstrand, 2016). Even then, very few marine organisms were sequenced. The genomic database cannot completely solve the issue; annotation plays a significant role in proteomics data search. Although several freely available tools exist, it is a labor and time-consuming task. Only a single work cannot produce a complete genomic database with annotation.

Molluscs is a taxa with only very few species sequenced and annotated. Recent works that aim to identify biomineralization proteins rely on genomic or transcriptome studies on the lining cells that secrete biomineralization proteins (Clark et al., 2020). These studies suffer from incomplete databases and annotations. To understand the complex biomineralization process by marine animals, in-depth proteome profiling is required, which is possible only when the database is complete.

3. High-resolution mass spectrometers and computational tools

It is fascinating that proteomics has advanced greatly from 1D and 2D electrophoresis to advanced tandem high-resolution mass spectrometers (HR-MS) within a decade. In the age of mass spectrometers, it evolved in all aspects, from ionization methods and analyzers to detectors. Current HR-MS is capable of high speed, sensitivity, and accuracy. To give an idea, the recent MS can identify more than 5,000 proteins in a 10-minute gradient.

Most of the research works make use of the HR-MS process clinical samples. They use it for biomarker discovery, drug development, understanding disease mechanisms, etc. Future works on biomineralization should consider using HR-MS, which would result in a complete catalog of proteins and the identification of several isoforms/proteoforms containing post-translational modification.

4. Proteomics workflows

Proteomics always refers to shotgun bottom-up proteomics. However, several complementary workflows exist that could provide more detailed information. To date, proteomics works using Data Dependent Acquisition (DDA) that fragments only the most intense peptides. The low-abundant peptides can be identified only with extensive chromatographic separation, fractionation, or enrichment steps. In data-independent acquisition (DIA), all the peptides were fragmented, resulting in less biased complete analysis (Gillet et al., 2012).

Similarly, the Top-down proteomics workflow complements the bottom-up, where the intact protein is directly analyzed in the mass spectrometer. It has a significant advantage over bottom-up in characterizing isoforms containing PTMs. Little to no DIA or top-down work has been carried out so far. These works might result in identifying new candidate proteins and isoforms in biomineralization, which would aid in understanding the biomineralization process better.

The recent greater leap in proteomics is single-cell proteomics. The HR-MS has advanced to the level of cataloging proteins from a single cell. Few instruments can precisely separate single cells and dispense nanoliters, making single-cell proteomics possible. Applying this workflow to epithelial cells lining the mollusk shells would provide valuable insight into the mechanism of the biomineralization process.

5. Automation and other advanced instrumentations

Sample preparation is the most critical step in any proteomics experiment. Most of the proteomics sample preparation involves enzymatic digestion that usually consumes several hours to days, slowing the overall workflow. Moreover, hundreds of sample preparations can easily introduce human error, leading to analytical inaccuracy (Fu et al., 2023). Recent advancements in instrumentation and automation minimizes sample handling error and fastens the sample preparation efficiently, facilitating the high throughput in the sample analysis. Several instruments, such as AssayMap Bravo, Andrew+, Protograph, etc., offer sample preparation from enrichment, denaturation, reduction, alkylation, and digestion till sample cleanup. With minimal programming, the automation can be taken to a level where human interference is not needed until result interpretation and manuscript preparation.

In marine biology research, especially in biomineralization, sample preparation is tedious, including demineralization and matrix separation. Also, there are still many questions to answer, including proteins specific to shell species, phenotype, region, and microstructure, which might demand huge sample preparation. Adding automation would enable high throughput sample preparation and analysis.

The important part of advanced mass spectrometers, workflows, and automation is the cost. Huge money is not only needed to build such infrastructure but also to analyze huge sets of samples. The pharmaceutical industry is one of the sectors that can spend several billion dollars each year on research and development. Hence, most of the above-mentioned cutting-edge technologies can be accessed only by the pharma industries. Exploring the ocean also has provided more valuable products such as antibiotics, anti-cancer, and anti-inflammatory substances. In the future, marine biology must be provided with access to cutting-edge proteomics techniques that would enable high throughput sample analysis, leading to the discovery of novel compounds. Biomineralization can lead to the discovery of complex mechanisms operating and biomimetic materials.

References

Arivalagan, J., Yarra, T., Marie, B., Sleight, V. A. V. A., Duvernois-Berthet, E., Clark, M. S. M. S., Marie, A., & Berland, S. (2017). Insights from the shell proteome. Biomineralization to adaptation. *Molecular Biology and Evolution*, 34(1), 66–77. https://doi.org/10.1093/molbev/msw219

Clark, M. S., Peck, L. S., Arivalagan, J., Backeljau, T., Berland, S., Cardoso, J. C. R., Caurcel, C., Chapelle, G., De Noia, M., Dupont, S., Gharbi, K., Hoffman, J. I., Last, K. S., Marie, A., Melzner, F., Michalek, K., Morris, J., Power, D. M., Ramesh, K., ... Harper, E. M. (2020). Deciphering mollusc shell production: the roles of genetic mechanisms through to ecology, aquaculture and biomimetics. *Biological Reviews*, 95(6). https://doi.org/10.1111/brv.12640

Fu, Q., Murray, C. I., Karpov, O. A., & Van Eyk, J. E. (2023). Automated proteomic sample preparation. The key component for high throughput and quantitative mass spectrometry analysis. In *Mass Spectrometry Reviews* (Vol. 42, Issue 2). https://doi.org/10.1002/mas.21750

Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., & Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition. A new concept for consistent and accurate proteome analysis. *Molecular and Cellular Proteomics*, 11(6). https://doi.org/10.1074/mcp.O111.016717

Heck, M., & Neely, B. A. (2020). Proteomics in Non-model Organisms. A New Analytical Frontier. In *Journal of Proteome Research* (Vol. 19, Issue 9). https://doi.org/10.1021/acs.jproteome.0c00448

Wetterstrand, K. A. (2016). DNA Sequencing Costs. Data from the NHGRI Large-Scale Genome Sequencing Program. *Www.Genome.Gov/Sequencingcostsdata*, 2.