University of Mons

Faculty of Sciences

Laboratory of Biology of Marine Organisms and Biomimetics

Morphological and molecular characterization of the adhesive system in the mussel *Mytilus edulis* and the tubeworm *Sabellaria alveolata* with a special emphasis on the enzymes involved in the maturation of adhesive proteins

---

Emilie Duthoo

A thesis submitted to the University of Mons in fulfillment of the requirements for the degree of Doctor in Sciences.

Mons, December 2024

Supervisors

Prof. Patrick FLAMMANG (University of Mons, Belgium)
Prof. Matthew J. HARRINGTON (McGill University, Canada)

Jury Members

Dr. Jérôme DELROISSE (University of Mons & University of Liège, Belgium)
Prof. Elise HENNEBERT (University of Mons, Belgium)
Prof. Ruddy WATTIEZ (University of Mons, Belgium)
Prof. Tim WOLLESEN (University of Vienna, Austria)
Dr. Franziska JEHLE (Max Planck Institute for Colloids and Interfaces, Germany)

Remerciements - Acknowledgements

Avant que vous ne lisiez ce manuscrit, j'aimerais prendre le temps de remercier toutes les personnes qui m'ont aidée et soutenue pendant toutes mes années d'études, et sans qui ce travail n'aurait jamais vu le jour.

Tout d'abord, je souhaiterais remercier mes promoteurs, le Professeur Patrick Flammang et le Professeur Matthew J. Harrington, qui m'ont épaulée, soutenue et guidée durant ces quatre dernières années. Merci de m'avoir supervisée dans la réalisation de ce travail. Merci à Patrick de m'avoir accueillie il y a cinq ans pour réaliser un mémoire, puis une thèse, et de m'avoir donné l'opportunité de continuer à travailler sur mon sujet pendant ces prochains mois. Je n'en aurais jamais espéré autant. Votre aide, vos idées, vos conseils, votre disponibilité, ainsi que nos discussions fructueuses, m'ont vraiment aidée à y voir plus clair et à mener à bien ce travail. Thank you, Matt, for all your support and for welcoming me into your laboratory several times. Thank you for all our discussions, your ideas, and, above all, your contagious passion for science. Thank you also for introducing me to the fascinating and complex subject of byssus adhesion.

Je tiens également à remercier le Professeur Igor Eeckhaut de m'avoir accueillie au sein du laboratoire de Biologie des Organismes Marins et Biomimétisme durant ces cinq dernières années, ainsi que de m'avoir transmis sa passion pour la biologie marine à travers ses nombreux cours.

Je remercie également les membres de mon jury qui ont pris le temps de lire, d'évaluer et de commenter ce travail. Thank you to Professor Tim Wollesen and Dr. Franziska Jehle for traveling all the way to Belgium and taking the time to read and evaluate this work. Your kindness and advice during my defense truly helped me. Franziska, it is notably thanks to you and Matt that I had the desire and courage to do a PhD thesis, thank you again for that.

Je souhaite ensuite exprimer ma gratitude au Docteur Elise Hennebert et aux membres du laboratoire de Biologie Cellulaire pour leur aide et leurs conseils, notamment concernant la technique des protéines recombinantes et la révélation des western blots.

Je tiens également à remercier le Professeur Ruddy Wattiez pour sa supervision et ses explications lors de l'utilisation du spectromètre de masse, ainsi que Cyril Mascolo et toute l'équipe du laboratoire de Protéomique et Microbiologie pour leur aide précieuse dans mes analyses protéomiques. Vos conseils m'ont été d'une grande utilité.

Merci également au Docteur Jérôme Delroisse d'avoir été présent dès mon arrivée au labo et de m'avoir encadrée lors de mon mémoire. C'est également grâce à toi que j'ai pu réaliser ce travail. Merci de m'avoir tout montré, de m'avoir aidée à comprendre mes manipulations et pour ta patience (surtout avec la phylogénie). C'est notamment grâce à toi que j'ai acquis autant de connaissances en bio mol !

Je souhaite également remercier le Professeur Jean-Marc Baele pour son aide lors de la mise en résine et du polissage de mes échantillons, ainsi que le GIGA Proteomics Facility de l'Université de Liège pour leur aide dans les analyses protéomiques.

Cette thèse n'aurait pas vu le jour sans la bourse de doctorat FRIA (Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture), octroyée par le Fonds de la Recherche Scientifique - FNRS. Je tiens également à les remercier pour l'octroi des nombreuses bourses [V3], qui m'ont permis d'effectuer des séjours à l'étranger, ainsi que pour les bourses de mobilité C17-C18, qui ont rendu possible ma participation à des conférences. Je remercie également l'Institut de Bioscience de l'Université de Mons, qui a contribué au financement de ces missions et conférences.

Ces remerciements n'auraient aucun sens si je ne mentionnais pas mes collègues de labo, qui sont depuis longtemps devenus bien plus que de simples collègues à mes yeux. Alexia, mille mercis de m'avoir rassurée et soutenue tout au long de ces quatre années (surtout vers la fin, en me disant que c'était normal de me sentir ainsi), pour nos séances de discussions-jacuzzi, nos séances sportives/fous rires, d'être la première à venir m'aider quand j'avais un souci, et de m'avoir introduite au monde de l'UDEM… juste, merci ! 💔 Merci aussi à toi et Guillaume pour nos soirées avec les doggo's, qui me faisaient tellement de bien. Guillaume, merci d'avoir été là, avec Alexia, pour m'écouter dans mes moments de doute. Merci surtout de m'avoir fait découvrir les différentes facettes de la vulgarisation scientifique et pour toutes nos discussions enrichissantes !

Merci Estelle et Lisa d'être arrivées et d'avoir illuminé le train-train quotidien du labo ! Je me souviendrai toujours de nos discussions potins et de nos « petites » pauses-café… Estelle, merci d'être là, pour nos discussions à cœur ouvert dans la voiture, les petits chocolat ou post-it surprises sur mon bureau et pour être toujours en synchro avec moi dans nos motivations à aller au crossfit et affronter Luc (« no no José » 😂). Lisa, merci aussi de m'avoir fait découvrir le doudou sous un autre angle ! Mais aussi pour ta joie de vivre, tes super bons plats faits maison, ton organisation, et surtout toutes les soirées passées ensemble (tes potions

resteront gravées dans ma mémoire 😂). Merci à vous deux également pour nos super sorties plongée carrière ! Nous avons réussi à vaincre le froid glacial de l'eau des carrières ahah. C'est grâce à vous deux, ainsi qu'à Killian, que je me suis remise à la plongée, une passion que j'avais mise de côté.

Killian, merci pour nos moments de soutien face à la complexité de nos sujets, ainsi que pour nos discussions laborieuses, rien que pour essayer de comprendre ce que nous faisions ! Louis, merci pour nos discussions, notamment sur nos débats cinéphiles, pour toutes les sorties piscines/carrières, ainsi que pour ton aide et tes précieux conseils sur l'écrit. Merci aussi à Noé (mon fournisseur officiel de Nalu, ma source de motivation avec le café !) et Antoine, mes partenaires de bureau ! Antoine, merci de m'avoir fait confiance pour lire ton super roman, de m'avoir aidée dans la réalisation de mes figures, et surtout pour ta bonne humeur infaillible. Noé, merci pour nos discussions techniques et pour ton aide précieuse dans la recherche d'équipement de labo (on l'a enfin eu, ce pH-mètre !). Merci aussi à toi, P'tit Gui ! Nos discussions, tes petites surprises déposées sur mon bureau, ainsi que tes mots ont toujours réussi à me réconforter.

Merci également à vous pour toutes nos soirées, nos teambuildings, et surtout de m'avoir suivie dans mes envies de voyage ! Ces petits séjours en Baie de Somme et à Naples étaient tout simplement parfait. Nos excursions, nos fous rires, nos soirées jeux de société et nos sorties au restaurant m'ont vraiment permis de me changer les idées.

Ali, thank you so much for our long discussions on the complex subject of marine adhesion; they really helped me see things more clearly. Thank you also for your kindness, for trusting me, for always encouraging me, and for all the good times spent at Oxford. Nathan, merci pour ta patience face à mes commandes interminables pour le labo, et pour nos discussions sur le sport, ahah. Antoine, mille mercis pour ton aide dans le rafistolage à la MacGyver du labo ! Et bien sûr, merci à tous les autres membres : Wendy (nos discussions sur les ISH), Youri (fais attention quand tu manipes, ahah), Némo (partenaire de crossfit !), Amandine (nos chasses à l'étoile de mer) et Léandro (nos discussions casse-tête sur nos super séquences) !

I'd also like to thank all the people I've met at conferences and during my missions, especially those at McGill. A thousand thanks to all the members of the Harrington Lab for your warm welcome, which truly made me feel at home among you. Thank you for letting me see Montreal from a new perspective and for introducing me to all the great restaurants and

pubs in the area. Gagan, I can't thank you enough for all the support you gave me during these missions, your good humor and cheerfulness were transmissible and put everyone in a good mood, and I'm so happy to count you among my friends! Thanks also to Stephen for those movie nights at your place, to Mathieu for our discussions about research, and to Hamideh for your invaluable help during my experiments. I'd also like to thank Lucia, Sam, Guido, Mostafa, Nathalie, Max, Amy, Tobias, and Jenaes for all our walks and excursions in Montreal!

Merci également à toutes les personnes que j'ai pu rencontrer au cours de ces quatre années de thèse ! Marie et Morgane, vous avez été parmi les premières à m'accueillir au sein du labo, merci infiniment ! Merci aussi pour nos discussions, nos fous rires, ainsi que pour vos conseils et votre aide précieuse lors de mes différentes expérimentations. Carla et Fabien, un grand merci pour votre aide inestimable dans mes recherches. Thank you Elin for your help and advice in solving my problems with recombinant protein. Enfin, merci aux différents mémorants et stagiaires que j'ai eu la chance de croiser au fil de ces années. Votre enthousiasme et votre collaboration ont été précieux.

Je tiens également à remercier toutes les personnes rencontrées lors des activités de vulgarisation scientifique. Aurélie (Agent A), merci de m'avoir fait confiance en m'associant à tes animations pour la jeunesse scientifique ! Florine, mille mercis de m'avoir permis d'animer des ateliers pour les enfants de l'UDEC et de partager avec eux ma passion ! Merci aussi à l'équipe de ma thèse en 180 secondes. Nos séances de coaching, nos discussions et nos exercices m'ont vraiment aidée à mieux présenter et vulgariser mon sujet !

Cette thèse n'aurait également jamais vu le jour sans ma famille, toujours présente pour moi, même quand je fais ma râleuse… Papa, maman, merci d'avoir toujours été là, de m'avoir rassurée et soutenue (je vous en ai fait faire des aller-retours à Mons), et surtout de vous être battus pour moi lors de mon accident ! 🍀 Merci de m'avoir forcée à poursuivre mes études et ma passion, et surtout de m'avoir encouragée à ne pas baisser les bras, même quand tout le monde était contre moi. 'Toine, merci d'avoir été là et de m'avoir remonté le moral à de nombreuses reprises, notamment lors de nos soupers « healthy » ou juste en me faisant des surprises avec des pâtisseries. 'Thom, merci de toujours me dire que tu es fier de ta grande sœur, de nos soirées, et surtout de toujours me suivre dans mes délires ! Bobonne, merci de m'avoir toujours soutenue, même si cela n'a pas toujours été facile pour toi ces dernières années… Sache aussi que je suis et resterai super fière de toi, ma petite sœur ❤️

Je remercie également tous les membres de ma famille qui m'ont toujours encouragée et soutenue depuis toute petite ! Mamy, Papy, j'espère que, de là-haut, vous êtes fiers de votre petite-fille. J'aurais tant espéré que vous soyez là.

Merci également à mes amis. Aurélie, merci d'être là et surtout d'être toi ! Je chéris nos souvenirs de secondaire, nos compétitions de natation et notamment notre stage d'entraînement parfois un peu douteux, ahah ! Je repense aussi avec tendresse à nos soirées films dans ton kot, et à nos moments/discussions autour d'une (ou plusieurs) délicieuse pâtisserie. Merci aussi d'avoir été là quand j'étais au plus mal. Merci aussi à toi Augustin ! Malgré nos agendas chargés, je chéris chaque moment de retrouvailles où nous rattrapons le temps écoulé. Merci également à Blandine, Lisa, Maïlys et Marie pour votre précieuse amitié !

Au total, j'aurai écrit 55 fois « merci » dans cette partie remerciements, mais je pense que cela ne sera jamais suffisant. Alors juste encore une fois, mille mercis à tout le monde (vraiment !), votre soutien infaillible m'a permis d'avancer, de ne pas perdre pied et de mener à bien ce gros projet !

Emilie

Remerciements - Acknowledgements

Summary

The intertidal zone is a unique environment where both chemical and physical conditions can change considerably. Many marine organisms have adapted to this harsh environment by developing adhesive systems, notably the blue mussel (*Mytilus edulis*) and the honeycomb worm (*Sabellaria alveolata*). The mussel secretes numerous proteinaceous threads, collectively called byssus, each ending in a plaque that facilitates adhesion to the substrate. On the other hand, the honeycomb worm lives in a tube it constructs by cementing sand grains together, then lining it internally with an organic sheath. The adhesion mechanisms of mussels and tubeworms have been studied for several decades. They involve adhesive proteins rich in post-translationally modified amino acids. These modifications result in functional groups that are crucial for adhesion and their production is often facilitated by specific enzymes that are poorly investigated.

Both model organisms have a DOPA-based adhesive system. DOPA (3,4-dihydroxy-L-phenylalanine) is produced by the post-translational hydroxylation of tyrosine residues within the adhesive proteins catalyzed by tyrosinase enzymes. The first part of the thesis focuses on these tyrosinase enzymes. Through proteo-transcriptomic analyses, we identified tyrosinase enzymes potentially involved in DOPA production and/or oxidation in the adhesive proteins of both organisms. Their roles were subsequently confirmed through localization via *in situ* hybridization, showing a high gland-specificity in mussels but not in tubeworms. Phylogenetic analyses were also conducted to address a long-lasting question regarding potential similarities in the adhesion mechanisms of these organisms. These analyses support the hypothesis of independent expansions and parallel evolution of tyrosinases involved in adhesive protein maturation in both lineages, supporting the convergent evolution of their DOPA-based adhesion.

One of the tyrosinase candidates of the blue mussel, involved in byssal plaque formation, was then selected for further investigation. To gain deeper insights into its enzymatic function, this tyrosinase was recombinantly produced in bacteria with the aim of assessing its enzymatic activity by measuring DOPA production using different substrates and conditions. However, although the enzyme was successfully produced, it could not be purified in a soluble form.

Another crucial enzyme involved in marine adhesion is kinase. The honeycomb worm adhesive comprises two serine-rich proteins in which most serine residues are converted to phosphoserines. We hypothesized that this phosphorylation of serine residues is catalyzed by

Summary

FAM20C kinase enzymes. We identified these enzymes through *in silico* analyses and confirmed their expression in the adhesive glands via *in situ* hybridization. Additionally, a better characterization of the honeycomb worm adhesive system was provided through a proteomic analysis of granules extracted from the cement glands, as well as morphological and elemental composition studies.

Finally, the tube lining, another important secreted system found in the honeycomb worm, has been studied. We examined the composition and ultrastructure of this complex fibrous sheath using electron microscopy, along with proteomic and molecular analyses. The morphology of the glands secreting this tube lining has also been investigated. The tube lining is a fibrous structure consisting of multiple layers of parallelly organized fibres embedded in a matrix. Three main types of secretory cells secrete this sheath: collagen-secreting, acidic mucopolysaccharide-secreting, and catechol-secreting cells. We identified eight main proteins within the tube lining. Interestingly, most of these proteins do not contain any known conserved functional domains nor share homology with any proteins from public databases. One of them, however, contains a collagen domain which shows some similarities with bacterial collagen triple helix repeat proteins.

Together, the works presented in this thesis significantly contribute to the understanding of marine adhesive systems, particularly the permanent adhesion of mussels and tubeworms. This knowledge may advance the design of new adhesive biomimetic materials, notably DOPA-based adhesive materials.

Table of contents

Chapter I

Chapter II

Chapter III

Chapter IV

Abbreviations

BLAST: Basic Local Alignment Search Tool

BLOSUM: BLOcks SUbstitution Matrix

Bp: Base pair

BSA: Bovine Serum Albumin

BUSCO: Benchmarking Universal Single-Copy Orthologs

CDD: Conserved Domain Database

CLANS: CLuster ANalysis of Sequences

DIG-dUTP: Digoxigenin (DIG)-11-deoxyuridine triphosphate (dUTP)

DOPA: 3,4-dihydroxyphenylalanine

DTT: Dithiothreitol

ECM: ExtraCellular Matrix

EDTA: Ethylenediaminetetraacetic acid

EDX: Energy-Dispersive X-ray spectroscopy

EGF-like: Epidermal Growth Factor

FDR: False Discovery Rate

FPKM: Fragments Per Kilobase Million

GuHCl: Guanidinium Chloride

HCl: Hydrochloric Acid

HPLC: High-Performance Liquid Chromatography

IB: Inclusion Bodies

IgG: Immunoglobulin G

IPTG: Isopropyl-β-D-thiogalactopyranoside

ISH: *In Situ* Hybridization

kDa: Kilodalton units

LB: Luria-Bertani broth

LC-MS/MS: Liquid Chromatography coupled to tandem Mass Spectrometry

LDL: Low-Density Lipoprotein

Mfp: Mussel foot protein

MW: Molecular Weight

NaCl: Sodium Chloride

NBT/BCIP: Nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt

NCBI: National Center for Biotechnology Information

OCT medium: Optimal Cutting Temperature medium

Abbreviations

ORF: Open Reading Frame

Osm: Osmoles

P.A.S.: Periodic Acid Schiff

PBS: Phosphate-Buffered Saline

PCR: Polymerase Chain Reaction

pI: Isoelectric point

psu: practical salinity unit

PTM: Post-Translational Modification

PTMP: Proximal Thread Matrix Proteins

RT: Room Temperature

SDS: Sodium Dodecyl Sulfate

SDS-PAGE gel: Sodium Dodecyl Sulfate–polyacrylamide gel

SEM: Scanning Electron Microscope

TEM: Transmission Electron Microscope

TMP: Thread Matrix Proteins

Tris: Tris(hydroxymethyl)aminomethane

General Introduction

Insights into Enzymatic Processes Involved in Marine Invertebrate Adhesive Systems: A Comprehensive Review

1. Preamble

Marine adhesion, a phenomenon encompassing all the glues secreted by invertebrate organisms in the sea, has always been and remains a challenge to understand. Marine invertebrates secrete chemical bioadhesives fulfilling various functions, such as allowing them to remain attached to surfaces and withstand the hydrodynamic forces of waves in the intertidal zone, to collect or capture their food, or to avoid being thwarted by predator attacks. Adhesion is therefore essential for them to survive in their environment. Because marine organisms possess the remarkable ability to adhere to a wide range of substrates quickly and firmly underwater, a feat that most current artificial chemical adhesives fail or struggle to achieve, their adhesive secretions have been studied for several decades. Understanding the mechanisms behind such adhesion could greatly aid in the development of biomimetic adhesives for use in industrial and medical fields, particularly in wet environments (Kamino, 2008; Lee *et al*., 2011; Zhong *et al.,* 2014; Zhao *et al*., 2016; Almeida *et al*., 2020). Furthermore, this understanding could contribute to the advancement of new antifouling technologies (Del Grosso *et al*., 2016; Amini *et al*., 2017).

Currently, aquatic species that use bioadhesive secretions have been identified in 28 metazoan phyla out of the 34 currently described (Delroisse *et al*., 2022). Both the evolutionary background and the biology of species influence the specific composition of attachment systems in each organism (Waite, 1983a; Gorb, 2008; Hennebert *et al*., 2015a; Li *et al*., 2021). Marine adhesion comes in such varied forms that it has been categorized into several types according to its time of action and composition; adhesion to the substratum can be instantaneous, permanent, temporary, or transitory (Flammang *et al*., 2005). Those characteristics are derived from the physicochemical properties of the adhesive proteins, notably from their post-translational modifications. These modifications of adhesive proteins fall into three main types: hydroxylation, phosphorylation, and glycosylation (Sagert *et al*., 2006; Hennebert *et al*., 2011; Stewart *et al*., 2011a; Petrone, 2013; Cui *et al*., 2017; Davey *et al*., 2021). However, aided by Omics technologies (Davey *et al*., 2021), many other modifications have been identified. All those modifications are often made possible thanks to specific enzymes. In parallel, surrounding factors such as bulk seawater, oxidative species, and microbial communities can diminish underwater adhesion through dilution, oxidation, and

degradation of biological holdfasts. However, certain marine organisms exhibit resistance to these effects through mechanisms that usually also involve enzymes (Liang *et al*., 2019a). Although numerous studies have explored the diversity of marine adhesive proteins, little or nothing is known about the enzymes involved in their maturation, including post-translational modifications. Yet, these modifications, occurring either before or after secretion, are important and allow the adhesive materials to fulfill their function. In this introduction, I first present the diversity of the adhesive systems found in marine invertebrate animals and then catalogue all the enzymes known to be involved in the post-translational modifications of their adhesive proteins. This review highlights the importance of these enzymes as well as the inherent challenges and suggests priorities for future research.

What is an enzyme?

It was in 1878 that Wilhelm Kühne used the word 'enzyme' for the first time, a term derived from the Greek words 'en', meaning 'within', and 'zume', meaning 'yeast'. Enzymes are defined as biological catalysts that speed up biochemical reactions in living organisms without being consumed during the reaction. They can catalyze the conversion of only one type or a range of similar types of substrate molecules into product molecules and are only required in very low concentrations. They are often found as proteins, but some enzymes are also RNAs, the so-called ribozymes, and play an important role in gene expression (Robinson, 2015). The Enzyme Commission of the International Union of Biochemistry classifies enzymes into six major groups according to the reactions catalyzed, e.g., transferases for enzymes that transfer an atom or a group. Enzymes are given specific nomenclature and are described by a four-part Enzyme Commission (EC) number (http://www.chem.qmul.ac.uk/iubmb/enzyme/). The polypeptidic chain of the enzyme can fold to form a specific three-dimensional structure, incorporating a small area known as the active site. The conformation of this active site and its charge properties enable the enzyme to bind to a single type of substrate molecule, giving specificity to its catalytic activity. The catalytic activity can be characterized by different parameters, e.g., Kcat represents the number of substrate molecules that can be converted to product by a single enzyme molecule per unit time. Some enzymes need a cofactor, which is necessary for the enzyme's catalytic activity.

2. Diversity of marine invertebrate adhesion

2.1. Instantaneous adhesion: How to be sticky within seconds

Instantaneous adhesion is characterized by rapidity, within seconds, and the destruction of tissue integrity when the adhesive secretion is released. Currently, only two adhesive systems are classified in this category: the Cuvierian tubules of sea cucumbers, and the colloblasts cell found on the tentacles of ctenophores (Flammang *et al*., 2005; Von Byern *et al*., 2010).

2.1.1. Ctenophores: the self-destructive adhesive system of colloblasts

Ctenophores, also called comb jellies, are worldwide marine predators and form the most basal animal lineage in phylogenetic analyses (Moroz *et al*., 2014). They have special adhesive cells called colloblasts. Colloblasts can be found on the surface of the ctenophore tentacles (Fig. 1A), with the exception of *Euchlora rubra* (new accepted name *Haeckelia rubra* (Kölliker, 1853)), in which nematocysts replace colloblasts (Tregouboff & Rose, 1957). These cells are responsible for capturing prey by releasing a fast, strong, and potentially reversible glue directed at the target. When in contact with the prey, the clusters of adhesive-containing granules within the colloblast rupture, resulting in the release of their contents and subsequent self-destruction of the cell (Fig.1B) (Franc, 1978; Von Byern *et al*., 2010).

While the morphology of colloblasts has been extensively described (Fig. 1C) (Storch & Lehnert-Moritz, 1974; Benwitz, 1978; Mackie *et al*., 1988; Von Byern *et al*., 2010), less information is available regarding their molecular composition. The adhesive substance primarily consists of lipoproteins (Bargmann *et al*., 1972), neutral sugars, and neutral proteins (Von Byern *et al*., 2010). It is also believed to be associated with a collagen-like, proline-rich molecule (Townsend & Sweeney, 2019). A recent study found evidence for the presence of catecholic amino acids closely related to L-DOPA (or 3,4-dihydroxyphenylalanine) within the colloblast granules of the ctenophore *Pleurobrachia bachei* A. Agassiz, in L. Agassiz, 1860, also called the Pacific Sea gooseberry (Townsend & Sweeney, 2019). DOPA is produced by the post-translational hydroxylation of tyrosine residues by tyrosinase enzymes (Ramsden & Riley, 2014). Additionally, the genome of this organism was found to contain a tyrosinase enzyme, supporting the hypothesis of a catechol-mediated adhesion mechanism in ctenophores (Moroz *et al*., 2014; Townsend & Sweeney, 2019). A study based on transcriptomic analysis was also carried out on various ctenophores, identifying colloblast-specific candidate genes (Babonis *et al*., 2018). Among these colloblast candidates, many sequences associated with secretion/cell membranes, and enzymes involved in cellular metabolism were identified. Some

putative adhesive genes were also found and share common domains with known adhesive proteins from other invertebrate marine species (Babonis *et al*., 2018). Moreover, it was shown that these genes are preferentially expressed during tentacle morphogenesis, and that both colloblast and neuron cells differentiate from the same progenitor cell. The ctenophore tentacles may also secrete toxins along with its adhesive secretion (Babonis *et al*., 2018).

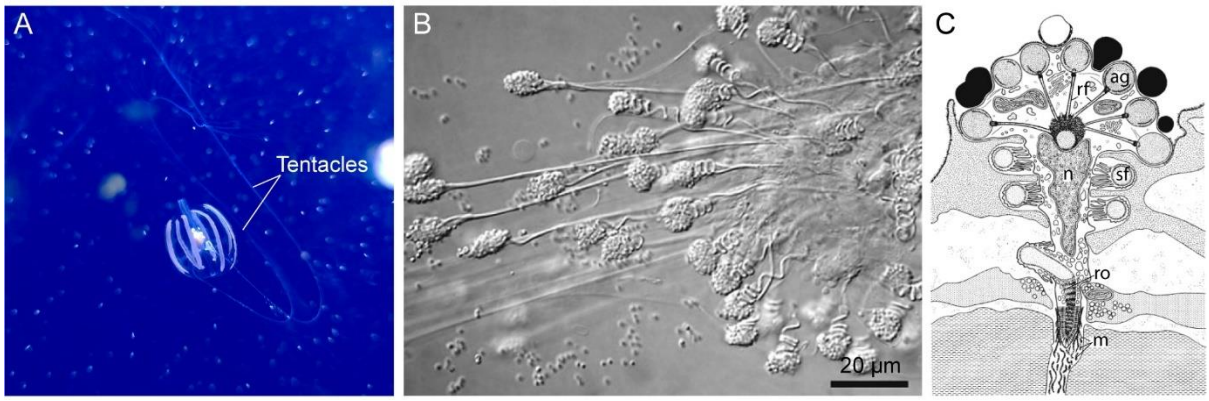


Figure 1. Picture of a comb jelly deploying its tentacles in the seawater (Picture courtesy of Noé Wambreuse) (A). Micrograph (B) and schematic drawing (C) of colloblasts. Legend: ag; adhesive granule; m – mesoglea; n – nucleus; rf – radiating fiber; ro – root; sf – spiral filament (adapted from Von Byern *et al*., 2010; Benwitz, 1978).

2.1.2. The particular Cuvierian tubules of sea cucumbers

Several species of Holothuroidea, all belonging to the order Holothuriida, are equipped with a particular defensive system (Flammang, 1996; DeMoor *et al*., 2003). When a sea cucumber perceives external pressure from a predator, it contracts its body and expels 10–20 elongated structures called Cuvierian tubules, triggered by an acetylcholine signaling system (Chen *et al*., 2023), through a tear in the cloacal wall (Fig. 2A). These tubules become sticky within seconds upon contact with any surface, entangling the predator (Hamel & Mercier, 2000; Becker & Flammang, 2010; Demeuldre *et al*., 2014). Before being expelled, Cuvierian tubules are white caeca attached to the basal part of the left respiratory tree with their distal parts floating freely in the coelomic cavity. They are composed of an inner epithelium surrounding a narrow lumen, a thick connective tissue layer, and finally, a mesothelium on the outside that contains both peritoneocytes and granular cells organized in superimposed layers (Fig. 2B) (Becker & Flammang, 2010; Demeuldre *et al*., 2014).

An early study showed that proteolytic treatment of these tubules in the species *Holothuria forskali* Delle Chiaje, 1824 decreases their adhesive strength, indicating that their adhesive is mainly composed of proteins (Müller *et al*., 1972). In this species, which is predominantly found in Europe and has been extensively studied, both peritoneocytes and

granular cells exhibit extensive immunoreactivity to antibodies raised against Cuvierian tubule adhesive material, and stain strongly for proteins, tyrosine residues, and disulfide bonds (Endean, 1957; DeMoor *et al*., 2003; Becker & Flammang, 2010). Moreover, the granular cells were tested positive for anti-phosphoserine antibodies, suggesting the presence of phosphoserine residues (Flammang *et al*., 2009). A study on the material left on the substratum after the mechanical detachment of the tubule collagenous core, called tubule prints, showed that this material is primarily composed of organic material, mainly proteins and carbohydrates in a 3:2 ratio by dry weight, with no lipids detected. Proteins extracted from the tubule prints are predominantly rich in glycine and acidic residues. SDS-PAGE gel analysis revealed the presence of 10 major protein species, with molecular weights ranging from 17 to 220 kDa (DeMoor *et al*., 2003). A more recent study using a proteo-transcriptomic approach identified 32 putative adhesive proteins, including one rich in glycine and one rich in cysteine residues (Bonneel, 2020). These appear to be the most abundant adhesive proteins, suggesting a two-component adhesive, likely related to the instantaneous adhesive system. Moreover, a C-type lectin was also identified, and numerous post-translational modifications, including N- and O-glycosylation and phosphorylation, were predicted in the two main adhesive proteins (Bonneel, 2020).

The tubule adhesive print was also studied in *Holothuria (Stauropora) dofleinii* Augustin, 1908, where seven protein bands were detected by SDS-PAGE gel (17–89 kDa). Via Edman degradation and mass spectrometry, several protein sequences were obtained, corresponding to various enzymes, two unannotated proteins, and a C-type lectin protein (Peng *et al*., 2011; Peng *et al*., 2014). Finally, a recent genomic study of the species *Holothuria leucospilota* (Brandt, 1835) found several Cuvierian tubules-specific proteins with unique long tandem-repeat sequences. These sequences are arranged to form intramolecular β-sheets that can be defined as a novel amyloid assembly. These functional amyloid-like fibrils are found on the surface of Cuvierian tubules (Fig. 1C).

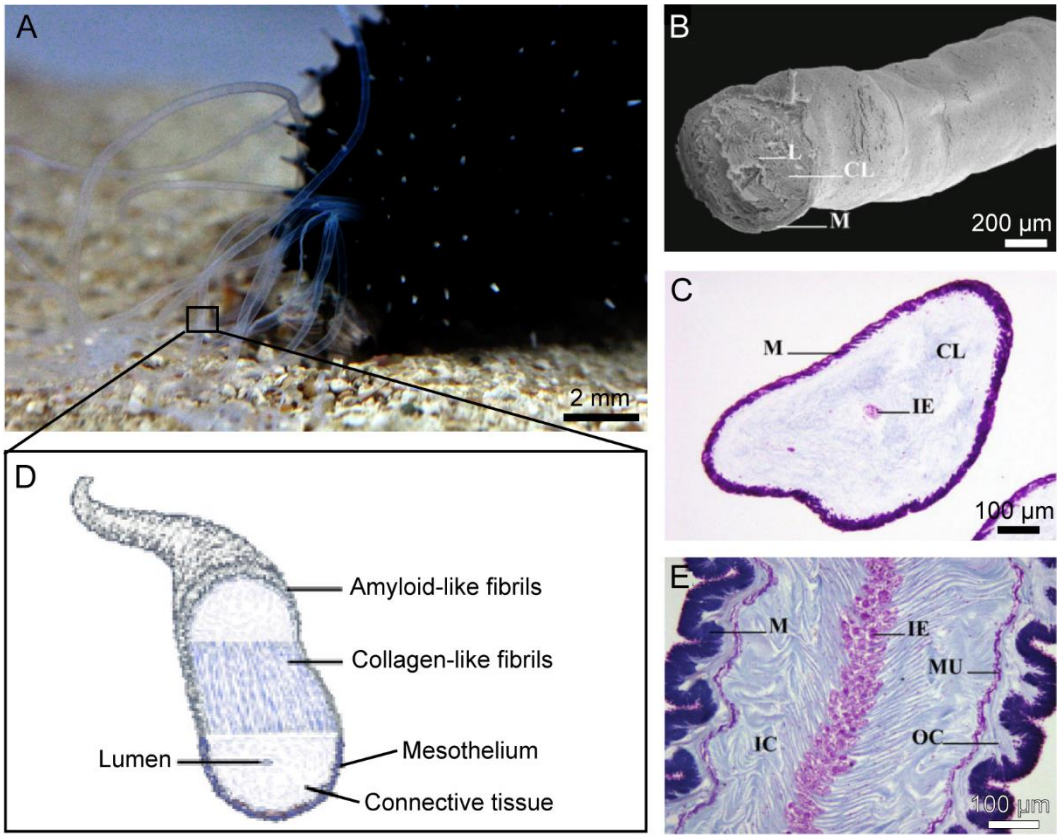


Figure 2. Picture of the sea cucumber *Holothuria forskali* with its Cuvierian tubules (A). SEM image of a tubule of *Holothuria (Stauropora) pervicax* Selenka, 1867 (B). Histological staining of a transverse section of a Cuvierian tubule of *Holothuria leucospilota* (Brandt, 1835) (C). Illustration of the composition of a Cuvierian tubule (D) and a longitudinal section of the tubule of *Bohadschia marmorata* Jaeger, 1833 (E). Legend: CL – connective tissue layer; IC – internal connective tissue layer; IE – inner epithelium; L – lumen; M – mesothelium; MU – muscles; OC – outer connective tissue layer (from Becker & Flammang, 2010; Chen *et al*., 2023).

2.2. Permanent adhesion: The adhesion of sessile marine organisms

Permanent adhesion is found mostly in sessile organisms. Although many phylogenetically diverse animals possess such adhesion, the characterization of such holdfast adhesion has mainly been limited to three organisms: the barnacle, the tubeworm, and the mussel (Kamino, 2008). All three primarily use a complex polyprotein cement, with the latter two relying mostly on DOPA-based and phosphoserine adhesion. This is not the case for barnacles, which differ from all other permanently gluing marine animals studied to date (Whittington & Cribb, 2001; Flammang *et al*., 2005; Jonker *et al*., 2012).

2.2.1. Molecular composition of adult barnacle cement

The life cycle of barnacles begins with planktonic larval stages, which include the nauplius (I-VI) and cyprid stages. During the latter, the larva actively explores substrates in search of a suitable place to settle down and undergo metamorphosis. Temporary adhesion is utilized

during this stage (Aldred & Clare, 2008; Aldred *et al*., 2010; Rosenhahn & Sendra, 2012). Once a suitable substrate is found, the barnacle larva secretes a permanent adhesive called cyprid cement to firmly attach itself and initiate metamorphosis into a juvenile. Eventually, it becomes an adult and establishes colonization. This attachment is long-lasting, even after the barnacle's death (Fig. 3A) (Saroyan *et al*., 1970). The adult uses a thin layer of cement to firmly attach itself to the substrate (Fyhn & Costlow, 1976). The animal first releases a fluid rich in lipids and oxidants to prepare the substrate (Fears *et al*., 2018), and then secretes cement through an extracellular canal system originating from unicellular glands, called cement gland, which are clustered in the basal region of the barnacle, close to the substrate. The barnacle cement proteins (BCPs) are synthesized in these glands and stored in large vacuoles (Fig. 3B) (Lacombe & Liguori, 1969; Lacombe, 1970; Walker, 1970; Power *et al*., 2010). There are two types of barnacles, stalked and acorn, both displaying a similar cement apparatus with minor differences due to their body structure (Power *et al*., 2010). However, the cement protein of acorn barnacles and its molecular function has been more studied than that of stalked barnacles.

The acorn barnacle cement is primarily composed of proteins (90%) (Kamino *et al*., 1996), including amyloid fibers rich in beta-sheet structures (Barlow *et al*., 2009, Barlow *et al*., 2010), but also contains carbohydrates, lipids, and inorganic ash (He *et al*., 2018). By using a high concentration of denaturing and reducing agents with heat treatment, 10 different BCPs have been identified in the barnacle *Megabalanus rosa* Pilsbry, 1916 (Kamino *et al*., 2000). These BCPs are named according to their approximate molecular weight estimated by SDS-PAGE gel. Cp20k, categorized as rich in cysteine and charged amino acids, may adhere via non-covalent interactions (Kamino, 2001). It was also proposed that Cp20k coordinates with Ca^{2+} and sugar chains of Cp52k in barnacle cement amyloid fibers (Liang *et al*., 2019a). Cp19k and Cp68k likely have both cohesive and adhesive roles and are grouped together based on biased amino acid composition (rich in serine, threonine, glycine, alanine, lysine, and valine). Cp100k and Cp52k may play a cohesive role by self-assembling into amyloid fibrils and are classified as hydrophobic proteins (Kamino, 2006; So *et al*., 2017). The BCPs were also characterized in *Amphibalanus amphitrite* (Darwin, 1854), in which a new protein named AaCP43, which has a theoretical MW of about 43 kDa but migrates unusually to around 63 kDa during electrophoresis, was identified (So *et al*., 2016). Several barnacle species, including stalked barnacles, have shown similar cement composition, except for Cp20k, which only exists in barnacles with a calcified baseplate (Fig. 3B) (Naldrett, 1993; Kamino *et al*., 1996; Jonker *et al*., 2014; Kamino, 2016; Liang *et al*., 2019a).

Introduction

For the stalked barnacle *Pollicipes pollicipes* (Gmelin, 1791 [in Gmelin, 1788-1792]), the cement and the gland have been characterized through proteomic analysis (Domínguez-Pérez *et al*., 2020; Domínguez-Pérez *et al*., 2021). The gland proteome was dominated by structural proteins and proteins involved in muscle contraction, as well as Cp100k and some uncharacterized proteins. The proteome of the cement revealed that half of the cement is composed of proteins that are not canonical cement proteins, but rather unannotated proteins. The major components of the cement proteome included BCPs, adhesion proteins of the extracellular matrix (ECM) and membranes, protease inhibitors, and proteases. One-third of the cement proteome corresponded to bulk adhesive proteins, with CP52k being the most abundant (Domínguez-Pérez *et al*., 2020). The proteome of the gland and cement of another stalked barnacle, *Lepas (Lepas) anatifera* Linnaeus, 1758, showed similar results. Most of the proteins present in the cement of *L. anatifera* were BCPs, with Cp-100k and Cp-52k being the most expressed bulk proteins, while Cp-43k-like, Cp-19k, and Cp-20k were the most expressed interfacial proteins, contrary to what was observed in the species *P. pollicipes*. New BCPs were also discovered (Domínguez-Pérez *et al*., 2021). Cp-20k was also found in both stalked barnacle species, indicating that this protein may not be exclusively found in calcareous-base acorn barnacles (Domínguez-Pérez *et al*., 2020; Domínguez-Pérez *et al*., 2021). Comparing the cement composition of acorn and stalked barnacles reveals similarities in the amino acid composition and type of bulk proteins, as well as in the glycosylation process. However, some differences in cement composition still exist between these barnacles, suggesting that the evolution of barnacle adhesive may have been constrained during species radiation (Schultzhaus *et al*., 2021).

As mentioned earlier, barnacle adhesion does not seem to involve significant post-translational protein modifications, only Cp52k is N-glycosylated and its function is primarily based on non-covalent intermolecular interactions (Kamino *et al*., 2012). A proteinase was detected in the barnacle cement, which is similar to human transglutaminase and bovine trypsin, suggesting that the mechanism of barnacle cement curing is a form of wound healing and is comparable to that of blood coagulation (Dickinson *et al*., 2009). Another study shows the involvement of peroxidase and lysyl oxidase enzymes in the barnacle *A. amphitrite* (So *et al*., 2017). In this same study, a novel full-length peroxinectin (AaPxt-1) was identified and thought to be responsible for oxidizing non-peptidyl catechol precursors present in cement layers into active semi-quinone. This semi-quinone would then react with the free amine of the BCPs, causing stable quinone protein crosslinking (Fig. 3C-E) (So *et al*., 2016; So *et al*., 2017).

Moreover, a major cement protein component, AaCp43, was found to contain ketone/aldehyde modifications via 2,4-dinitrophenylhydrazine (DNPH) derivatization (So *et al*., 2017). It was also found that the cement of barnacle is rich in phosphoproteins at the interface (Dickinson *et al*., 2016). However, this and the potential involvement of a quinone cross-linking process, such as the one found in sandcastle worms and mussel adhesion (see hereunder), in the curing of barnacle cement is still controversial and experimental evidence is lacking (Clancy *et al*., 2016; Liang *et al*., 2019a; Xu *et al*., 2022).

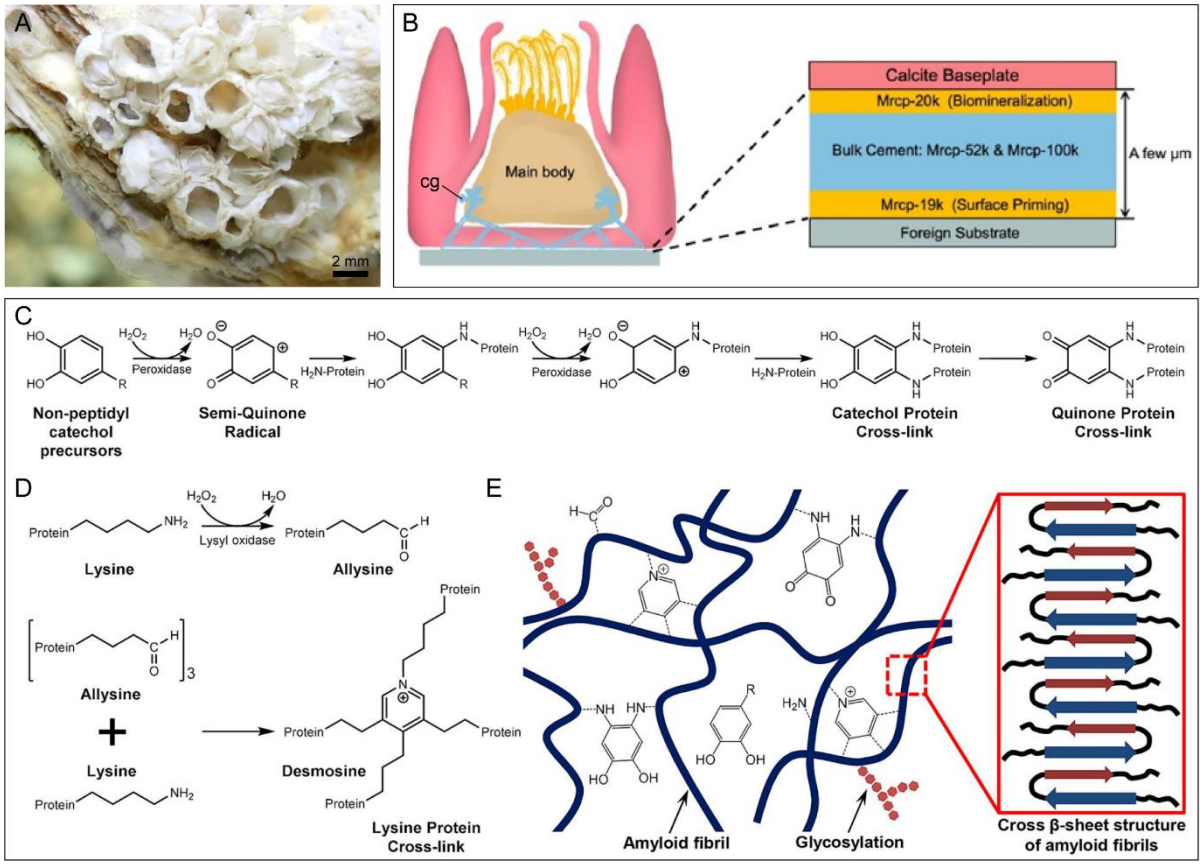


Figure 3. Picture of several acorn barnacles attached on the shell of an oyster (A). Illustration of the thin cement layer found between the substrate and the calcified base of the barnacle, showing its different BCP components (B). A focus on two described molecular processes of the cross-linking of barnacle cement. First, non-peptidyl catechol precursors are oxidized by peroxidase into quinones. These quinones then react with the free amine groups of BCPs to form stable quinone protein cross-links (C). Secondly, lysyl oxidase converts the lysine residues of the BCPs into allysine, which further reacts with other lysine residues to form multivalent cross-links (D). Together, these elements cross-link with the self-assembled amyloid fibrils BCPs to form interwoven fibers (from So *et al*., 2017; Liang *et al*., 2019a; Mondarte *et al*., 2023). Legend: cg – cement gland.

2.2.2. The tube-building worms and their remarkable adhesive cement

Some polychaete worms can build their own tubes, protecting themselves from the hydrodynamic forces of the waves. In the family Sabellariidae, worms catch particles in the

water column and use a special cement to hold them together, creating a secure home. In some cases, the tubes of different individuals are closely imbricated, resulting in reef-like structures that can extend over several hectares (Fig. 4A) (Vovelle, 1965; Noernberg *et al*., 2010). Two tube-building worm species, *Sabellaria alveolata* (Linnaeus, 1767), which is found along the European coasts (Vovelle, 1965; Gruet, 1972), and *Phragmatopoma californica* (Fewkes, 1889), which is located off the American west coast (Abbott & Reish, 1980), have received attention due to their impressive adhesive abilities. Both species possess a unique organ known as the building organ, which secretes a protein-based cement capable of hardening in less than 30 seconds (Fig. 4B) (Stevens *et al*., 2007). This organ consists of two external appendages, or lobes, enclosing cellular processes that originate from two types of cement glands: cement gland containing homogeneous granules and cement gland containing heterogeneous granules filled with inclusions (Fig. 4C-D). The cell bodies of the cement glands are located in the parathoracic region, around the worm's gut and within its parapodia (Vovelle, 1965; Wang & Stewart, 2012).

The adhesion in both species relies primarily on three main types of adhesive proteins categorized as (1) GY-rich proteins, with sequences containing many glycine and tyrosine residues; (2) H-repeat proteins, which contain several histidine, glycine, and tyrosine residues; and (3) poly(S) cement proteins containing a significant amount of serine. The first two types of adhesive proteins are basic whereas the third one, due to the phosphorylation of serine into phosphoserine (between 60–90 mol% serine), tends to be very acidic (Stewart *et al*., 2004; Zhao *et al*., 2005). Another important post-translational modification in these tubeworms involves the hydroxylation of most tyrosine residues into DOPA (or 3,4-dihydroxyphenylalanine) in the first two categories of adhesive proteins by a tyrosinase enzyme (Waite *et al*., 1992). These adhesive proteins display highly repetitive and blocky primary structures (Zhao *et al*., 2005; Endrizzi & Stewart, 2009; Becker *et al*., 2012; Stevens *et al*., 2007; Wang & Stewart, 2013). In the case of *P. californica*, a higher number of adhesive proteins with characteristics of structural glue proteins have been identified (Endrizzi & Stewart, 2009). There is also a final category of adhesive proteins that contain low sequence complexity cement proteins and do not fall into the first three categories, classified as eclectic miscellaneous proteins. These proteins contain high amounts of glycine, tyrosine, lysine, histidine, alanine, or serine residues (Endrizzi & Stewart, 2009; Buffet *et al*., 2018). These modified amino acids, along with DOPA, play a crucial role in the adhesion and cohesion of both Sabellariidae species (Zhao *et al*., 2005; Sun *et al*., 2007; Stewart *et al*., 2011b). Within

the cement glands, the adhesive proteins are condensed into secretory granules through a process called complex coacervation, wherein oppositely charged proteins and divalent cations (such as Ca^{2+} and Mg^{2+}) coalesce. Secretion is accompanied by a pH change from 5 (the pH within the cells) to 8 (the pH of seawater), resulting in the formation of an adhesive with a foam-like structure (Fig. 4E) (Stewart *et al*., 2004; Stevens *et al*., 2007).

One tyrosinase enzyme responsible for producing DOPA has been characterized in *P. californica* and is expressed in both types of cement glands (Wang & Stewart, 2012). In *S. alveolata*, 50 tyrosinase sequences were identified, along with a peroxidase and BPTI/Kunitz domain-containing proteases. Several enzymes contributing to the DOPA metabolic pathway were also revealed (Buffet *et al*., 2018). Another study revealed the presence of 2-chloro-DOPA, a halogenated derivative of DOPA, in the cement of *P. californica* (Sun *et al*., 2009). This compound has a lower pKa and thus is as good as DOPA for strong surface interactions, as it can stabilize metals by coordination at lower pH. It was also proposed that chloro-DOPA could play a role in defending the cement against microbial fouling and degradation, as well as in surface interactions (Sun *et al*., 2009).

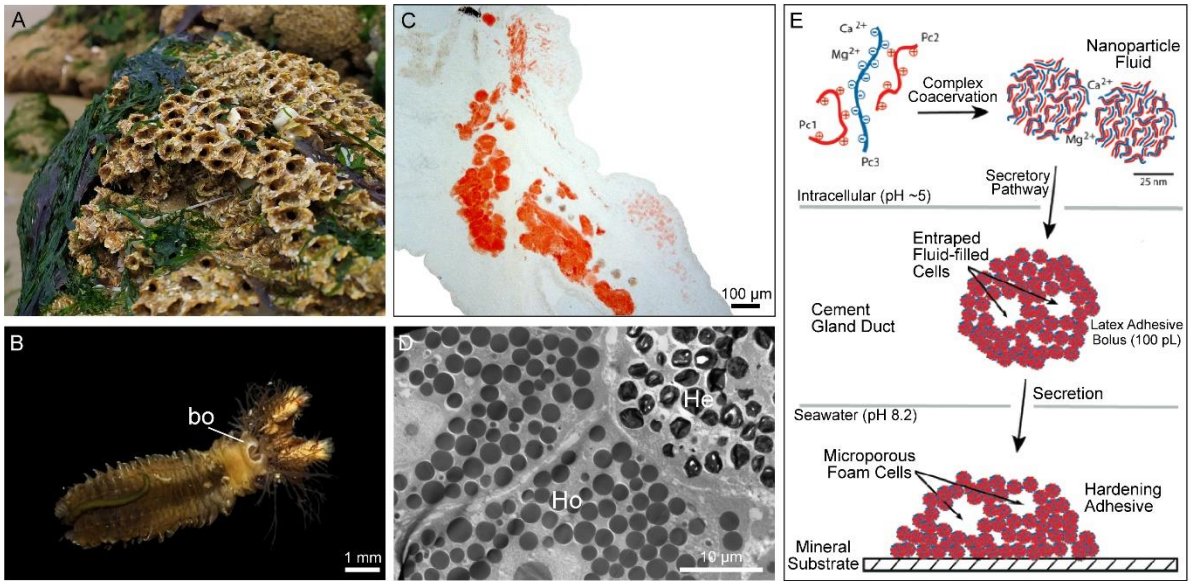


Figure 4. Picture of a part of a *Sabellaria alveolata* reef (courtesy of Alexia Lourtie) (A) and of *S. alveolata* with its building organ secreting the cement highlighted (B). Arnow staining shows the DOPA content of the two types of adhesive glands: the cement gland containing homogeneous granules and the one containing heterogeneous granules (C). A TEM image shows both types of cement glands, where the sub-granules of the heterogeneous granules can be distinguished (D). These cement glands contain oppositely charged adhesive proteins, as well as divalent cations. The interaction of these components leads to a coacervation phenomenon. A change in pH during secretion allows the formation of a foam-like, hard structure (E) (from Stevens *et al*., 2007). Legend: bo – building organ; He – adhesive gland containing heterogeneous granules; Ho - adhesive gland containing homogeneous granules.

2.2.3. The complex structure and function of mussel byssus

Mussels inhabit the intertidal zone where they face the challenge of staying attached to substrates to avoid being dislodged by the waves. To achieve this, mussels secrete a set of proteinaceous thread collectively called byssus (Waite, 1983a). The complexity of the byssus has been studied for decades, notably in members of the genus *Mytilus* (*Mytilus edulis* Linnaeus, 1758 (Fig. 5A), *Mytilus galloprovincialis* Lamarck, 1819, and *Mytilus californianus* Conrad, 1837) and in the species *Perna viridis* (Linnaeus, 1758). The adhesive properties developed by these species' byssi are among the strongest known in invertebrate marine organisms and have inspired research into the development of biomimetic adhesives for medical and industrial applications (Harrington *et al*., 2018).

The byssus is composed of several byssal threads, between 50-100 threads each 2-6 cm in length, terminating in a flattened plaque that allows adhesion to the substrate. Each byssal thread is covered by a thin, strong coating called the cuticle. The byssal threads are produced by the self-assembly of several proteins present in secretory vesicles that are stockpiled in specialized glands: the plaque gland (or the phenol gland), the core gland (also called the collagen or white gland), and the cuticle gland (also called the accessory or the enzyme gland) (Tamarin & Keller, 1972; Tamarin *et al*., 1976; Price, 1983; Waite, 2017). These glands release their content into a narrow groove running along the mussel's foot and in its distal depression to form new byssal threads (Fig. 5C).

The threads contain a unique type of collagen called pre-pepsinized collagens, or preCOLs. These collagens are distributed differently along the threads (Mascolo & Waite, 1986), categorized into distal (D), proximal (P), and nongradient (NG) distributions, known as preCOL-D, preCOL-P, and preCOL-NG, respectively (Waite *et al*., 2002). PreCOLs are trimeric, featuring super secondary triple helical collagen cores flanked by domains resembling silk-like β-sheets in preCOL-D, disordered elastin in preCOL-P and flagelliform silk in preCOL-NG (Qin & Waite, 1995; Coyne *et al*., 1997; Qin *et al*., 1997; Waite *et al.,* 1998). All preCols N- and C-termini are rich in histidine that form metal coordination complexes, which mechanically enhance fiber toughness and self-healing properties through reversible sacrificial bonds (Schmitt *et al*., 2015b). These preCOLs are found with thread matrix proteins (TMPs), a family of proteins rich in glycine, tyrosine, and asparagine, as well as proximal thread matrix proteins (PTMPs) which seem to be glycosylated by α-D-manno- or α-D-glucopyranosyl residues (Sagert & Waite, 2009; Arnold *et al*., 2013; Suhre *et al*., 2014). TMPs form a

viscoelastic matrix that supports the collagenous microfibrils during tension-induced deformation.

The mussel foot proteins (Mfp) are specific proteins produced in the plaque gland. More than 20 mfps have been identified in the plaque and cuticle glands (Fig. 5B) (DeMartini *et al*., 2017). The best characterized are mfp-2 to mfp-6, most of which are rich in glycine, are intrinsically disordered, and contain abundant DOPA (or 3,4-dihydroxyphenyl-L-alanine) (Waite, 2017). Mfp-3 and mfp-5 are responsible for plaque adhesion as they form bonds with substrates, notably due to their DOPA content. Mfp-5 also contains some phosphoserine residues in its sequence (Waite & Tanzer, 1981; Vreeland *et al*., 1998; Waite & Qin, 2001; Zhao & Waite, 2006; Zhao *et al*., 2006; Danner *et al*., 2012). Mfp-2 is a cohesive protein that forms links within the adhesive material itself, while mfp-4, a histidine-rich protein, links the plaque to the thread (Vreeland *et al*., 1998). Finally, mfp-6 is responsible for the reducing activity of the plaque thanks to the thiol group of its cysteine residues, preventing the oxidation of DOPA to dopaquinone which would otherwise inhibit the adsorption of the other mfps on the surface (Zhao & Waite, 2006; Nicklisch *et al*., 2016). Once all the mfp-containing vesicles of the plaque gland are secreted into the distal depression of the foot, they undergo condensation as well as fluid–fluid phase separations, forming a solid porous material, the plaque (Waite, 2017; Renner-Rao *et al*., 2022). The adhesive properties of mfps are highly dependent on post-translational modifications, and these high rates of modification suggest an enzyme-dependent co- or post-translational processing (Waite, 2017). The most abundant modified amino acid is DOPA. DOPA can form hydrogen bonds, covalent cross-links, electrostatic interactions and coordination complexes with metal ions on surfaces, as well as forming bonds with the adhesive material itself, increasing the cohesive strength of the adhesive material (Fig. 5D) (Lee *et al*., 2011; Petrone, 2013; Waite, 2017; Priemel *et al*., 2020). Other modifications occur, such as the production of 7-hydroxytryptophan and 4-hydroxyarginine (Zhao *et al*., 2006; Silverman & Roberto, 2007; Zhao *et al*., 2009), or the hydroxylation of proline residue to form hydroxyproline in the cuticle and the byssal thread collagen (Waite, 1983b; Mascolo & Waite, 1986). All these additional hydroxyl groups could increase hydrogen bonding with surfaces (Silverman & Roberto, 2007).

The third main part of the byssus is a 2-5 µm thin protective proteinaceous cuticle that is extensible yet hard, thanks to metal-coordination links such as DOPA-Fe^{3+} or -V^{4+} complexes (Holten-Andersen *et al*., 2007; Holten-Andersen *et al*., 2009; Schmitt *et al*., 2015a). The secretory vesicles from the cuticle gland undergo a maturation process resembling phase

separation during secretion to finally form a prefabricated granule surrounded by a matrix (Harrington *et al.*, 2010). Mfp-1 is found inside the cuticle gland, contains many DOPA residues, and up to 60% of all its amino acids are affected by post-translational modifications (Waite, 2017). Mfp-1 is secreted along with other mfps that are characterized as cysteine-rich (Jehle *et al.*, 2020).

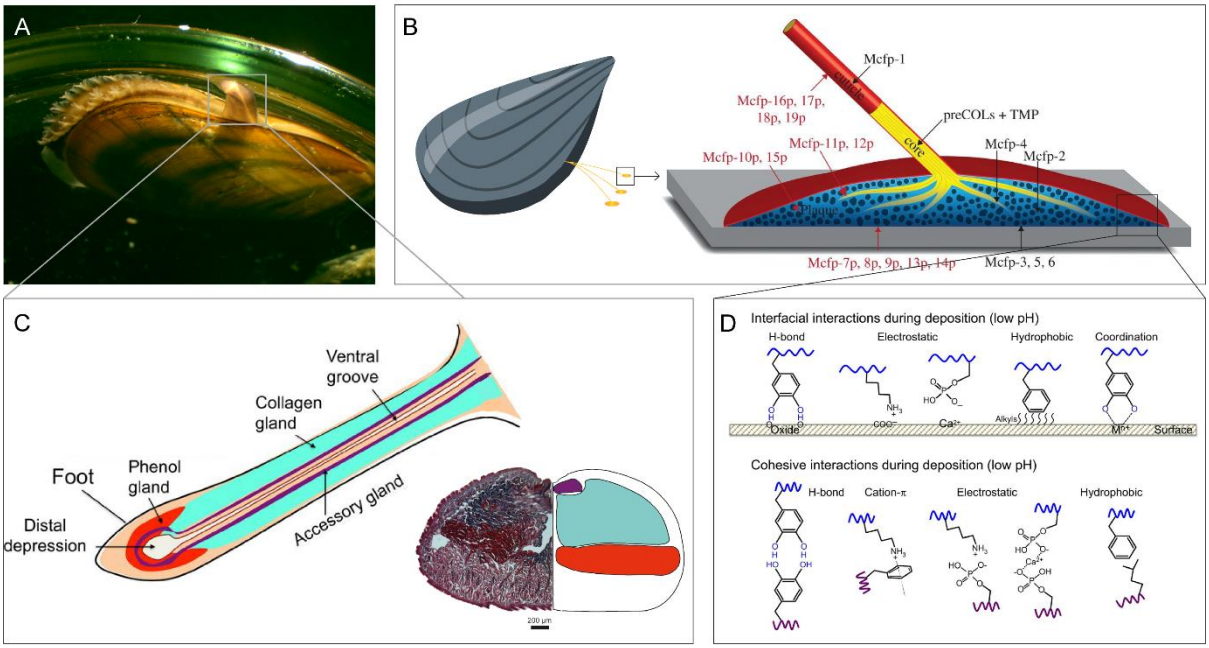


Figure 5. Picture of a blue mussel extending its foot (A), along with an illustration of its byssus showing the different mussel foot protein compositions (B). Illustration showing the different glands found inside the foot of the blue mussel: the collagen gland in blue that secretes the threads, the plaque gland in red that secretes the plaque, and finally, the cuticle gland in purple that secretes the cuticle. Additionally, a transverse histological section of a foot stained with Heidenhain's azan stain is included (C). Image showing the different adhesive and cohesive bonds made by DOPA and other amino acids at low pH (D) (from DeMartini *et al.*, 2017; Waite, 2017).

2.2.4. Examining the Byssal Threads of Other Molluscs

Another bivalve mollusc able to produce and secrete a byssus is for example the pearl oyster *Pinctada fucata* (A. Gould, 1850). Its byssus differs from the one of the mussel as it consists of 3-10 threads, each 1-2 cm in length. Byssal threads can be divided into three regions: the proximal and distal parts of the threads, and the terminal plaque that allows adhesion to surfaces (Fig. 6A-B) (Liu *et al.*, 2015). For some other oyster species, the byssus is secreted only during the larval stage. This byssus is also secreted by the foot, but its morphology is slightly different. It was shown in the larva of the European flat oyster *Ostrea edulis* Linnaeus, 1758, that the foot comprises nine different gland cells, some of which open into the byssus duct and discharge their contents during the last larval stage for settlement (Cranfield, 1973).

A study on the oyster *Pinctada fucata* showed that the proximal region of its byssus is composed of foam-like granules with hollow spaces around them (Liu *et al*., 2015). The distal region, responsible for its extensibility, constitutes 80% of the total byssal thread length. The extensibility of the byssus can be attributed to the interaction between Ca^{2+} and the vWFA domain of a thrombospondin-1 containing protein (TSP-1). This interaction increases the β-sheet conformation of the byssal proteins and stabilizes the nanocavities between the protein fibrils in the core. The protein fibrils are more compact in the cuticle (Fig. 6C). By extracting proteins from the distal region of the byssus and identifying the SDS-PAGE bands through LC-MS/MS analyses, 14 different proteins were identified. These included proteins with repeated low-complexity domains, foot proteins 1 (P-FU1), P-FU2, calcium-transporting ATPase, and ATPase (Liu *et al*., 2015). More recently, a transcriptomics and proteomics study identified 16 proteins in the adhesive plaque of the byssus of *P. fucata*, where TSP-1 was found. It is speculated that the vWFA domain-containing protein plays a structural role in the adhesives together with TSP-1 proteins. Proteins without BLAST homology were described as *Pinctada* unannotated foot proteins (PUF3-6), and other proteins such as tyrosinase, mucin-like proteins, and protease inhibitors were also identified (Liu & Zhang, 2021).

Another *Pinctada* species, *Pinctada maxima* (Jameson, 1901), was also investigated (Whaite *et al*., 2022). This oyster is famous for its high-quality pearl production. By performing proteo-transcriptomic analyses of native dried byssal threads, a proteome of 49 proteins was obtained, of which perlucin-like foot protein (Pmfp1), Balbiani ring-like protein 3 (BR3), peroxidase and thrombospondin 1 were identified as being highly expressed, and the presence of some of them was confirmed in the foot by immunodetection. TSP1-like and TSP1-domain-containing peptides were also found to be abundant in the *P. maxima* proteome. The foot tissue of this oyster also exhibited high expression of glycoprotein-like sequences, as well as a byssal metalloproteinase inhibitor and a sequence containing a tyrosinase domain (Whaite *et al*., 2022). A byssus-producing scallop species, *Chlamys farreri* (K. H. Jones & Preston, 1904), was also studied. Based on a proteo-transcriptomic study, seven foot-specific scallop byssal protein (Sbp) components were identified, most of which are rich in lysine and arginine. Several phosphorylation sites were also predicted (Miao *et al*., 2015).

Another bivalve species that produces a byssus is *Pinna nobilis* Linnaeus, 1758, which is a very large mussel, growing up to 1 meter in length, that lives partially buried in sandy areas of the Mediterranean Sea. Unlike other bivalve species, it possesses long and thin byssal threads that act like anchoring plant roots. The ultrastructure of its byssus was investigated and

shows that it is comprised of a peculiar organization of globular proteins, densely packed and arranged helically into nanofibers dispersed in a matrix. *Atrina pectinata* (Linnaeus, 1767) possesses the same kind of byssus ultrastructure, but the nanofibrils group into bundles in the byssus of *A. pectinata*, while the nanofibers are more dispersed in the byssus of *P. nobilis* (Pasche *et al*., 2018). The ultrastructure of the byssus of the giant clam *Tridacna maxima* (Röding, 1798) was also investigated and consists of a bundle of hundreds of individual threads, each 100 µm in diameter, and is composed of proteins arranged in a four-stranded alpha-helical coiled coil. Its amino acid composition features unusually high levels of acidic and basic residues, indicating that the four-helix structure is stabilized by strong ionic interactions (Miserez *et al*., 2012). However, for these three species, no molecular data is currently known for their byssal content, with the exception of a foot protein found in *A. pectinata*, called apfp-1, which contains DOPA and a sugar-binding domain. The combination of these two has been shown to play an important role in providing a robust load-bearing device (Yoo *et al*., 2016).

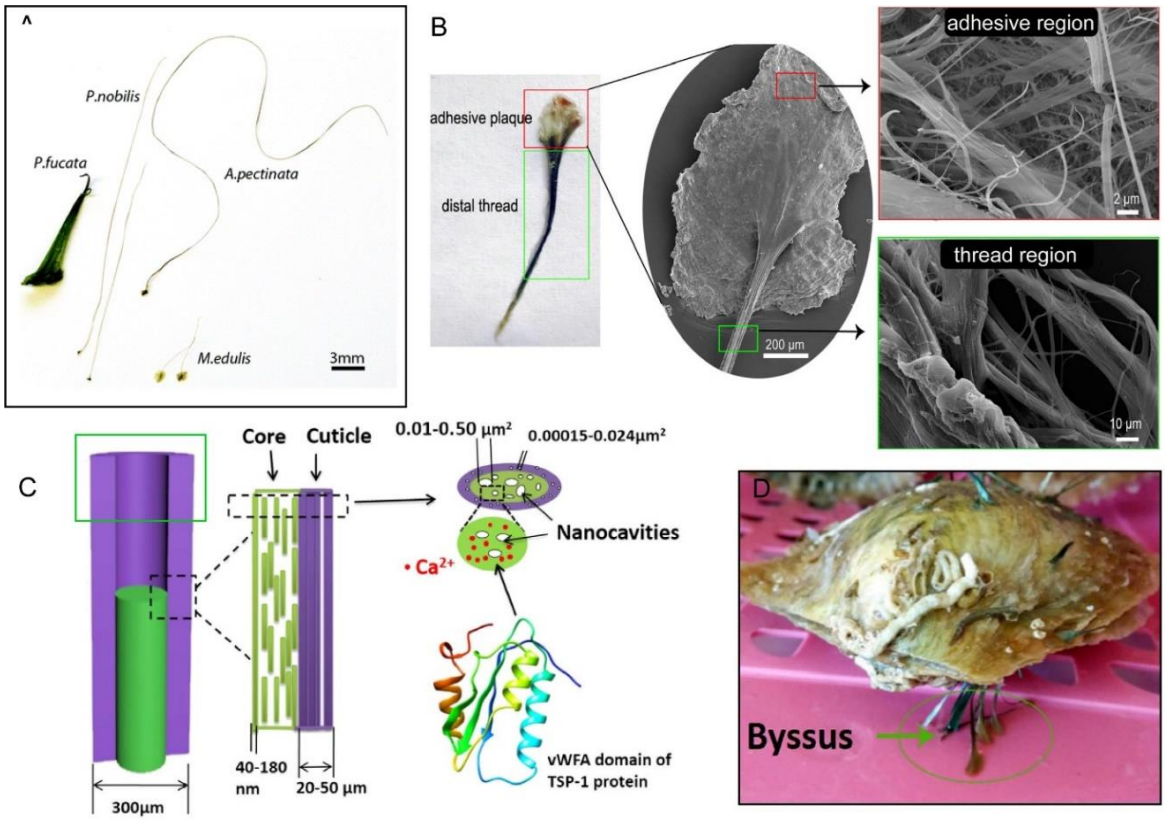


Figure 6. Morphology of three different bivalve byssal threads (*P. fucata*, *P. nobilis*, and *A. pectinata*) compared to that of the blue mussel *M. edulis* (A). Pictures of the *P. fucata* byssal threads with SEM images of the distal part and adhesive plaque of its byssus (B). In the core of *P.fucata* byssal threads, the protein fibrils are separated by nanocavities, whereas in the cuticle, the fibrils are compact. Interaction between Ca^{2+} and the vWFA domain of TSP-1 protein enhances the β-sheet conformation

of byssal proteins, stabilizing the nanocavities between fibrils (C). Picture of the byssus of a pearl oyster (D) (from Liu *et al*., 2015; Pasche *et al*., 2018; Liu & Zhang, 2021).

2.2.5. Adhesion mechanisms in ascidians

Ascidians, or sea squirts, are marine biofouling invertebrates that, during their juvenile stage, secrete adhesive proteins from their adhesive papillae, located at the anterior end of the head, to achieve permanent adhesion during metamorphosis. This initial adhesion is called “rapid adhesion.” Then, vascular-like ampullae form near the epidermis of the attached juvenile and spread out over the substrate, so that "slow adhesion" can subsequently occur (Torrence & Cloney, 1981; Ueki *et al*., 2018). Their adhesive papillae are composed of collocytes, which are elongated mucus-secreting cells, and of sensory and axial columnar cells (Cloney, 1977; Dolcemascolo *et al*., 2009). Once adults, they are tightly attached to the substrate and can reattach with varying tenacity depending on the type of substrate, through the growth of new tissue (Fig. 7B) (Edlund & Koehl, 1998).

Only a few studies have been conducted on the adhesive system of the adult sea squirt compared to the other permanent adhesives from marine invertebrates. A study on the ascidian *Styela plicata* (Lesueur, 1823), shows that its adhesive is composed of several proteins, which constitute the highest content, as well as polysaccharides and lipids (Saad *et al*., 2017). Adult *Ciona robusta* Hoshino & Tokioka, 1967, have also been studied and can tightly adhere to underwater surfaces thanks to a holdfast structure known as the stolon. This stolon has exceptional strength, allowing them to withstand significant changes in marine environments. Using a proteomic approach, 26 representative stolon proteins were identified, including ascidian stolon protein 1 (or ASP-1), which contains a von willebrand factor type A (vWFA) domain proven to be involved in adhesion by recombinant protein experiments. It was also suggested that the adsorption of ASP-1 is DOPA-dependent (Li *et al*., 2019). Another quantitative proteomics study on this fouling ascidian identified 16 adhesive protein candidates with increased expression in the stolon. Among these, ascidian adhesive protein 1 (AAP1) was identified as an important interfacial protein capable of cross-linking via hydrogen bonds and salt bridges among vWFA domains (Fig. 7B) (Li *et al*., 2024).

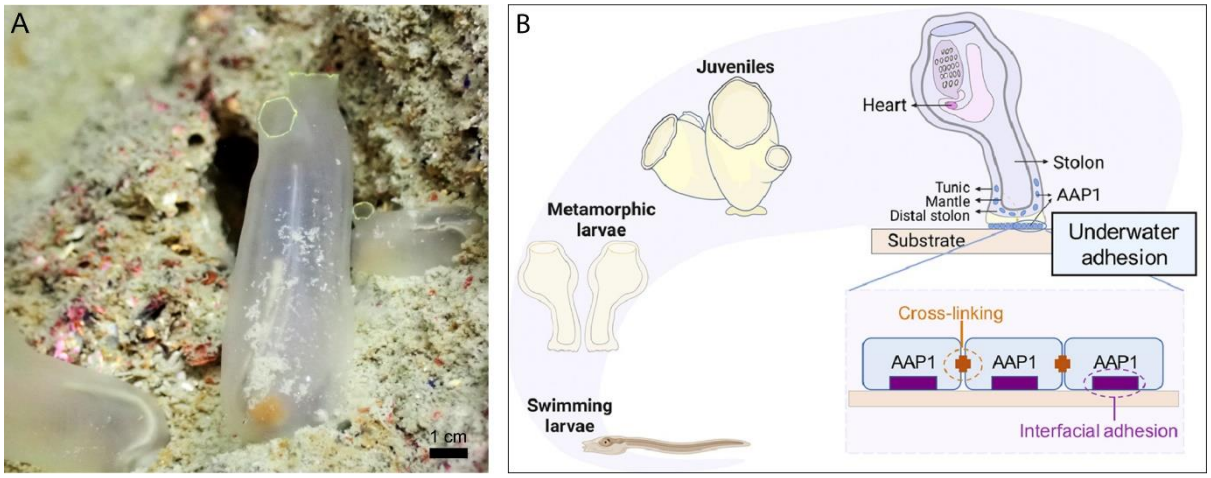


Figure 7. Picture of *Ciona robusta* (A) with an illustration of its different life stages, and its AAP1-mediated stolon permanent adhesion once adult. AAP-1 is an interfacial protein capable of cross-linking via hydrogen bonds and salt bridges among vWFA domains (B) (from Li *et al*., 2024).

2.3. Temporary adhesion: an attachment-detachment mechanism

Many small invertebrates living in the interstitial spaces of soft sediments, as well as some macro-invertebrates such as certain cnidarians and most echinoderms, secrete an adhesive material that allows them to stay firmly attached to the substrate, but only momentarily, as they can detach later on. This “attach/detach repeatedly” mechanism is a type of adhesion called temporary adhesion (Walker 1987; Flammang, 1996; Flammang *et al*., 2005). The temporary adhesive system comprises two kinds of gland cells and is termed the "duo-gland system" where one cell type produces the adhesive secretion and the other produces a secretion that causes detachment (Tyler, 1976; Whittington & Cribb, 2001). Different types of detachment mechanisms have been proposed: a "competition model," where a secretion outcompetes the binding between the adhesive layer and the adhesive organ surface; an "enzymatic model," involving detachment through enzymes; and a "mechanical model," where detachment occurs by the action of muscular contractions, proposed for animals without an additional secretion to facilitate detachment (Lengerer & Ladurner, 2018). The attachment-detachment mechanism usually leaves an adhesive "footprint" on the surface, with a topography similar in most temporarily adhering animals (Lengerer *et al*., 2018a).

2.3.1. Exploring flatworm adhesion mechanisms

Free-living marine flatworms from the turbellarian orders Haplopharyngida, Macrostomida, Polycladida, Rhabdocoela, and Proseriata possess duo-gland systems (Tyler, 1976). One or more adhesive gland cells with electron-dense granules form the adhesive, while one or more releasing gland cells possess smaller, less dense granules (Lengerer *et al*., 2014). The flatworm

Introduction

Macrostomum lignano Ladurner, Schärer, Salvenmoser & Rieger, 2005 is commonly found in the meiofauna, living between the sand grains of the intertidal zone and is a well-studied model (Fig. 8A). It has about 130 adhesive organs localized at the tip of its tail, and each is composed of three cell types: an anchor cell, an adhesive cell, and a releasing cell (Fig. 8B-D). These flatworms attach themselves by pressing their tail on the surface. By conducting a specific transcriptome analysis of the posterior end of the animal, an intermediate filament-like gene, which was referred to as macif1, was identified in the anchor cells (Lengerer *et al*., 2014). In addition, later studies revealed two large proteins playing an essential role in the animal's adhesion, namely Mlig-ap1 and Mlig-ap2, for *M. lignano* adhesion protein 1-2, which are localized within the adhesive vesicles of the adhesive cells. Mlig-ap2 is proposed to displace water molecules and attach to the surface, while Mlig-ap1 is a cohesive protein that connects Mlig-ap2 to the microvilli of the anchor cell (Fig. 8E). Moreover, Mlig-ap1 shares similar domains with the sea star cohesive protein sfp1. Both proteins have a high cysteine content, large repetitive regions, and several protein-carbohydrate and protein-protein interaction domains. It was also shown that flatworm adhesion was not possible on highly hydrated surfaces, and a small negatively charged molecule that interferes with Mlig-ap1 may induce detachment (Fig. 8F) (Wunderer *et al*., 2019).

The adhesive system of the flatworm *Minona ileanae* Curini-galletti, 1997, which also lives in the sand of the intertidal zone and possesses characteristics like those of *M. lignano* adhesive organs, was also studied. A differential transcriptome analysis between intact and tail-amputated flatworms, along with the use of Oxford Nanopore sequencing of gDNA, allowed the identification of five adhesive proteins. Among them, Mile-ap1 (*M. ileanae* adhesive protein 1) showed high similarity to Mlig-ap1 with respect to the presence of protein domains (C-type lectin domain (CTL), a von Willebrand type D domain (vWD), a C8-domain (C8), and a trypsin inhibitor-like domain (TIL)). One transcript, Mile-if1 (*M. ileanae* intermediate filament protein 1), shows homology with the *M. lignano* anchor cell-specific intermediate filament protein Mlig-if1 and is expressed in the anchor cells. It was also shown that several different proteins from *M. ileanae* correspond to one single *M. lignano* proteins; for example, Mile-ap1 had a very high resemblance to the core region of Mlig-ap1 with an identical order of protein domains, while the KR-rich regions of Mlig-ap1 were part of the proposed Mile-ap3a/b protein (Pjeta *et al*., 2019).

Another study on the adhesive system of the small sand-dwelling flatworm, *Theama mediterranea* Curini-Galletti, Campus & Delogu, 2008, identified and characterized 15

different genes, four of which are homologous to adhesion-related genes found in other flatworm species (Bertemes *et al*., 2022). Three adhesive proteins, named Tmed-ap1 to -3, were discovered, with one being O-glycosylated, along with a Kringle-domain-containing protein and a tyrosinase enzyme. However, no L-DOPA was detected in the footprint of this species, suggesting that the role of this enzyme is not the hydroxylation of tyrosine to L-DOPA, but rather a cross-linking role. A model for the detachment process of this flatworm was proposed, where the animal secretes the three adhesive proteins that bind to the surface. The tyrosinase may possibly crosslink these secreted proteins, and the lysine residue of the third adhesive protein interacts with the Kringle-domain-containing protein. Later, a negatively charged molecule is secreted and releases the animal from the surface by interacting with the positively charged protein Tmed-ap3, outcompeting the lysine residues that were bound to the Kringle domains (Bertemes *et al*., 2022).

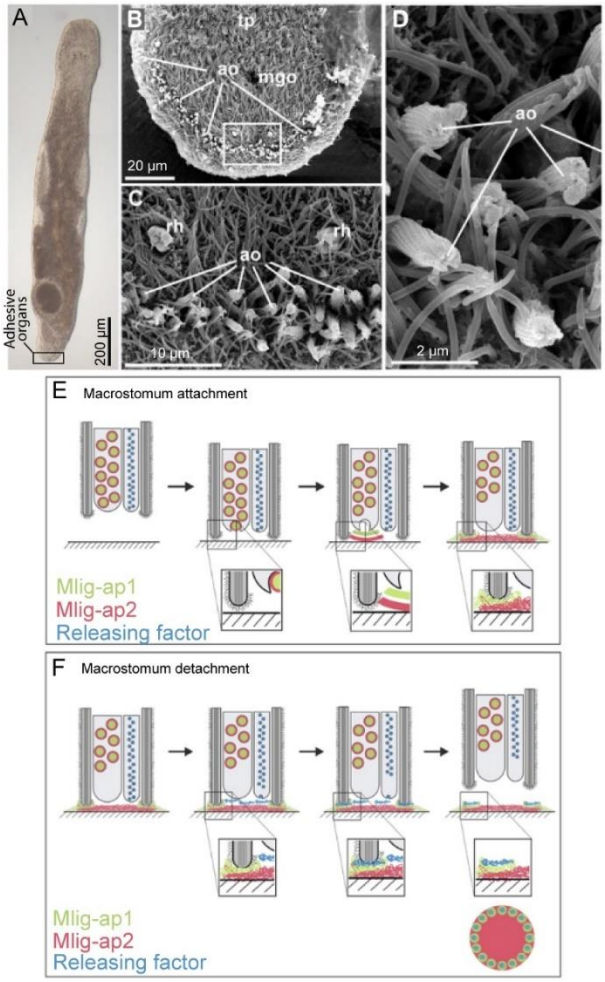


Figure 8. Picture of the flatworm *Macrostomum lignano* with a frame indicating the position of its adhesive glands in the tail plate (A). SEM images of its the tail-plate showing several adhesive organs (B) and a detailed view of some of them (C). Details of its adhesive organs reveal a microvilli collar on the anchor cells (D). An illustration of the attachment mechanism of a single adhesive organ shows that

upon attachment, *M. lignano* discharges releasing vesicles containing Mlig-ap1 and Mlig-ap2. Mlig-ap2 displaces the water and attaches to the surface, while Mlig-ap1 is a cohesive protein that connects Mlig-ap2 to the microvilli of the anchor cells (E). An illustration of the detachment mechanism shows that it is based on the release of negatively charged substances that react with Mlig-ap1, leaving a footprint on the surface (F) (from Lengerer *et al*., 2014 & Wunderer *et al*., 2019). Legend: ao - adhesive organ; mgo - male genital opening; tp – tail plate; rh – rhabdites.

2.3.2. Understanding sea star adhesion

Sea stars are marine organisms that display temporary but firm adhesion during locomotion, feeding, and burrowing (Flammang, 1996; Whittington & Cribb, 2001). It was in 1926 that a sticky secretion was first described as playing a role in the adhesion of the sea star, rather than suction, as was generally thought (Paine, 1926). For their adhesion, sea stars rely on numerous small appendages called tube feet or podia found along each arm on the oral side (Fig. 9A) (Flammang *et al*., 1994). The podia are made of two parts: the stem, which is connected to the arm and bears the tensions applied to the animal by external forces (Santos *et al*., 2009a), and the disc, which is the distal part in contact with the surface (Fig. 9B-D). The adhesive glands are found in the epidermis of the disc and comprise different cell types including two different types of adhesive gland cells (AC1 and AC2), and a de-adhesive gland cell (DAC) (Fig. 9G). The adhesive gland cells are distinguishable by the ultrastructure of their secretory granules (Flammang *et al*., 1994; Flammang *et al*., 1998). The podia are covered externally by a glycocalyx-like cuticle whose outermost layer, called “the fuzzy coat”, can protect against invading microorganisms (Ameye *et al*., 2000; Schröder & Bosch, 2016). Once the podia are detached, the adhesive material is left on the surface as a footprint (Fig. 9E) (Flammang, 1996; Santos *et al*., 2009a).

Asterias rubens Linnaeus, 1758 is the sea star that has been the most extensively studied. Its adhesive footprints are composed of a thin homogeneous layer produced by AC2 covered by a meshwork whose components are secreted by AC1 (Hennebert *et al*., 2008). By combining mass spectrometry analysis and lectin blots, researchers demonstrated that this secretion mainly comprises proteins, proteoglycans, and glycoproteins with oligosaccharide moieties including galactose, N-acetylgalactosamine, fucose, and sialic acid residues. These components are thought to confer both cohesion and adhesion through electrostatic interactions facilitated by the polar and hydrogen-bonding functional groups of their glycan chains (Hennebert *et al*., 2011). Thanks to proteo-transcriptomic analyses, 34 footprint protein candidates have been identified (Hennebert *et al*., 2015c), with 22 of them localized exclusively in the secretory cells of the disc epidermis (Lengerer *et al*., 2019). Sfp1, characterized as one

of the most abundant proteins in the footprint material, has been identified as a cohesive protein and localized in AC1 and the fibrillar meshwork of the footprint. This large protein, consisting of 3853 amino acids, has a high cysteine content prone to form disulfide bonds and undergoes autocleavage into four subunits before secretion. It also possesses domains that mediate protein-protein, protein-carbohydrate, and protein-metal interactions (Hennebert *et al*., 2014). One of the Sfp proteins contains an astacin-like domain, is secreted by DAC glands, and may act as a proteinase responsible for the detachment mechanism (Fig. 9F) (Hennebert *et al*., 2015c; Algrain *et al*., 2022).

Comparative studies based on morphological and immunohistochemical studies on adhesive material from different species of sea star, have shown a certain degree of conservation between species (Flammang *et al*., 1994; Santos *et al*., 2005; Lengerer *et al*., 2018a). The predicted cohesive protein (Sfp1) was showed to be conserved among distantly related sea star species, independently of inhabited environments and morphological adaptations of their tube feet (Lengerer *et al*., 2019). Furthermore, it has been demonstrated that in the adhesive system of another sea star species, *Asterina gibbosa* (Pennant, 1777), the presence of glycoconjugates with α-linked mannose residues within the secreted adhesive material is similar to that of *A. rubens*. However, the specific role of glycosylation and carbohydrate residues in the adhesion process of sea stars and most other temporary adhesive animals remains unknown (Lengerer *et al*., 2018a).

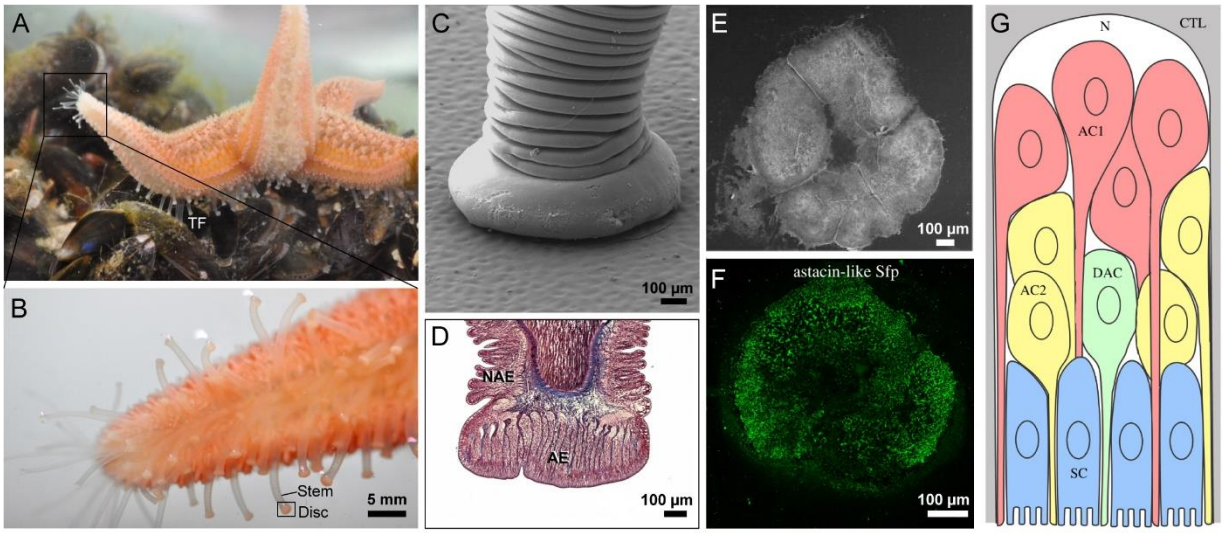


Figure 9. Pictures of *Asterias rubens* on a bed of mussels (A), with a zoom on its tube feet (B) (Photo courtesy of Amandine Deridoux), and an SEM image of an attached tube foot (C). Heidenhain's azan staining of a longitudinal section through a podium (D). SEM image of a footprint left by a podium of *A. rubens* once detached from the substrate (E), with an immunofluorescence localization picture of the footprints labeled with antibodies directed against Astacin-like Sfp (F). Illustration of a longitudinal section through a radial epidermal strip located between two adjacent connective tissue laminae,

showing the disposition of the different adhesive and deadhesive glands (from Hennebert *et al*., 2008; Hennebert *et al*., 2014; Algrain *et al*., 2022). Legend: AC1 - type 1 adhesive gland cell; AC2 - type 2 adhesive gland cell; AE – adhesive epidermis; CTL - connective tissue layer; DAC - de-adhesive gland cell; M - mucus gland; N - nerve plexus; NAE – nonadhesive epidermis; SC - support cell; TF – tube feet.

2.3.3. Investigating sea urchin attachment

Temporary adhesion is found in all echinoderms, but for practical reasons, only a few have been studied. Parallelly to sea stars, the adhesive system of sea urchins has been studied, notably the one from the adoral tube feet of *Paracentrotus lividus* (Lamarck, 1816). Only the tube feet located on the oral surface of the test are specialized in locomotion and temporary attachment (Fig.10A) (Santos *et al.,* 2013). Sea urchins can attach strongly to surfaces thanks to enlarged and flattened apical discs of these adoral tube feet (Märkel & Titschack, 1965; Santos & Flammang, 2007). Like the tube feet of asteroids, their adhesive epidermis is composed of four cell categories: support cells, sensory cells, one type of adhesive cells (compared to two in asteroids), and de-adhesive cells (Fig. 10B-C) (Santos *et al*., 2009a). Here also, the epidermis is covered by a well-developed cuticle made up of a multilayered glycocalyx (Ameye *et al*., 2000). The enzymatic activity of the de-adhesive secretions could be released within the cuticle and causes the removal of its outermost layer, the "fuzzy coat," leaving most of the adhesive material strongly attached to the substrate in the form of a footprint after detachment (Flammang & Jangoux, 1993; Flammang, 1996; Flammang *et al*., 1998). It was shown that the *P. lividus* footprint consists mainly of inorganic residues (45.5%) and only a few proteins (6.4%) compared to that of the sea stars (20.6%). It also contains neutral sugars (1.2%) and lipids (2.5%) (Santos *et al*., 2009b). The protein fraction had slightly more non-polar (57.4%) than polar (42.6%) amino acids, with equivalent amounts of charged (20.2%) and uncharged (22.4%) residues (Santos *et al*., 2009b).

A study conducted on the proteome of *P. lividus* identified a pool of 163 proteins overexpressed in the disc, including several enzymes involved in the adhesion of the sea urchin tube feet (Table 1). Most of the identified proteins are actins (27.9%), histones (24.4%), tubulins (11.9%), ribosomal proteins (7.9%), and myosins (1.4%). Other proteins, such as certain ECM components involved in tissue repair after detachment, were identified. As well as proteins involved in the citric acid cycle, which are important for the energy consumption required for adhesive material secretion (Lebesgue *et al*., 2016). Three Nectin-like proteins, which are overexpressed in the tube feet adhesive discs and the secreted adhesive footprint, are proposed to have a cohesive role as they have six galactose-binding discoidin-like domains that

can bind molecules bearing galactose and N-acetylglucosamine (Lebesgue *et al*., 2016; Ventura *et al*., 2023). Indeed, it has been previously suggested that a high molecular weight glycoprotein (>180 kDa) with oligomers of N-acetylglucosamine (GlcNAc) in the form of chitobiose is the main protein of the adhesive system (Fig. 10D). In addition, N-acetylglucosamine is also conjugated to two smaller glycoproteins (Simão *et al*., 2020). It is proposed that there are different variants produced, probably to cope with different substrate chemistry (Ventura *et al*., 2023). Toposome was also proposed as a potential adhesive protein (Santos *et al*., 2013). By conducting a proteome remapping of the previous results (Lebesgue *et al*., 2016), combined with a differential RNAseq analysis of RNA from tube foot disc and stem tissue, 16 putative adhesive or cohesive genes were identified and localized. Six of them showed similarity with sea star orthologue adhesive proteins (Pjeta *et al*., 2020). Moreover, it was shown that five of these proteins are glycosylated with N-acetylglucosamine and N-acetylgalactosamine residues (Ventura *et al*., 2023). Other proteins, such as protease inhibitors or Alpha-tectorin-like proteins, known to be involved in the high cysteine content of the adhesive material (Algrain *et al*., 2022), and different enzymes were found within the adhesive material (Fig. 10D) (Ventura *et al*., 2023). Finally, it was also shown that several proteases and glycosylases are potential components of the de-adhesive secretions (Lebesgue *et a*l., 2016; Lengerer & Ladurner, 2018).

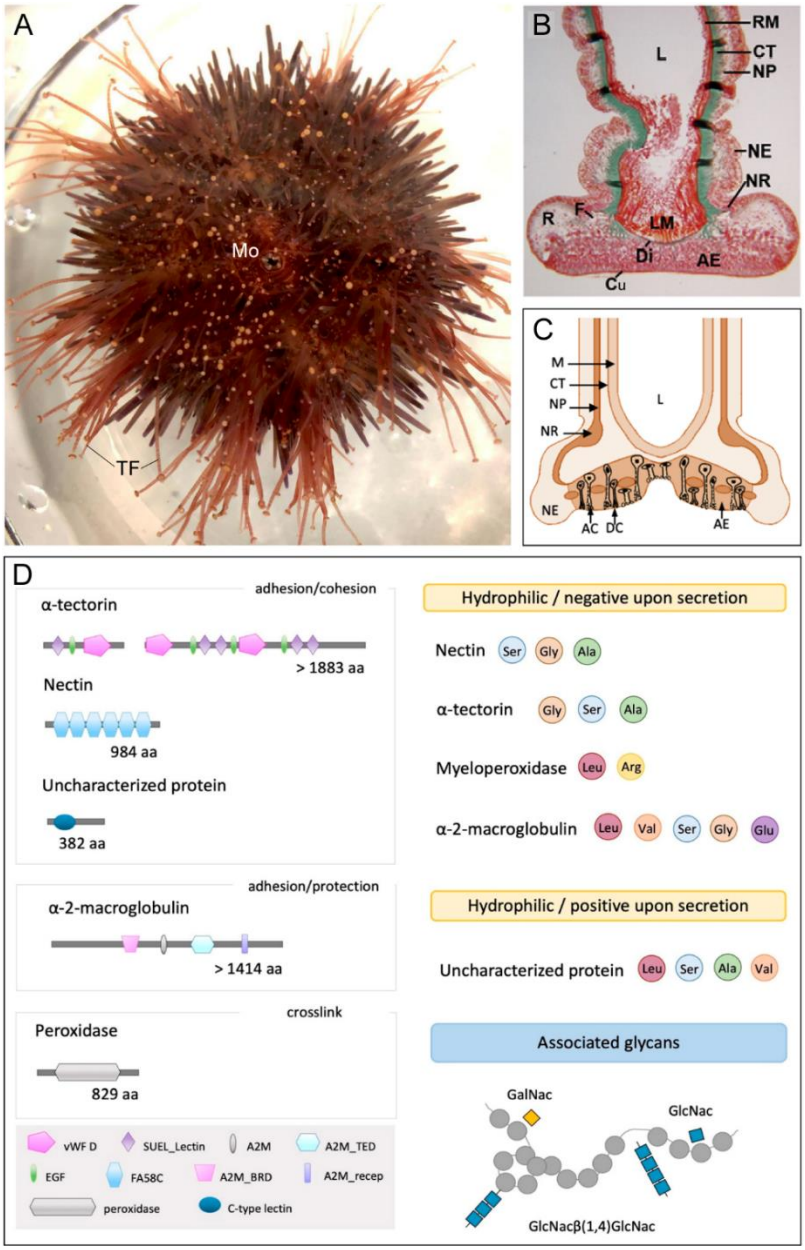


Figure 10. Picture of *Paracentrotus lividus* showing its oral surface with the tube feet specialized in locomotion and temporary attachment (photo courtesy of Inês Ventura) (A). Histological staining of a longitudinal section through a tube foot (B), and illustration of the duo-glandular adhesive epidermis with adhesive and de-adhesive cell glands present in the tube feet (C). Illustration showing the different adhesive and cohesive proteins identified in the *P. lividus* tube feet by lectin pulldowns, mass spectrometry, and *in silico* characterization (D) (from Santos *et al*., 2013; Pjeta *et al*., 2020; Ventura *et al*., 2023). Abbreviations: A2M - alpha macroglobulin domain; A2M_BRD - alpha macroglobulin bait region; A2M_recep - alpha macroglobulin receptor binding domain; A2M_TED - alpha macroglobulin receptor binding domain; EGF - EGF-like calcium-binding domain; FA58C - discoidin domains; GalNAc - N-acetylgalactosamine; GlcNAc - N-acetylglucosamine; Glc- NAcβ(1,4)GlcNAc - N-acetylglucosamine in a specific chitobiose arrangement; SUEL_Lectin, SUEL lectin domain; vWF D – von Willebrand factor type D domain. Legend: AC - adhesive secretory cell; AE - adhesive epidermis; CT - connective tissue; Cu, cuticle; Di - diaphragm; DC - de-adhesive secretory cell; F - frame ; L - lumen; LM - levator muscle ; M - myomesothelium; Mo – mouth; NE - non-adhesive epidermis; NP- nerve plexus; NR - nerve ring; R - rosette; RM - retractor muscle; TF - tube feet.

2.3.4. Other animals with temporary adhesion

Marine gastrotrichs are microscopic animals, measuring a few hundred micrometers, and can be found from the littoral region to the deep sea in the mesopsammon (Schmidt-Rhaesa, 2014). Except for the genus *Neodasys*, all gastrotrichs have a temporary adhesive system. Transmission microscopy analysis of several species showed the two gland cell types that are characteristic of the duo-gland adhesive system, comparable to that of flatworms (turbellarians). The characteristic of their adhesive system helps in their taxonomic placement (Tyler & Rieger, 1980). It has also been shown that locomotion in one species of nematode, *Theristus caudasaliens* Adams & Tyler, 1980, uses the tail musculature and caudal adhesive glands (Adams & Tyler, 1980). However not much is known about their adhesive system.

2.4. Transitory adhesion: being attached while moving

Transitory adhesion enables simultaneous adhesion and locomotion. The animals attach themselves by a viscous film deposited between their body and the substrate, and crawl on the film, which is left behind as they move (Walker, 1987; Flammang, 1996; Whittington & Cribb, 2001). The animal's movement is usually allowed by waves of muscular contractions travelling along the attached foot. This type of adhesion is characteristic of soft-bodied invertebrates such as sea anemones and gastropod molluscs.

2.4.1. Limpets and their powerful attachment

Limpets can actively attach to the substrate with remarkable force by using a suction attachment mechanism prior to locomotion at high tide, but they can also secrete a glue-like adhesion using an adhesive mucus for powerful long-term attachment at low tide (Smith, 1991; Smith, 1992). Several studies have been conducted to characterize their adhesive secretion using histochemistry and gel electrophoresis, which have shown that it is predominantly water-based (90-95% of its composition), with the remainder consisting of proteins, carbohydrates, and inorganic materials (Davies *et al*., 1990; Smith *et al*., 1999; Smith, 2016). The limpet *Lottia limatula* (P. P. Carpenter, 1864) secretes a specific form of slippery mucus for locomotion or for use during suction. The mucus can be modified with more protein and carbohydrate when the limpet needs a stronger attachment. There are no differences in the inorganic material or protein compositions between the two types of mucus, except for a protein of 118 kDa that was found only in the adhesive mucus, and a 68 kDa protein that was found only in the slippery, non-adhesive mucus (Smith, 1992; Smith *et al*., 1999).

Another species, the common limpet *Patella vulgata* Linnaeus, 1758 (Fig. 11A), can cling strongly to rocks and move around while being still attached, thanks to a special pedal mucus that allows the limpet to switch between temporary and transitory adhesion. In this limpet species, there is no sub-pedal lower pressure for attachment compared to other species; instead, they secrete a strong and reversible specialized pedal mucus for glue-like adhesion (Kang *et al*., 2020). This mucous-type adhesive forms a strong, reversible or degradable attachment for locomotion. Three types of mucus can be collected separately: IPAM ("interfacial primary adhesive mucus"), which is a thin layer left on the surface when the limpet is detached; BPAM ("bulk primary adhesive mucus"), which is the primary adhesive mucus secreted when the limpet is attached, directly collected from the pedal sole; and SAM ("secondary adhesive mucus"), which is the newly secreted mucus collected by wiping the pedal sole 30 minutes after limpet detachment (Fig. 11D) (Kang *et al*., 2020).

The limpet adhesive system resides in the foot, which can be divided into three distinct regions: the sole, the peripheral region, and the sidewall (Fig. 11B). The adhesive glands are located in the pedal sole, which is ciliated and separated from the peripheral region in the anterior third of the foot by the marginal groove. *P. vulgata* possesses nine different pedal glands, six of which secrete onto the sole (Grenon & Walker, 1978). A lectin-binding assay and transcriptome-assisted proteomics study identified 171 sequences from the three types of limpet pedal mucus, 27% of which were found in all three types. Among the identified proteins, there are proteins likely involved in oligomerization, ligand binding, peptide stabilization through disulfide bridges, enzymes, inhibitors, and proteins with dual functions. Five of these were categorized as enzymes, as they contain at least one enzymatic domain (sugar-cleaving glycoside hydrolase and metallopeptidase). Moreover, all detected proteins were predicted to have post-translational modifications (PTMs), with at least one type of glycosylation and numerous phosphorylation sites. The pedal mucus has molecular constituents similar to those of temporary bio-adhesives but with some distinctions. The lectin staining suggests that α-Mannose and α-Fucose linked to N-acetylchitobiose are specific to the pedal sole mucus, with a dorsoventral gradient staining in which longer subepithelial glands are stained for N-acetylgalactosamine or sialic acid. Shorter glands closer to the epithelium surface express other residues, such as mannose, fucose, chitobiose, sialic acid, or N-acetylglucosamine (Fig. 11C). This pattern may indicate that specific sugars are involved in transitory adhesion (Kang *et al*., 2020).

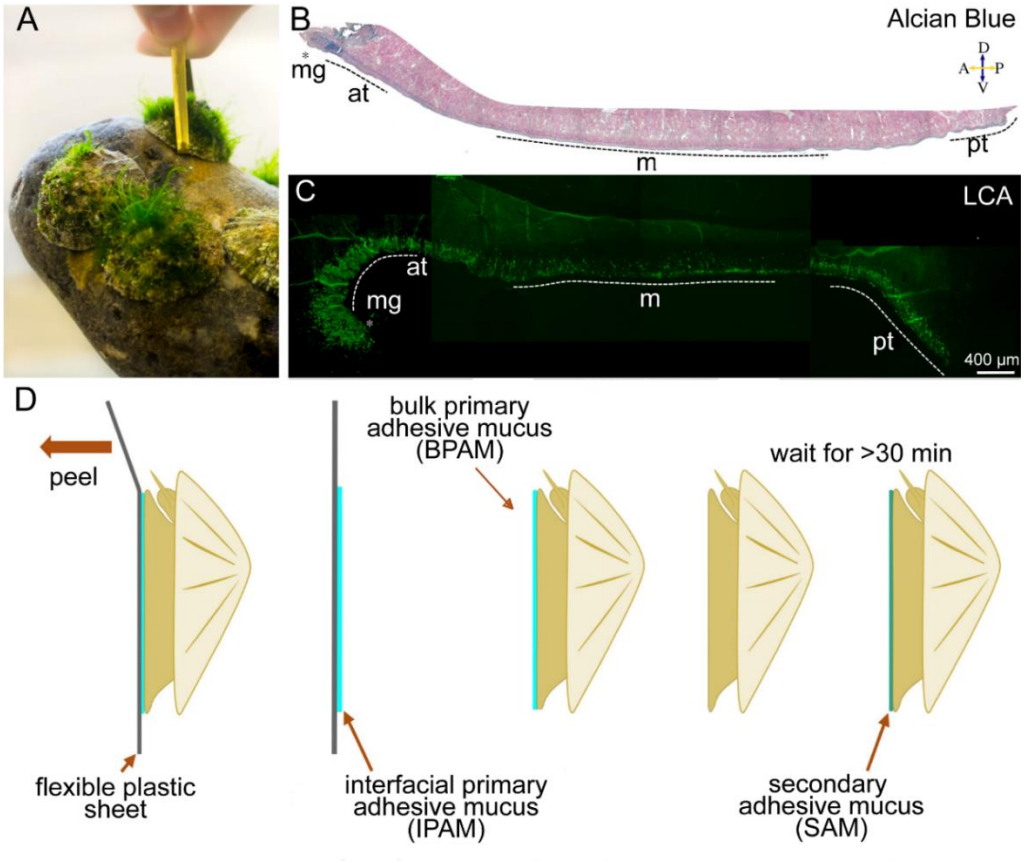


Figure 11. Picture of the limpet *Patella vulgata* (A) with Alcian blue staining of its foot tissue showing carboxylated and sulphated compounds in blue, and phloxine in red (B). LCA lectin staining of the foot tissue showing the part of the tissue with sequences containing α-linked mannose residues (C). Illustration of the three types of mucus collected in *P. vulgata* (D) (from Kang *et al*., 2020). Legend: at – anterior; m - middle; mg - marginal groove; pt - posterior.

2.4.2. Pedal disc adhesion in sea anemones

Sea anemones, belonging to the order Actiniaria, are predatory marine invertebrates that can settle on a variety of surfaces using their pedal disc. Most of them are solitary polyps attached to surfaces, and early observations of sea anemones noted their swimming movements and their ability to detach and resettle (Yentsch & Pierce, 1955; Hoyle, 1960; Ross & Sutton, 1964a; Ross & Sutton, 1964b; Ellis *et al*., 1969). Their body wall is made of three primary parts, from inside out: the endoderm, the mesoglea, and the ectoderm which is in contact with the outer environment. The adhesive secretion comes from the ectodermal part of the pedal disc (Clarke *et al*., 2020). Only a few studies exist on the molecular mechanism of the adhesion of cnidarian species, although it would be interesting to study this phenomenon from an evolutionary point of view, as this phylum diverged early from the other metazoans (Davey *et al*., 2019).

The sea anemone *Exaiptasia pallida* (new accepted name *Exaiptasia diaphana* (Rapp, 1829)) is a tropical species that is fast-growing and lives in symbiosis with zooxanthellae

(Harrison & Booth, 2007). This species, which secretes a mucus on its whole-body surface, can attach firmly to a substrate and then detach for locomotion (Clarke *et al*., 2020). This will leave a footprint on the surface, shown to be a meshwork formed by the contours of individual smooth cells and pores between the ectodermal basal disc cells (Rodrigues *et al*., 2016; Clarke *et al*., 2020). The ectoderm of the pedal disc contains four types of secretory cells: typical cnidarian mucocytes, two types of secretory cells containing distinct vesicles and a gland-like structure, and microvilli-bearing cells containing small electron-dense granules in the apical cytoplasm (Clarke *et al*., 2020). The adhesion of sea anemones is protein-based and was shown to be strongly associated with protein-protein interactions (Young *et al*., 1988; Davey *et al*., 2019). A recent study based on comparative transcriptomic and mRNA-sequencing analysis has identified multiple genes putatively involved in the adhesion of the sea anemone *E. pallida*. This study proposed that the adhesive secreted by the pedal disc is like an ECM-like proteinaceous matrix that contains collagen, glycoproteins, proteoglycans, lectins, and enzymes such as metalloproteinases and serine proteinases that can degrade and remodel this ECM to facilitate the detachment-attachment of *E. pallida* from surfaces (Fig. 12B). Their adhesion is also thought to be helped by mucocytes with β-glycan, sulfated and carboxylated polysaccharide complexes playing a role in adhesion (Clarke *et al*., 2020). Moreover, cross-linking by oxidoreductase reactions may also occur to strengthen adhesion (Davey *et al*., 2019), and four tyrosinases were identified in the pedal disk of the species *Haliplanella luciae* (new accepted name *Diadumene lineata* (Verrill, 1869)), another sea anemone species (Wang *et al*., 2020). Later, a multi-omics study of the pedal disc of *H. luciae* identified 1262 adhesive components, including tyrosinases, structural proteins related to the extracellular matrix (ECM), and hypothetical novel proteins (Wang *et al*., 2022). Among them, a cysteine-rich thrombospondin-1 type I repeat-like (TSRL) protein was also identified that can form coordination bonds with Ca^{2+} (Wang *et al*., 2022).

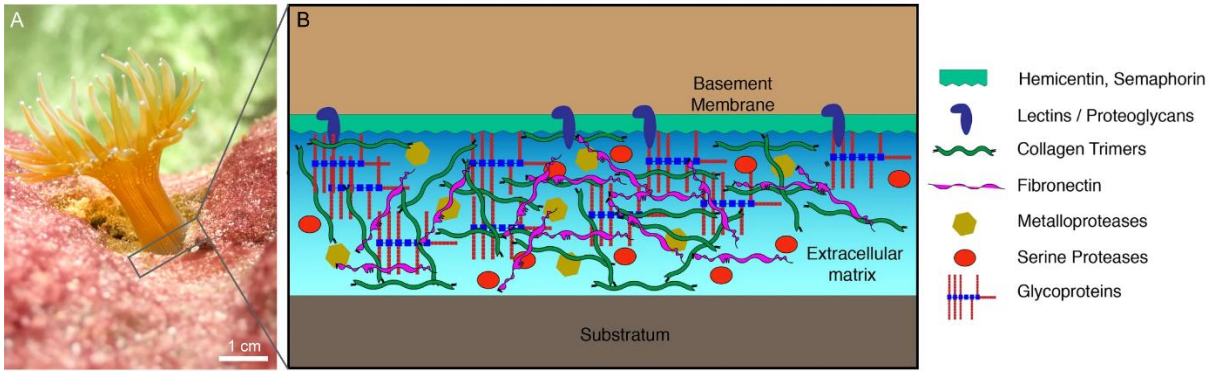


Figure 12. Picture of a sea anemone spp. (A) with an illustration of its ECM-like proteinaceous adhesive matrix, which includes collagen, glycoproteins, proteoglycans, lectins, and enzymes such as metalloproteinases and serine proteinases. These enzymes can degrade and remodel the ECM, enabling the sea anemone to detach and reattach to surfaces (B) (from Davey *et al*., 2019).

3. Diversity of enzymes involved in marine invertebrate adhesion

The first part of this introduction described most of the invertebrate marine organisms with adhesive systems documented in the literature. In this second part, I will focus more on the enzymes that have been identified as being involved in their adhesive systems. Thanks to multi-omics technology, many discoveries have been made regarding the enzymes involved in marine invertebrate adhesion. Most of these enzymes have been identified directly through mass spectrometry analysis of secreted adhesive materials (or from proteins extracted from adhesive materials and separated by SDS-PAGE), and then compared with the transcriptome of adhesive organs ("proteo-transcriptomic analyses"). However, some studies used more indirect approaches to tentatively identify enzymes involved in the adhesive systems of marine invertebrates. Differential proteomic analysis between different parts of the adhesive organ has been used to highlight these enzymes ("differential proteome analysis"). Some enzymes were also discovered by examining the differential expression of genes through transcriptomic analysis of adhesive organ tissue compared to other tissues in the organism ("differential transcriptomic analysis"). In addition, certain enzymes have been identified by the sequencing and assembly of a genome or a transcriptome from the adhesive organ (“genomic or transcriptomic analysis”). Some enzymes are more hypothetical ("hypothetically present”), for example, we can expect the presence of tyrosinase by finding a high concentration of DOPA, or the presence of glycosyl transferase may be suggested by identifying sugar moieties on the adhesive proteins.

Those enzymes are listed in Table 1, where each enzyme is described along with its associated EC number and reaction type. Regarding the EC number, the first Enzyme Commission, in its 1961 report, established a classification system for enzymes that also serves as the basis for assigning code numbers to them. During this commission, the International Union of Biochemistry and Molecular Biology (I.U.B.M.B.) categorized the known enzymes into six main classes: class 1 – Oxidoreductases, class 2 – Transferases, class 3 – Hydrolases, class 4 – Lyases, class 5 – Isomerases, and class 6 – Ligases. These code numbers, prefixed by EC, are now widely used and consist of four elements separated by periods, with the following meanings:

(i) The first number indicates which of the six main classes the enzyme belongs to.

(ii) The second number specifies the subclass.

(iii) The third number identifies the sub-subclass.

(iv) The fourth number is the serial number of the enzyme within its sub-subclass (https://www.enzyme-database.org/rules.php).

There are also several classification and nomenclature rules, but they are not cited here (Webb, 1992; McDonald *et al*., 2009).

Table 1. List of the enzymes potentially involved in the adhesive systems of the marine invertebrates described in the first part of this introduction. The EC number and reaction type were obtained from different enzyme databases, i.e., ExplorEnz – the Enzyme Database (https://www.enzyme-database.org/index.php), BRENDA (https://www.brenda-enzymes.org/index.php), and Expasy – Enzyme nomenclature database (https://enzyme.expasy.org/).

Enzyme common names	Enzyme class [EC number]	Reaction type	Organism	Adhesive protein target	Function in the adhesion	Identification method	References
Instantaneous adhesion							
Aldolase	Lyase [EC 4.1.2.13]	It catalyzes a reversible reaction that splits the aldol, fructose 1,6-bisphosphate, into the triose phosphates dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P).	Sea cucumber (*Holothuria dofleinii*)	NA	The fructose-bisphosphate aldolase was identified in the adhesive prints left on glass after the removal of Cuvierian tubules. This enzyme may mediate adhesion to the surface.	Proteomic analysis of SDS-gel band from the Cuvierian tubule print extract.	Peng *et al.*, 2014
Glucose-6-phosphate isomerase	Isomerase [EC 5.3.1.9]	This enzyme catalyzes the isomerization of glucose-6-phosphate (G6P) in fructose-6-phosphate (F6P).	Sea cucumber (*Holothuria dofleinii*)	NA	The glucose-6-phosphate isomerase is known to be part of the pentose phosphate pathway or glycolysis. It is thought to perform more than one function in adhesion and attachment.	Proteomic analysis of SDS-gel band from the Cuvierian tubule print extract.	Peng *et al.*, 2014
Kinase	Transferase [EC 2.7.11]	Catalyzes the phosphorylation of a serine residue to a phosphoserine.	Sea cucumber (*Bohadschia subrubra, Holothuria forskali, Pearsonothuria graeffei*)	NA	The adhesive system of sea cucumbers has been found to be rich in phosphoserine, but its role is still unknown.	Enzyme hypothetically present because anti-phosphoserine antibody labelling of the granular cells and western blotting show phosphorylation sites.	Flammang *et al.*, 2009
Transaldolase	Transferase [EC 2.2.1.2]	It catalyzes this reaction: sedoheptulose 7-phosphate + glyceraldehyde 3-phosphate ⇌ erythrose 4-phosphate + fructose 6-phosphate	Sea cucumber (*Holothuria dofleinii*)	NA	A transaldolase-like protein was identified in the H7 band of the gel and is also known to be an enzyme of the pentose	Proteomic analysis of SDS-gel band from the Cuvierian tubule print extract.	Peng *et al.*, 2014

					phosphate pathway. It may have the same role as above.		
Transketolase protein	Transferase [EC 2.2.1.1]	This enzyme catalyzes this reaction: D-glyceraldehyde 3-phosphate + D-sedoheptulose 7-phosphate = aldehydo-D-ribose 5-phosphate + D-xylulose 5-phosphate	Sea cucumber (*Holothuria dofleinii* & *Holothuria forskali*)	NA	This protein is present in band gel H3 in *H. dofleinii* and in the peritoneocytes of *H. forskali*.	Proteomic analysis of SDS-gel band from the Cuvierian tubule print extract.	Peng *et al*., 2014; Bonneel, 2020
Peptidase	Hydrolase [EC 3.4]	This enzyme acts on breaking peptide bonds.	Sea cucumber (*Holothuria forskali*)	NA	Putative peptidase M20 was found present in the peritoneocytes but has an unknown role.	Proteomic analysis of the tubule print material detached from glass surface.	Bonneel, 2020
Peptidyl-prolyl cis-trans isomerase	Isomerase [EC 5.2.1.8]	This enzyme catalyzes the isomerization between the cis and trans forms of peptide bonds.	Sea cucumber (*Holothuria forskali*)	NA	Peptidyl-prolyl cis-trans isomerase was found in the peritoneocytes but has an unknown role.	Proteomic analysis of the tubule print material detached from glass surface.	Bonneel, 2020
Protein disulfide-isomerase	Isomerase [EC 5.3.4.1]	This enzyme catalyzes the formation and breakage of disulfide bonds.	Sea cucumber (*Holothuria forskali*)	NA	The protein disulfide isomerase is present in the granular cells but also has an unknown role.	Proteomic analysis of the tubule print material detached from glass surface.	Bonneel, 2020
Tyrosinase	Oxidoreductases [EC 1.14.18.1]	This enzyme catalyzes the hydroxylation of tyrosine residue into DOPA, and the oxidation of DOPA to Dopaquinone.	Ctenophore (*Pleurobrachia bachei*)	Adhesive proteins	The formation of DOPA allows the adhesive proteins to make links with various substrates.	Genome sequencing and transcriptome profiling of the tentacles and other organs of several species of ctenophores.	Moroz *et al*., 2014; Townsend & Sweeney, 2019
Permanent adhesion							
Beta-glucosidase	Hydrolase [EC 3.2.1.21]	This enzyme hydrolyzes the non-reducing β-D-glucosyl residues, releasing β-D-glucose.	Barnacle (*Pollicipes pollicipes*)	NA	This enzyme was found to be highly expressed in the cement proteome, but its role is unknown.	Proteomic analysis of the cement.	Domínguez-Pérez *et al*., 2020

Glycosyl transferase	Transferase [EC 2.4]	N-glycosylation is the linkage of an oligosaccharide to the nitrogen atom of an asparagine residue.	Barnacle (*Megabalanus rosa*)	BCP52k	The N-glycosylation of Cp52k may help the crosslinking of barnacle cement by associating its sugar chain with the amyloid, thanks to the protein Cp20k.	The presence of this enzyme is hypothetical, *as in silico* analyses and specific staining have shown that Cp52k is N-glycosylated.	Kamino *et al*., 2012
Lysyl oxidase	Oxidoreductase [EC 1.4.3.13]	This enzyme converts the lysine residues into its aldehyde derivative allysine.	Barnacle (*Amphibalanus Amphitrite*, *Pollicipes pollicipes* & *Lepas anatifera*)	BCPs	The barnacle cement is crosslinked through the formation of allysine by lysyl oxidase at the adhesive interface.	Proteomic analysis of the solubilized adhesive, SDS-PAGE staining and fluorometric enzyme activity assay.	So *et al*., 2017; Domínguez-Pérez *et al*., 2020; Domínguez-Pérez *et al*., 2021
Peroxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes this type of oxidative reaction: $AH_2 + H_2O_2 \rightarrow A + 2\ H_2O$	Barnacle (*Amphibalanus amphitrite*)	BCPs	This enzyme is involved in the cross-linking of the cement through the oxidation of phenol groups during the development of the adhesive interface in the barnacle.	Proteomic analysis of the solubilized adhesive, SDS-PAGE staining and fluorometric enzyme activity assay.	So *et al*., 2016; So *et al*., 2017
Dual oxidase	Oxidoreductase [EC 1.6.3.1]	This enzyme requires heme and calcium and catalyzes this reaction: $H^+ + NADH + O_2 = H_2O_2 + NAD^+$.	Sandcastle worm (*Phragmatopoma caudata*)	NA	One gene of dual peroxidase was found in *P. caudata* and is thought to be involved in the cross-linking of adhesive proteins.	Differential transcriptomic analysis of genes expressed in the parathoracic region.	Buffet *et al*., 2018
Kinase	Transferase [EC 2.7.11]	Catalyzes the phosphorylation of a serine residue to a phosphoserine.	Sabellariidae species (*Phragmatopoma caudata* & *Sabellaria alveolata*)	Poly(S) cement proteins	A cluster of proteins containing a kinase domain was identified and overexpressed in *S. alveolata* and *P. californica*. This modification is responsible for the cross-linkage via ionic bonds between poly(S) proteins.	Differential transcriptomic analysis of genes expressed in the parathoracic region.	Buffet *et al*., 2018

Laccase	Oxidoreductase [EC 1.10.3.2]	This enzyme is a copper oxidase that operates on phenolic compounds forming reactive quinones or quinone methides that crosslink nucleophilic sidechains between adjacent cuticle proteins and chitin microfibrils.	Sandcastle worms (*Phragmatopoma californica*)	NA	This enzyme operates on phenolic compounds and possesses a broad substrate specificity. It may be involved in sclerotization and/or pigmentation.	Sequencing of random clones from cDNA libraries of the parathoracic region of the worm.	Endrizzi & Stewart, 2009
Peroxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes this type of oxidative reaction: $AH_2 + H_2O_2 \rightarrow A + 2\ H_2O$.	Sabellariidae species (*Sabellaria alveolata* & *Phragmatopoma caudata*)	GY-rich and H-repeat cement proteins	Peroxidase can oxidize catechol (L-DOPA) with hydrogen peroxide to form DOPAquinone and thus could be involved in the attachment to surfaces.	Differential transcriptomic analysis of genes expressed in the parathoracic region.	Buffet *et al*., 2018
Protease	Hydrolase [EC 3.4]	This enzyme acts on breaking peptide bonds.	Sabellariidae species (*Sabellaria alveolata* & *Phragmatopoma caudata*)	NA	Several BPTI/Kunitz domain-containing proteases were found to be overexpressed in the parathorax of both species.	Differential transcriptomic analysis of genes expressed in the parathoracic region.	Buffet *et al*., 2018
Tyrosinase	Oxidoreductase [EC 1.14. 18.1]	Catalyzes the hydroxylation of tyrosine residue to DOPA, and the oxidation of DOPA to Dopaquinone.	Sabellariidae species (*Phragmatopoma caudata*, *Phragmatopoma californica* & *Sabellaria alveolata*)	GY-rich and H-repeat cement proteins	This enzyme is important for making links and in the curing mechanisms of their adhesive system.	Transcriptomic analysis, and sequencing of random clones in a constructed cDNA libraries.	Endrizzi & Stewart, 2009; Buffet *et al*., 2018; Wang & Stewart, 2012
4-arginyl-hydroxylase or arginine-hydroxylase	Oxidoreductase [EC 1.14.11.73]	This enzyme catalyzes the hydroxylation of an arginine residue to produce 4-hydroxyarginine.	Mussels (*Mytilus californianus*, *Mytilus edulis* & *Mytilus galloprovinciali s*)	Mfp-3	Mfp-3 is an adhesive protein that forms a link between the surface and the plaque. The addition of hydroxyl groups increases hydrogen bonding with surfaces.	Enzyme hypothetically present as 4-hydroxyarginine was detected by collision-induced dissociation, amino acid analysis and mass spectrometry.	Papov *et al*., 1995; Zhao *et al*., 2006 ; Waite, 2017

Amine oxidase	Oxidoreductase [EC 1.4.3.21]	This enzyme catalyzes this reaction: a primary methyl amine + H_2O + O_2 = an aldehyde + H_2O_2 + $NH^4(+)$	Mussel (*Mytilus coruscus*)	NA	Amine oxidase was identified in *M. coruscus* byssus, but its role is unknown.	Proteomic (shotgun-LTQ) analysis on different part of the byssus.	Qin *et al*., 2016
Bifunctional arginine demethylase	Oxidoreductase [EC 1.2]	It removes the methyl group from proteins.	Mussel (*Perna viridis*)	NA	This enzyme is potentially involved in the metabolism and processing of byssal proteins.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Glycosyl-hydrolase	Hydrolase [EC 3.2.1.17]	It catalyzes the hydrolysis of glycosidic bonds in polysaccharides.	Mussel (*Mytilus coruscus*)	NA	Glycosyl-hydrolase was identified in the proximal thread of *M. coruscus* byssus, but its role is unknown.	Proteomic (shotgun-LTQ) analysis on different part of the byssus.	Qin *et al*., 2016
Glycosyl transferase	Transferase [EC 2.4]	O-glycosylation is the linkage of an oligosaccharide to the oxygen of a serine or threonine residue, while N-glycosylation is the linkage of an oligosaccharide to the nitrogen atom.	Mussel (*Mytilus edulis*)	PTMP-1	PTMP is glycosylated by α-D-mannopyranosyl or α-D-glucopyranosyl residues.	Glycosylation was demonstrated by Western blot on purified PTMP using biotinylated concanavalin A and sugar analysis.	Sun *et al*., 2002
Kinase	Transferase [EC 2.7.11.1]	This enzyme phosphorylates serine residues to produce phosphoserine.	Mussel (all the studied Mytilidae species)	Mfp-5 & -6	Several serine/threonine protein kinases were identified in the foot and fought to be involved in the phosphorylation process of serine residue.	Whole-genome and foot-specific transcriptomic analyses.	Waite & Qin, 2001; Zhao & Waite, 2006; Inoue *et al*., 2021
Lysozyme	Hydrolase [EC 3.2.1.17]	This enzyme catalyzes the hydrolysis of the bond between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in a peptidoglycan.	Mussel (*Perna viridis*)	NA	A gene encoding a lysozyme was found to be expressed in the foot and is thought to be involved in the antibacterial defense of the byssus.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Lysyl-hydroxylase	Oxidoreductase [EC 1.14.11.4]	This enzyme catalyzes the hydroxylation of lysine to hydroxylysine.	Mussel (*Perna viridis*)	NA	This enzyme was found among the proteins of the foot and is potentially involved in	Whole-genome and foot-specific	Inoue *et al*., 2021

					metabolism and processing of byssal proteins.	transcriptomic analyses.	
Glycosyl transferase	Transferase [EC 2.4]	This enzyme catalyzes the link of a sugar molecule to the oxygen atom of a threonine (Thr) residues in a protein.	Mussel (*Perna viridis*)	Pvfp-1	The O-glycosylation of a threonine residue by mannose, galactose, galactosamine, glucose, or fucose may increase conformational stability or mediate protein association to stabilize the protein assembly of the byssus.	Lectin-blot, proteomic and whole-genome and foot-specific transcriptomic analyses.	Ohkawa *et al*., 2004 ; Inoue *et al*., 2021
Peptidyl-prolyl cis-trans isomerase	Isomerase [EC 5.2.1.8]	This enzyme catalyzes the isomerization between the cis and trans forms of peptide bonds.	Mussel (*Perna viridis*)	PreCOLs	This enzyme helps for the construction of specific helical structures of collagens in byssal threads.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Peroxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes this type of oxidative reaction: $AH_2 + H_2O_2 \rightarrow A + 2\ H_2O$	Mussel (*Mytilus coruscus*)	NA	Peroxidase is a heme-containing enzyme that helps maintain the redox balance during the conversion between tyrosine and DOPA and/or between DOPA and dopaquinone in mussel byssus.	Proteomic (shotgun-LTQ) analysis on different part of the byssus.	Qin *et al*., 2016
Phenylalanine-4-hydroxylase	Oxidoreductase [EC 1.14.16.1]	This enzyme catalyzes the formation of a tyrosine residue by the hydroxylation of the aromatic side-chain of phenylalanine.	Mussel (*Perna viridis*)	NA	Enzyme found in the foot transcriptome of *P.viridis* and potentially involved in metabolism and processing of byssal proteins.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
prolyl-4-hydroxylases	Oxidoreductase [EC 1.14.11.2]	This enzyme catalyzes the formation of trans-4-hydroxyproline. It requires Fe^{2+} and ascorbate	Mussels (*Mytilus edulis*)	PreCols & Mfp-1	Hydroxyproline play a crucial role in formation of a stable collagen triple helix of preCols and can be detected in the cuticle as well.	Enzyme extraction and enzymatic activity test.	Marumo & Waite, 1987

Protein arginine N-methyltransferase	Transferase [EC 2.1.1.321]	This enzyme catalyzes the methyl transfer to the arginine residues of protein substrates.	Mussel (*Perna viridis*)	NA	Also annotated as an enzyme potentially involved in the metabolism and processing of byssal proteins.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Protein disulfide-isomerase	Isomerase [EC 5.3.4.1]	This enzyme catalyzes the formation and disruption of disulfide bonds between cysteine residues within proteins during their folding process.	Mussel (*Perna viridis*)	PreCOLs	This enzyme helps to promote collagen helix formation in byssal threads.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Protein phosphatase	Hydrolase [EC 3.1.3.16]	Those enzymes removed the serine- or threonine-bound phosphate group from a wide range of phosphoproteins.	Mussel (*Perna viridis*)	NA	This enzyme is possibly involved in the metabolism and processing of byssal protein.	Whole-genome and foot-specific transcriptomic analyses.	Inoue *et al*., 2021
Superoxide dismutase	Oxidoreductase [EC 1.15.1.1]	This enzyme catalyzes this reaction: 2 superoxide + 2 H^+ = O_2 + H_2O_2	Mussel (*Mytilus coruscus*)	NA	Two superoxide dismutases were found in the thread and the plaque, potentially preventing damage from oxygen-mediated free radicals by catalyzing the dismutation of superoxide into molecular oxygen and hydrogen peroxide.	Proteomic (shotgun-LTQ) analysis on different part of the byssus.	Qin *et al*., 2016
Tryptophan hydroxylase	Oxidoreductase [EC 1.14.16.4]	This enzyme employs a mononuclear iron (II) as cofactor and is responsible for addition of the -OH group to form the amino acid 7-hydroxytryptophan. It is activated by phosphorylation.	Mussel (*Perna viridis*)	Pvfp-1	7-Hydroxytryptophan may have the same role as DOPA and chelate metal ions. It can be further mannosylated to block cleavage near residues modified by exoproteases and endoproteases, making the proteins more resistant to degradation.	Enzyme hypothetically present as hydroxytryptophan was detected by tandem mass spectrometry of isolated tryptic decapeptides of Pvfp-1 and by UV-vis spectrophotometry.	Zhao *et al*., 2009

Tyrosinase	Oxidoreductase [EC 1.14. 18.1]	This enzyme catalyzes the hydroxylation of tyrosine residue to DOPA and to TOPA and oxidize the DOPA to Dopaquinone.	Mussel (all the studied Mytilidae species)	All mfps, PreCOLs and TMP	DOPA is an important modified amino acid because it can form links between proteins and has a cohesive role, as well as with the substrates and has an adhesive role.	Proteomic analyses and RT-qPCR in different regions of the foot.	Burzio & Waite, 2002 ; Guerette *et al*., 2013 ; Qin *et al*., 2016
Arylsulfatase	Hydrolase [EC 3.1.6.1]	This enzyme catalyzes this reaction: aryl sulfate + H_2O = phenol + sulfate.	Mollusc (*Chlamys farreri*)	NA	A specific sequence from the foot of the scallop was annotated as arylsulfatase, but its role is unknown.	Proteomic analyses of SDS-PAGE gel bands from proteins extracted from the byssus and byssal threads.	Miao *et al*., 2015
ATPase	Hydrolase [EC 3.6.1.3]	ATPase or Adénosinetriphosphatase catalyze the decomposition of ATP into ADP and a free phosphate ion.	Oyster (*Pinctada fucata*)	NA	ATPase, ATP synthase subunit beta, and calcium-transporting ATPase were identified in the distal byssus protein extraction.	Proteomic analyses of SDS-PAGE gel bands from extracted byssal proteins.	Liu *et al*., 2015
Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1	Transferase [EC 2.4.1.122]	This enzyme catalyzes this reaction: UDP-alpha-D-galactose + N-acetyl-alpha-D-galactosaminyl-R = UDP + beta-D-galactosyl-(1->3)-N-acetyl-alpha-D-galactosaminyl-R.	Mollusc (*Chlamys farreri*)	NA	This enzyme was identified as a specific gene in the scallop foot tissue, but its function in the scallop adhesion is unknown.	Proteomic analyses of SDS-PAGE gel bands from proteins extracted from the byssus and byssal threads.	Miao *et al*., 2015
Kinase	Transferase [EC 2.7.11]	This enzyme catalyzes the phosphorylation of a Thr/Ser residue to a phosphoserine.	Mollusc (*Chlamys farreri*)	Sbp-5-2, Sbp-7 and Sbp-8-2	Several different kinases were identified in genes highly expressed in the foot, suggesting that phosphorylation may be a crucial posttranslational modification (PTM) for scallop byssal proteins (Sbps).	Proteomic analyses of SDS-PAGE gel bands from proteins extracted from the byssus and byssal threads.	Miao *et al*., 2015

Peroxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes this type of oxidative reaction: $AH_2 + H_2O_2 \rightarrow A + 2\ H_2O$.	Oyster (*Pinctada fucata*, *Pinctada maxima*, *Chlamys farreri*)	NA	Peroxidase and peroxidasin were found in the byssus of different species of oysters, suggesting a requirement for these enzymes, as well as superoxide dismutase and proteinase inhibitors, to prevent degradation by an oxidative environment or by proteinases from microorganisms in the seawater. This enzyme can also catalyze the oxidation of DOPA to DOPAquinone and strengthen cross-links among the foot proteins during adhesive plaque formation.	Proteomic analysis of SDS-PAGE gel bands from proteins extracted from the byssus and byssal threads. And proteomic analysis of byssal threads.	Liu *et al*., 2015; Miao *et al*., 2015; Whaite *et al*., 2022
Superoxide dismutase	Oxidoreductase [EC 1.15.1.1]	This enzyme catalyzes this reaction: 2 superoxide + 2 H^+ = $O_2 + H_2O_2$.	Oyster (*Pinctada fucata*)	NA	A copper/zinc superoxide dismutase was identified in the byssus of *P. fucata* and may also help to prevent degradation by an oxidative environment.	Proteomic analyses of SDS-PAGE gel bands from extracted byssal proteins.	Liu *et al*., 2015
Tyrosinase	Oxidoreductase [EC 1.14. 18.1]	This enzyme catalyzes the hydroxylation of tyrosine residues to DOPA and oxidizes the DOPA to DOPAquinone.	Oyster (*Pinctada fucata*, *Pinctada maxima*, *Chlamys farreri Atrina pectinata*)	NA	Tyrosinases were identified in the adhesive plaques and in the foot of *P. fucata*, *P. maxima* and *C. farreri* and are thought to be responsible for the hydroxylation of the tyrosine residue of the adhesive proteins.	Proteomic analysis of the byssus and Arnow assay in *A. pectinata*.	Miao *et al*., 2015; Yoo *et al*., 2016; Liu & Zhang, 2021; Whaite *et al*., 2022
Tyrosinase	Oxidoreductase [EC 1.14.18.1]	It catalyzes the hydroxylation of tyrosine residue into DOPA, and the oxidation of DOPA to Dopaquinone.	Ascidians (*Ciona robusta*)	ASP-1	In adult *C. robusta*, it was suggested that the adhesion of ASP-1 is DOPA-dependent.	Enzyme hypothetically present because DOPA was found to be present in the	Li *et al*., 2019; Zeng *et al*., 2019

						modified protein by NBT staining.	
Temporary adhesion							
Glycosyl transferase	Transferase [EC 2.4]	O-glycosylation is the linkage of an oligosaccharide to the oxygen of a serine or threonine residue.	Flatworm (*Macrostomum lignano*, *Minona ileanae* & flatworm *Theama mediterranea*)	Mlig-ap2, Mile-ap2 & Tmed-ap2	It was suggested that the high threonine content of Mile-ap2a and Mlig-ap2 could indicate O-linked glycosylation. Moreover, lectin labeling confirmed the presence of sugar moieties.	Enzyme hypothetically present because an analysis of the footprint by lectin staining revealed the presence of sugar moieties.	Pjeta *et al*., 2019; Bertemes *et al*., 2022
Tyrosinase	Oxidoreductase [EC 1.14.18.1]	This enzyme catalyzes the hydroxylation of tyrosine residue into DOPA, and the oxidation of DOPA to Dopaquinone.	Flatworm (*Theama mediterranea*)	NA	This tyrosinase may crosslinks the adhesive materials during the attachment process.	Identified by generating a tail-specific positional RNA sequencing dataset.	Bertemes *et al*., 2022
Glycosyl transferase	Transferase [EC 2.4]	O-glycosylation is the linkage of an oligosaccharide to the oxygen of a serine or threonine residue, while N-glycosylation is the linkage of an oligosaccharide to the nitrogen atom.	Sea star (*Asterias rubens*)	Sfps	In their adhesives, some sfp are glycosylated and the carbohydrate fraction of adhesive proteins may enhance adhesion through electrostatic interactions between the polar and hydrogen bonding functional groups of glycan chains (such as ethers, amides, hydroxyls, and carboxylates) and underwater surfaces.	Enzyme hypothetically present because adhesive proteins were shown to contain sugar moieties through lectin analysis.	Hennebert *et al*., 2011
Ovoperoxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes the oxidation of tyrosine, leading to the formation of di-tyrosine residues.	Sea star (*Asterias rubens*)	NA	A footprint protein show similarity with ovoperoxidase and may help for the formation of cross-links between the adhesive proteins, thereby improving footprint cohesion.	Proteomic analysis of the adhesive footprint.	Hennebert *et al*., 2015c

Protease - Astacin	Hydrolase [EC 3.4.24.21]	This enzyme hydrolyses the peptide bonds.	Sea star (*Asterias rubens*)	NA	Two proteases were found in the footprint proteome (metalloproteinase SpAN and tolloid-like protein 2) and present a metalloendopeptidase activity. They were proposed as involved in the detachment process of the sea star.	Proteomic analysis of the adhesive footprint.	Hennebert *et al*., 2015c; Lengerer & Ladurner, 2018; Algrain *et al*., 2022
Glycosylase	Hydrolase [EC 3.2]	This enzyme acts on breaking glycosyl groups.	Sea urchin (*Paracentrotus lividus*)	NA	N-(beta-n-acetylglucosaminyl)-l-asparaginase, Carbohydrate-binding family 9-like, and Sialidases are over-expressed in the disc and may act on the detachment of the tube feet.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Glycosyl transferase	Transferase [EC 2.4]	O-glycosylation is the linkage of an oligosaccharide to the oxygen of a serine or threonine residue, while N-glycosylation is the linkage of an oligosaccharide to the nitrogen atom.	Sea Urchin (*Paracentrotus lividus*)	NA	Most of the adhesive proteins found in the sea urchin are glycosylated and contain (GlcNAc) N-Acétylglucosamine and N-Acetylgalactosamine (GalNAc).	Hypothetically present as glycosylation was shown present by lectin blotting analysis complemented with mass spectrometry analysis.	Ventura *et al*., 2023
Glycosyltransferase	Transferase [EC 2.4.1.221]	This enzyme adds O-fucose through an O-glycosidic linkage to serine or threonine residues.	Sea Urchin (*Paracentrotus lividus*)	NA	The glycosyltransferase O-fucosyltransferase 1 protein was found overexpressed in the tube foot disc.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Peptidyl-prolyl cis-trans isomerase	Isomerase [EC 5.2.1.8]	This enzyme catalyzes the isomerization between the cis and trans forms of peptide bonds.	Sea Urchin (*Paracentrotus lividus*)	NA	This enzyme is involved in the folding and stability of proteins during the biosynthesis of adhesive proteins.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016

Peroxidase	Oxidoreductase [EC 1.11.1.7]	This enzyme catalyzes this type of oxidative reaction: $AH_2 + H_2O_2 \rightarrow A + 2\ H_2O$.	Sea urchin (*Paracentrotus lividus*)	NA	Tryparedoxin, Ovoperoxidase, Peroxiredoxin 4, Myeloperoxidase and Dual oxidase 1 & 2 are heme-dependent peroxidases involved in the polymerization of echinoderm temporary adhesives. They are involved either in curing or act as antioxidants, metabolizing hydrogen peroxide into water molecules.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016; Ventura *et al*., 2023
Phosphoacetylglucosamine mutase	Isomerase [EC 5.4.2.3]	This enzyme is involved in the phosphorylation of amino sugars.	Sea Urchin (*Paracentrotus lividus*)	NA	This enzyme was found overexpressed in the tube foot disc of the sea urchin.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Protease	Hydrolase [EC 3.4]	This enzyme acts on breaking peptide bonds.	Sea urchin (*Paracentrotus lividus*)	NA	Aminopeptidases, Dipeptidases, Bleomycin hydrolase-like and Cathepsin z are over-expressed in the tube foot disc and can act on the detachment of the tube feet.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Protein disulfide-isomerase A5	Isomerase [EC 5.3.4.1]	This enzyme catalyzes the rearrangement of disulfide bonds.	Sea Urchin (*Paracentrotus lividus*)	NA	This enzyme is involved in the folding and stability of proteins during the biosynthesis of adhesive proteins.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Protein phosphatase	Hydrolase [EC 3.1.3.16]	Those enzymes remove the serine- or threonine-bound phosphate group from a wide range of phosphoproteins.	Sea Urchin (*Paracentrotus lividus*)	NA	Protein-serine/threonine phosphatase and protein phosphatase were found in the tube foot disc and are proposed to regulate the exocytosis of adhesive and de-adhesive granules through calcium, in combination with the modulation of exocytosis	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016

					by protein phosphorylation and dephosphorylation.		
Serine/threonine kinase and phosphorylase activity	Transferase [EC 2.7.11]	This enzyme phosphorylates proteins while converting ATP to ADP.	Sea Urchin (*Paracentrotus lividus*)	NA	The significant presence of proteins with serine/threonine kinase and phosphorylase activity has been proposed to regulate the exocytosis of adhesive and de-adhesive granules by calcium, combined with the modulation of exocytosis by protein phosphorylation and dephosphorylation.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Sulfatase	Hydrolase [EC 3.1.6]	This enzyme catalyzes the hydrolysis of sulfate ester bonds from different substrates such as glycosaminoglycans, and steroid sulfates.	Sea urchin (*Paracentrotus lividus*)	NA	Two sulfatases, arylsulfatase and N-acetylglucosamine-6-sulfatase, were highly overexpressed in the tube foot disc and may play a role in binding the sea urchin's secreted adhesive to the disc cuticle or in enhancing the cohesion of the adhesive film by binding its components.	Differential proteome analysis of the tube foot disc and stem.	Lebesgue *et al*., 2016
Transitory adhesion							
Astacin (Peptidase family M12A)	Hydrolase [EC 3.4.24.21]	Peptidase M12A is a zinc-dependent metallopeptidase.	Limpet (*Patella vulgata*)	NA	An astacin-like protein, named P-vulgata_14, contains numerous domains for recognizing, binding, and degrading sugars and peptides, and may acts as an antibacterial agent within the pedal mucus.	Proteomic analysis of the BPAM mucus.	Kang *et al*., 2020
Glycoside hydrolase-like protein	Hydrolase [EC 3.2.1]	This enzyme cleaves O- or S-glycosides bonds within carbohydrates.	Limpet (*Patella vulgata*)	NA	A Glycoside hydrolase-like protein (P-vulgata_9) was found in IPAM and is a	Proteomic analysis of the IPAM mucus.	Kang *et al*., 2020

					secreted protein with a glycoside hydrolase domain that may be involved in the active degradation of pedal mucus, facilitating the transition from stationary to locomotive states.		
Metallopeptidase reprolysin	Hydrolase [EC 3.4.24]	This endopeptidase enzyme has zinc as cofactor and cleaves peptides.	Limpet (*Patella vulgata*)	NA	This metallopeptidase (P-vulgata_8) is a secreted protein found in IPAM mucus and features the reprolysin domain, a type of metallopeptidase also known as adamalysin M12B peptidase. However, it is unclear whether P-vulgata_8 has any enzymatic function.	Proteomic analysis of the IPAM mucus.	Kang *et al.*, 2020
Glycosyl transferase	Transferase [EC 2.4]	O-glycosylation is the linkage of an oligosaccharide to the oxygen of a serine or threonine residue, while N-glycosylation is the linkage of an oligosaccharide to the nitrogen atom.	Sea anemone (*Exaiptasia pallida*)	Extracellularly secreted proteins	The glycosylation of the proteins has the potential to facilitate non-covalent cross-linking, increasing cohesion and adhesion strength in the sea anemone. Uromodulin was identified and shown to be an upregulated extracellularly secreted protein in the pedal disc. This protein undergoes heavy glycosylation and promotes protein-protein interactions, resulting in gelation.	mRNA-sequencing of the pedal disc compared to a reference genome.	Davey *et al.*, 2019
Metalloprotease	Hydrolase [EC 3.4]	This enzyme breaks peptide bonds with a metal in its active site.	Sea anemone (*Exaiptasia pallida*)	NA	Several matrix metalloproteinases were identified and suggested to play a role in tissue regeneration and the	mRNA-sequencing of the pedal disc compared to a reference genome.	Davey *et al.*, 2019

					degradation of ECM proteins in pedal disc adhesion.		
Peptidase	Hydrolase [EC 3.4]	This enzyme acts on breaking peptide bonds.	Sea anemone (*Exaiptasia pallida*)	NA	Several genes associated with peptidase activity, such as serine proteases, protease inhibitors, and trypsin-like serine proteases, were identified in the pedal disc adhesion and thought to be involved in the attachment/detachment system, either by disrupting cross-linking in proteins or by having a role in recognizing, degrading, and remodeling the ECM-proteinaceous matrix.	mRNA-sequencing of the pedal disc compared to a reference genome.	Davey *et al*., 2019
Tyrosinase	Oxidoreductase [EC 1.14. 18.1]	This enzyme catalyzes the hydroxylation of tyrosine residue to DOPA, and the oxidization of DOPA to Dopaquinone.	Sea anemone (*Haliplanella luciae*)	NA	Four tyrosinases were identified as up regulated in the pedal disc of *H. luciae* and may occur to strengthen its adhesion.	Comparative transcriptomic and proteomic analysis of the cement deposited on glass slides.	Wang *et al*., 2020

3.1. Oxidoreductase

The largest class, and first class of enzymes [EC 1], comprises oxidoreductases which catalyze biological oxidation-reduction reactions. This enzyme class has 26 subclasses (https://enzyme.expasy.org/enzyme-byclass.html) and can be found in different biomes of life. These enzymes catalyze the exchange of electrons between donor and acceptor molecules in reactions involving electron transfer, proton/hydrogen extraction, hydride transfer, oxygen insertion, etc (Younus, 2019). Within this group, there are four main classes: oxidases, dehydrogenases, peroxidases, and oxygenases. These enzymes have been mainly studied because developing practical biocatalytic applications for oxidoreductase enzymes has been a goal for many years (May, 1979; May, 1999; Cárdenas-Moreno *et al*., 2023).

In the oxidoreductase reaction, the substrate that is oxidized is considered the hydrogen donor. The systematic naming convention for this class of enzymes is based on the donor: acceptor oxidoreductase format:

$$\mathbf{A^- + B \rightarrow B^- + A}$$

These enzymes are generally referred to as dehydrogenases; but the term reductase can be used. The term oxidase is specifically used when O_2 is the acceptor (McDonald *et al*., 2009). Most oxidoreductases need a cofactor for their catalytic activity (May, 1999), and they have a wide range of natural substrates, as they are involved in a variety of physiological functions (Marcia *et al*., 2010; Cárdenas-Moreno *et al*., 2023).

The enzymes from marine adhesive systems belonging to this class are tyrosinases, peroxidases, dual oxidases, lysyl oxidases, laccases, lysyl- or prolyl-4-hydroxylases, phenylalanine-4-hydroxylases, tryptophan hydroxylases and superoxide dismutases. Most of these enzymes have been identified in all types of adhesion in invertebrate animals, but particularly those with a permanent adhesive system. In this type of adhesion, the secreted adhesive is mainly proteinaceous, forming strong bonds with surfaces, preventing the animal from being dislodged by waves or predators. The most well-known post-translational modification in permanent adhesion is the hydroxylation of tyrosine residues to form DOPA. DOPA is crucial for the effective adhesion of these animals, as it can make hydrogen bonds, covalent cross-links, electrostatic interactions, and coordination complexes with metal ions and metal oxides. Additionally, it contributes to the cohesive strength of the adhesive material (Lee *et al*., 2011; Petrone, 2013; Waite, 2017). The enzymes found in this class are involved in forming molecules that can create strong bonds with various types of substrates.

Introduction

One of the main enzymes studied in this thesis is tyrosinase [EC 1.14.18.1], an oxidoreductase belonging to the type-3 copper protein family. Tyrosinase is a bifunctional enzyme that operates through two distinct oxidation mechanisms: it catalyzes the hydroxylation of monophenols to o-diphenols (monophenolase activity) and the subsequent oxidation of o-diphenols to o-quinones (diphenolase activity) (Sánchez-Ferrer *et al*., 1995; Ullrich & Hofrichter, 2007). Tyrosinase has broad substrate specificity and can react with exposed tyrosyl side chains in polypeptides to form reactive quinones enabling protein-protein cross-linking (Thalmann & Lötzbeyer, 2002; Fairhead & Thöny-Meyer, 2012). The active centers of these proteins consist of two copper atoms, each coordinated by three histidine residues, to which dioxygen is bound in a peroxy configuration (Fig. 13A) (Ramsden & Riley, 2014). This enzyme is further characterized by four distinct oxidation states, deoxy-, oxy-, met-, and inactivated deact-tyrosinase, as determined by the oxidation state of the two copper atoms in the active site (Ramsden & Riley, 2014). Tyrosinase can also exist in a latent form, exhibiting a "lag period" before activation when used *in vitro* (Ramsden & Riley, 2014). All type-3 copper proteins contain a conserved pair of copper-binding sites, called Cu(A) and Cu(B). These sequences can be classified into subclasses according to the presence of other conserved domains or motifs in their sequence (Aguilera *et al*., 2013) (Fig. 13B).

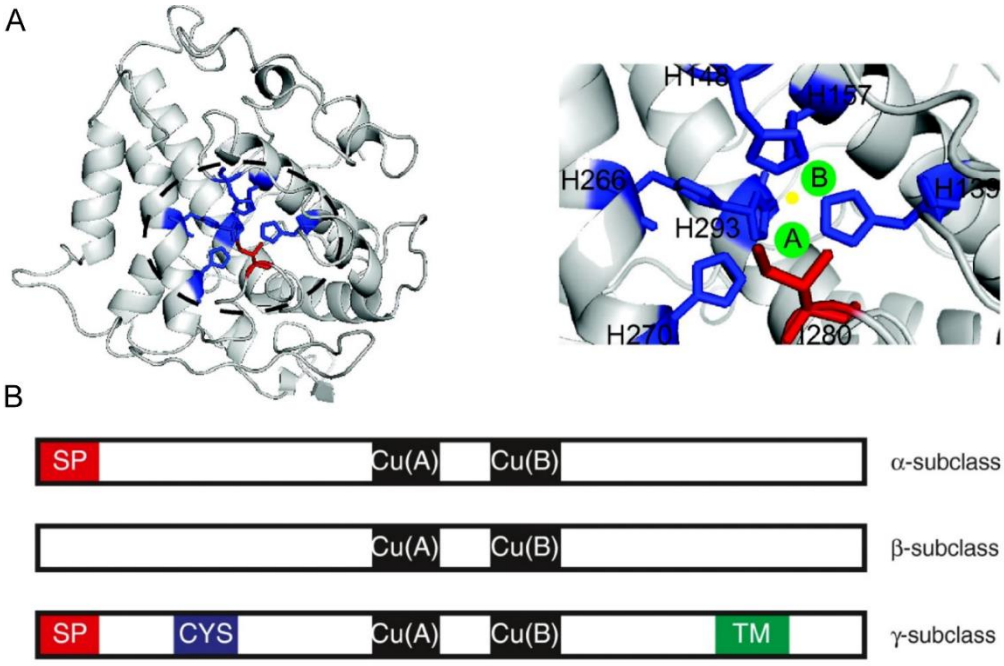


Figure 13. 3D view of a mollusc tyrosinase protein with a zoomed-in view of its binuclear active site, highlighting the placeholder residue: Copper ions are shown in green, with Cu(A) on the left and Cu(B) on the right. The yellow sphere represents a dioxygen molecule. The six histidine ligands coordinating the copper ions and stabilizing the structural conformation of the active site are depicted in dark blue. The residue entering the active site above Cu(A), blocking the substrate-binding pocket and preventing

substrate entry, is shown in red (A). An illustration of the domain architecture of different subclasses of type-3 copper proteins (B) (from Aguilera *et al*., 2013).

During their production, some tyrosinase enzymes undergo N-glycosylation. For example, the human tyrosinase protein contains several N-glycosylation sites and cysteine residues that participate in disulfide bond formation (Újvári *et al*., 2021). Additionally, certain tyrosinases can undergo maturation through a proteolytic mechanism (Van Gelder *et al*., 1997). For example, the tyrosinase from the mushroom *Agaricus bisporus* has been shown to be activated by a serine protease (Espín *et al*., 1999; Wichers *et al*., 2003).

Tyrosinases are widely distributed among animals, plants, and fungi, where they play a central role in melanin biosynthesis (Esposito *et al*., 2012). These tyrosinases differ in their sequences, size, glycosylation patterns, and mechanisms of activation (Jaenicke & Decker, 2003). However, it has been suggested that tyrosinase and tyrosinase-related protein gene family evolved from a common ancestral tyrosinase (Esposito *et al*., 2012). In animals, the typical substrate is L-tyrosine, and in vertebrates, tyrosinase serves as the principal enzyme involved in melanogenesis, a process occurring almost exclusively within specialized intracellular organelles known as melanosomes (Borovansky & Riley, 2011). In plants and many invertebrates, tyrosinase contributes to wound healing and the primary immune response (Van Gelder *et al*., 1997; Cerenius & Söderhäll, 2004). In arthropods, the enzyme also plays a critical role in the sclerotization of the cuticle following molting or injury (Claus & Decker, 2006).

3.2. Transferase

The second class of enzymes, EC 2 or transferases, are enzymes that transfer a functional group from a donor molecule, often a cofactor (coenzyme), to an acceptor molecule. This class of enzymes has 10 subclasses (https://enzyme.expasy.org/enzyme-byclass.html), which are notably based on the functional groups they transfer. The transferase reaction follows a donor: acceptor group transfer scheme, similar to this type of redox reaction (Palmer & Bonner, 2011):

AX+ B < = > BX + A

The most common transferase enzymes found in the adhesive systems of invertebrates are glycosyl transferases and kinases. Glycosylation has been shown to be involved in different types of adhesion: permanent (as seen in ascidians, barnacles, and mussels), temporary (as seen in flatworms, sea stars, and sea urchins), and transitory (as seen in limpets and sea anemones). Carbohydrates can be found linked to proteins, but they can also exist as free polysaccharides

(Hennebert *et al*., 2011). The current knowledge of the function of these sugars in adhesion is limited. However, it has been suggested that glycoconjugates contribute to adhesion and cohesion, possibly through electrostatic interactions and/or cross-linking (Hennebert *et al*., 2015c; Ventura *et al*., 2023). In permanent adhesion, it was shown in barnacles that glycoconjugate molecules participate in the cross-linking of barnacle cement (Kamino *et al*., 2012; Liang *et al*., 2019a). And in sea anemones, glycoproteins have been shown to be part of their ECM-proteinaceous matrix, where polysaccharide complexes establish links for adhesion (Davey *et al*., 2019).

The second most prevalent type of transferase enzyme is kinase, which is notably found in the adhesion systems of sandcastle worms and mussels. In the cement of Sabellariidae species, between 60 and 90% of the poly(S) cement proteins are serine residues that are phosphorylated by a kinase, leading to their conversion to phosphoserine. This phosphorylation makes the proteins very acidic and allows for a coacervation process with other components of the adhesive secretion (Buffet *et al*., 2018). A similar kind of phosphorylation occurs in mussel foot protein mfp-5, a protein that forms links with the substrate, but the amount of serine in mfp-5 is lower (Waite & Qin, 2001; Zhao & Waite, 2006). In temporary adhesion, the presence of proteins with serine/threonine kinase and phosphorylase activity has been proposed to regulate the exocytosis of adhesive and de-adhesive granules by calcium, combined with the modulation of exocytosis by protein phosphorylation and dephosphorylation (Lebesgue *et al*., 2016). Finally, it was shown that the adhesive system of sea cucumbers may also contain kinase, as indicated by the presence of phosphorylation sites (Flammang *et al*., 2009). The transketolase and transaldolase enzymes are also categorized as transferase enzymes and have been identified in the adhesion of the sea cucumbers (Peng *et al*., 2014). In summary, transferase enzymes are present in all types of adhesion and play an important role in the well-functioning of the adhesive mechanism.

3.3. Hydrolase

The third class of enzymes, EC3 or hydrolases, contains enzymes that catalyze the hydrolysis of chemical bonds in biomolecules:

$$\mathbf{R\text{-}R' + H_2O \rightleftharpoons R\text{-}OH + R'\text{-}H}$$

This is the most diverse class, and its members are categorized according to the type of bond they cleave. Their systematic names are formed as “substrate hydrolase,” but more commonly, their names take the form “substratease.” For example, nuclease refers to an

enzyme that hydrolyzes nucleic acids. These enzymes are involved in many different degradative reactions (Bairoch, 2000; Shukla *et al*., 2022).

Numerous hydrolases are found in the form of peptidases, also called proteases, which act on breaking peptide bonds, in all types of adhesive systems such as those of sea cucumbers (instantaneous), sandcastle worms (permanent), and sea urchins (temporary). They are also found as metalloproteases in limpets and sea anemones (transitory). In temporary adhesion, proteases such as an astacin-like protein with metalloendopeptidase activity have been identified and proposed to be involved in the detachment process of sea stars (Algrain *et al*., 2022) and sea urchins (Lebesgue *et al*., 2016). In transitory adhesion, hydrolases have also been suggested to be involved in the detachment of limpets, with glycoside hydrolases being involved in the active degradation of pedal mucus, facilitating the transition from stationary to locomotive states (Kang *et al*., 2020). In sea anemones, peptidase activity has been associated with the attachment/detachment system, either by disrupting cross-linking in proteins or by recognizing, degrading, and remodeling the ECM-proteinaceous matrix (Davey *et al*., 2019).

3.4. Lyase

The fourth class of enzyme, EC 4, are the lyase enzymes. These enzymes break C-C, C-O, C-N, and other bonds through mechanisms other than hydrolysis (EC 3) or oxidation (EC 1). This class has 10 subclasses (https://enzyme.expasy.org/enzyme-byclass.html), and unlike other enzymes, they involve two (or more) substrates in one reaction direction and result in one fewer compound in the reverse direction. When acting on a single substrate, they eliminate a molecule, creating either a new double bond or a new ring. Lyase enzymes are widely used in industry, and to biotransform specific pollutants of various natures (Demarche *et al*., 2012; Manubolu *et al*., 2018; Nunes *et al*., 2021).

Only one lyase enzyme has been found in the adhesive system of the sea cucumber: a fructose-bisphosphate aldolase, or aldolase. This enzyme was identified in a gel electrophoresis band from the adhesive prints left on glass after the removal of Cuvierian tubules and is suggested to mediate adhesion to the surface in a process similar to that found in bacteria (Peng *et al*., 2014). As only one enzyme of this class has been found, it is possible that lyases are rarely present in the adhesive systems of marine organisms and so this type of enzyme is of little use in adhesive mechanisms. Alternatively, as enzymes are not the primary focus of the adhesive research, more in-depth studies may be needed to identify them.

3.5. Isomerase

The isomerase enzymes, or EC 5 enzyme class, catalyze the isomerization of a molecule, the changes within a single molecule:

A → B

This process involves altering the bond connectivity or spatial arrangements within a single substrate molecule, resulting in a product with the same chemical formula, known as an isomer. This is a small class involving unimolecular reactions that are often fundamental in biology as isomerase participates in the central metabolism of organisms. There are seven different classes of isomerases. To characterize all isomerase reactions, we can examine the type of isomerism between substrate and product, highlighting three groups of similar reactions: enantioisomerism (racemases and epimerases, or EC 5.1), cis-trans isomerism (cis-trans isomerases, or EC 5.2), and structural isomerism (intramolecular oxidoreductases or EC 5.3, intramolecular transferases or EC 5.4, and intramolecular lyases or EC 5.5). Other isomerases include both stereoisomers and structural isomers (EC 5.6 and 5.99) (Martínez Cuesta *et al*., 2014; Martínez Cuesta *et al*., 2016).

Several isomerase enzymes have been identified in marine adhesive systems. For instance, in instantaneous adhesion, glucose-6-phosphate isomerase (EC 5.3) is believed to perform more than one function within the adhesive system of sea cucumbers (Peng *et al*., 2014). Additionally, peptidyl-prolyl cis-trans isomerase (EC 5.2) and protein disulfide-isomerase (EC 5.3) were detected in the peritoneocytes of the Cuvierian tubules of sea cucumbers, although their specific roles remain unknown (Bonneel, 2020). In permanent adhesion, peptidyl-prolyl cis-trans isomerase (EC 5.3) promotes collagen helix formation in the byssal threads of mussels (Inoue *et al*., 2021). In temporary adhesion, phosphoacetylglucosamine mutase (EC 5.4) is overexpressed in the disc of sea urchins, while protein disulfide-isomerase 3 (EC 5.3) catalyzes secreted proteins in the ER, playing a crucial role in protein folding and regulating protein glycosylation in sea urchins (Santos *et al*., 2013; Lebesgue *et al*., 2016). No isomerase was identified in the transitory adhesive system. Most of these enzymes belong to the third class of isomerases, which catalyze the oxidation of one part of a molecule with a corresponding reduction of another part of the same molecule. For example, disulfide-isomerases catalyze the rearrangement of disulfide bonds within proteins during folding, playing a significant role in metabolic reactions.

3.6. Ligase

The Ligase, or EC 6 enzyme class, catalyzes the joining of two molecules by forming new chemical bonds, coupled with the breakdown of the pyrophosphate bond in a nucleoside triphosphate. This class possesses seven subclasses, each defined by the type of bond formed, and the resulting bonds are high-energy bonds. Even though this class performs many biologically essential reactions, it is the smallest class of enzymes, with around 167 different reactions currently defined (Palmer & Bonner, 2011; Holliday *et al*., 2014). This type of enzyme was not found among the known enzymes involved in the adhesive system of marine invertebrates. This may be because this type of enzyme is mainly involved in metabolism, such as DNA ligases or aminoacyl-tRNA ligases. And it is often the secreting materials, notably their proteins content, that are investigated in adhesive materials studies.

Introduction

Research objectives

Research objectives

Many marine invertebrates from the intertidal zone have developed adhesive strategies to deal with the dynamic ocean environment by permanently or temporarily attaching themselves to surfaces. Their protein-based underwater adhesives are known for their superior strength and durability compared with man-made materials. The best-studied examples are the permanent adhesives of the mussels and the sabellariid tubeworms. Mussels produce multiple protein-rich threads, known collectively as the byssus, with each thread ending in a plaque that allows the adhesion to the substrate. On the other hand, the tubeworms construct their own tube by cementing particles together, then line its interior surface with an organic sheath. Both organisms produce adhesive proteins rich in post-translationally modified amino acids. Two of them, DOPA (3,4-dihydroxyphenylalanine) and phosphoserine, have been extensively studied and are known for their important interfacial adhesive as well as bulk cohesive roles. However, unlike adhesive proteins, the enzymes responsible for these modifications (tyrosinases and kinases, respectively) have been little studied, even though they are essential to catalyze the formation of functional groups in adhesive proteins. Several tyrosinases have been identified in the adhesive glands of mussels and tubeworms but their exact distribution and roles are unknown. As for the kinases involved in adhesive protein phosphorylation, they remain to be discovered.

In **Chapter I**, we investigate the diversity and expression patterns of tyrosinases in the blue mussel *Mytilus edulis* and in the honeycomb worm *Sabellaria alveolata* and address the long-lasting question about a potential convergent evolution of the DOPA-based adhesive mechanisms in both species. Using proteo-transcriptomic analyses, we identified and characterized a catalogue of tyrosinase candidates that may play a role in the adhesive protein maturation of *M. edulis* and *S. alveolata*. The role of some of these enzymes was further confirmed by localizing them in the different adhesive glands through *in situ* hybridization. Finally, phylogenetic analyses were conducted to explore whether the blue mussel and the honeycomb worm have developed homologous adhesive mechanisms to adapt to the hydrodynamic challenges of their habitats.

In the blue mussel, the gland distribution of some tyrosinases suggest they can be specific for a few adhesive proteins. In **Chapter II**, we focus on one of the previously identified tyrosinases of *M. edulis*. This candidate was localized exclusively in the plaque gland and was found within vesicles present in the gland, reinforcing its putative role in the maturation of plaque adhesive proteins. To evaluate its enzymatic activity, this enzyme was produced recombinantly with the goal of measuring DOPA production using UV-vis spectrometry.

Research objectives

Contrary to the mussel byssus, the knowledge about tubeworm adhesive system is more fragmentary. The European tubeworm species *S. alveolata* has been less investigated than the closely related Californian species, *Phragmatopoma californica*, and several differences between the two species have already been pointed out. **Chapter III** aims at increasing our understanding of the adhesive mechanisms in *S. alveolata* through proteomic analyses of granules extracted from its cement glands, as well as morphological and elemental composition studies. The enzymes responsible for the production of phosphoserine, a post-translational modification characteristic of tubeworm adhesive proteins, were also investigated using *in silico* analyses, and the transcripts encoding these enzymes were localized via *in situ* hybridization.

The tube lining is another aspect of the adhesive system of the honeycomb worm. Old histochemical studies indicate the presence of proteins, mucopolysaccharides and a tyrosinase in the glands producing the tube lining, but the exact nature of these constituents remains unknown. In **Chapter IV**, the organic sheath that covers the inside surface of the tube is investigated. This part of the thesis aims to characterize this complex structure, which has received less attention compared to tube cement, at the molecular level through proteo-transcriptomic analyses. The localization of the transcripts encoding the proteins identified in these analyses was also performed using *in situ* hybridization. Additionally, the ultrastructure of the tube lining and the glands that produce it were investigated.

Adhesives produced by marine organisms are a major source of inspiration for the development of biomimetic adhesives either for the medical or industrial fields. However, a deeper understanding of their fundamental mechanisms is essential to advance these applications. The work presented in this thesis aims at making a significant contribution towards our understanding of marine adhesive systems, in particular the enzymes involved in the permanent adhesion of mussels and tubeworms. This knowledge may lead to advances in the design of new biomimetic adhesive materials, particularly those based on DOPA

Chapter I

Diversity and evolution of the tyrosinase enzymes involved in the adhesive systems of mussels and tubeworms

Emilie Duthoo, Jérôme Delroisse, Barbara Maldonado, Fabien Sinot, Cyril Mascolo, Ruddy Wattiez, Pascal Jean Lopez, Cécile Van de Weerdt, Matthew J. Harrington and Patrick Flammang

Publication: Published in the Journal *iScience*, 111443. (https://doi.org/10.1016/j.isci.2024.111443)

Contribution: This study is based on preliminary data obtained by Barbara Maldonado a few years ago. Based on her results, I have redone all the *in silico* analyses. I carried out the proteomic analyses with the help of Cyril Mascolo and under the supervision of Ruddy Wattiez. I also performed all the *in situ* hybridization and phylogenetic analyses under the supervision of Jérôme Delroisse and Patrick Flammang. The results were interpreted by myself, Jérôme Delroisse, Matthew J. Harrington and Patrick Flammang.

Chapter I

Diversity and evolution of tyrosinase enzymes involved in the adhesive systems of mussels and tubeworms

Emilie Duthoo[1], Jérôme Delroisse[1,2], Barbara Maldonado[1,3], Fabien Sinot[1], Cyril Mascolo[4], Ruddy Wattiez[4], Pascal Jean Lopez[5], Cécile Van de Weerdt[3], Matthew J. Harrington[6] and Patrick Flammang[1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cellular and Molecular Immunology, GIGA Institute, University of Liège, 11 avenue de l'hôpital, 4000 Liège, Belgium

[3] Molecular Biomimetic and Protein Engineering Laboratory, GIGA, University of Liège, 11 avenue de l'hôpital, 4000 Liège, Belgium

[4] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

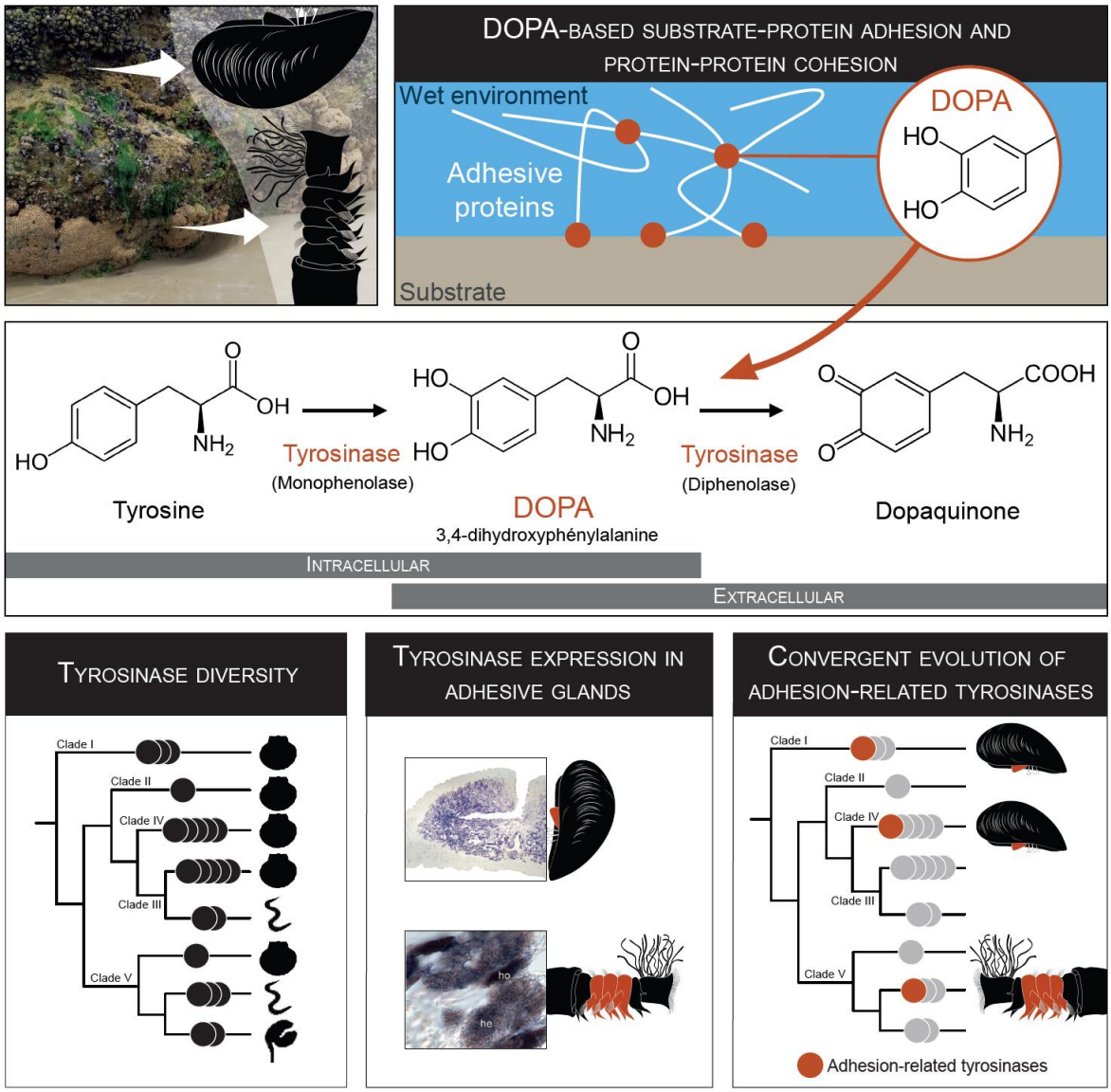
[5] UMR Biologie des Organismes et des Ecosystèmes Aquatiques, MNHN/CNRS-7208 Sorbonne Université/IRD-207/UCN /UA, 43 rue Cuvier, Paris 75005, France

[6] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

Summary

Mussels and tubeworms have evolved similar adhesive systems to cope with the hydrodynamics of intertidal environments. Both secrete adhesive proteins rich in DOPA, a post-translationally modified amino acid playing essential roles in their permanent adhesion. DOPA is produced by the hydroxylation of tyrosine residues by tyrosinase enzymes, which can also oxidize it further into dopaquinone. We have compiled a catalog of the tyrosinases potentially involved in the adhesive systems of *Mytilus edulis* and *Sabellaria alveolata*. Some were shown to be expressed in the adhesive glands, with a high gland specificity in mussels but not in tubeworms. The diversity of tyrosinases identified in the two species suggests the coexistence of different enzymatic activities and substrate specificities. However, the exact role of the different enzymes needs to be further investigated. Phylogenetic analyses support the hypothesis of independent expansions and parallel evolution of tyrosinases involved in DOPA-based adhesion in both lineages.

Graphical abstract



Highlights

- Tyrosinases are multifunctional enzymes playing numerous important roles in metazoans.
- We identified tyrosinases involved in adhesive protein maturation in mussels and tubeworms.
- Some tyrosinases are gland specific in mussels but not in tubeworms.
- The phylogeny of tyrosinases indicates a convergent evolution of the adhesive systems.

1. Introduction

Many marine organisms have evolved diverse attachment strategies to cope with their hydrodynamic environment (Delroisse *et al*., 2023). In particular, marine invertebrates rely on proteinaceous underwater adhesives for both permanent and temporary attachment to surfaces in the intertidal zone (Almeida *et al*., 2020; Li *et al*., 2021). These adhesives are known for their superior strength and durability compared with man-made materials and can therefore provide inspiration for the design and implementation of robust underwater adhesive strategies (Hofman *et al*., 2018). Two of the most extensively investigated organisms in this context are mussels and tubeworms (Fig. 1A-C). To attach themselves to rocks, mussels produce a byssus, which consists of a set of protein threads, each connected proximally to the base of the animal's foot (Fig. 1B), inside the shell, and ending distally in a flattened plaque sticking to the substratum (Waite, 2017). Byssal threads are formed by the auto-assembly of proteins secreted by three distinct glands enclosed in the mussel foot: the plaque gland, the core gland, and the cuticle gland (Priemel *et al*., 2017; Waite, 2017). To build and expand the tube in which they live, tubeworms of the family Sabellariidae collect mineral particles from their surroundings, dab them with spots of cement, and then add them to the opening of their tube (Vovelle, 1965; Stewart *et al*., 2004). The cement consists mostly of several different proteins produced by two types of parathoracic unicellular glands (cells with homogeneous granules and cells with heterogeneous granules) (Fig. 1C) (Becker *et al*., 2012; Wang & Stewart, 2012). The presence of DOPA (3,4-dihydroxy-L-phenylalanine) is a distinctive feature common to proteins identified in the adhesive systems of both mussels and tubeworms (Stewart *et al*., 2011a; Petrone, 2013; Davey *et al*., 2021). This post-translationally modified amino acid fulfils crucial roles in both interfacial adhesive and bulk cohesive interactions within the adhesive secretions (Sagert *et al*., 2006; Waite, 2017). DOPA is produced by the post-translational hydroxylation of tyrosine residues of the adhesive proteins by tyrosinase enzymes (Silverman & Roberto, 2007; Priemel *et al*., 2020).

Tyrosinases belong to the type-3 copper protein family alongside hemocyanins (Aguilera *et al*., 2013). These oxygen-transferring copper metalloproteins catalyze the *o*-hydroxylation of monophenols (e.g., tyrosine) into o-diphenols (e.g., DOPA) and the further oxidation of *o*-diphenols to *o*-quinones (Fig. 1E) (Ullrich & Hofrichter, 2007; Aguilera *et al*., 2013). Consequently, these enzymes exhibit both cresolase (monophenol monooxygenase, EC 1.14.18.1) and catecholase (catechol oxidase, EC 1.10.3.1) activities (Ramsden & Riley, 2014). Tyrosinases are distributed throughout the tree of life (Aguilera *et al*., 2013). They are among

the most widespread enzymes in nature as they are key enzymes for pigment synthesis in many organisms (Del Marmol & Beermann, 1996). They are also involved in many other biological processes such as shell formation in molluscs (Huan *et al*., 2013) or wound healing (Sugumaran *et al*., 1996; Theopold *et al*., 2004), parasite encapsulation (González-Santoyo & Córdoba-Aguilar, 2012) and cuticle sclerification in insects (Dennell, 1958; Andersen, 2010). Regardless of their biological function, all tyrosinases share a common origin traceable to an ancestral tyrosinase gene (Polivares & Solano, 2009). Lineage-specific gene duplication events took place during tyrosinase evolution, potentially contributing to functional diversification in specific taxa (Esposito *et al*., 2012; Aguilera *et al*., 2013).

Compared to the detailed knowledge available on mussel and tubeworm adhesive proteins, little information exists about the tyrosinases present in the adhesive systems of these organisms. In the mussel *Mytilus edulis*, Waite (1985), and later Hellio and collaborators (2000), reported the partial purification and kinetics of tyrosinases extracted from the foot and the byssus and which displayed a catechol oxidase activity (Waite, 1985; Hellio *et al*., 2000). More recently, developments in transcriptomics and proteomics allowed the discovery of new foot- or byssus-specific tyrosinase sequences. Guerette *et al*. (2013) identified 5 isoforms from the foot of the green mussel *Perna viridis* while Qin *et al.* (2016) detected 6 isoforms in the byssus of *Mytilus coruscus* (Guerette *et al*., 2013; Qin *et al*., 2016). Their distribution within the foot or the byssus suggests they might have specific substrates and/or functions. In tubeworms, one tyrosinase was identified in *Phragmatopoma californica* and shown to be expressed in the two types of adhesive glands (Wang & Stewart, 2012). A later study from the same authors demonstrated that the enzyme was a catechol oxidase that catalyzes the covalent cross-linking of L-DOPA (Wang & Stewart, 2013). In two other species, *Phragmatopoma caudata* and *Sabellaria alveolata*, a differential transcriptomic study highlighted 23 tyrosinase transcripts overexpressed in the region of the body enclosing the adhesive glands (Buffet *et al*., 2018). Many studies therefore point to an expansion of the tyrosinase repertoire linked to the maturation of adhesive proteins in both mussels and tubeworms.

The adhesive secretions of mussels and tubeworms provide a molecular model for studying the evolution of tyrosinases and the diversification of their functions. Despite belonging to two distinct phyla, mussels and tubeworms are phylogenetically close, both being lophotrochozoans (Fig. 1D). Moreover, co-occurring in some intertidal habitats, they are subjected to the same environmental conditions and selective pressures, and their adhesive mechanisms are notably similar, primarily relying on DOPA. Several reviews have hinted at

this similarity (Endrizzi & Stewart, 2009; Flammang *et al*., 2009; Stewart *et al*., 2011a; Hofman *et al*., 2018). However, due to the short and intrinsically disordered nature of the adhesive protein sequences in both mussels and tubeworms, no homology can be traced between them, and they are thought to have evolved independently (Kamino, 2010). Examining the enzymes involved in the maturation process of adhesive proteins, especially tyrosinases, in these two taxa could provide insights into the evolutionary relationships between their adhesive systems. In this study, we investigated the catalogue of tyrosinases potentially involved in the maturation of adhesive proteins in the blue mussel *M. edulis* and the honeycomb worm *S. alveolata* by performing proteo-transcriptomic analyses. *In situ* hybridization was then used to confirm the expression of these candidate enzymes in adhesive glands, validating their role in bioadhesion. Finally, we conducted phylogenetic analyses to address the question of the evolution and diversification of tyrosinases in these two lineages.

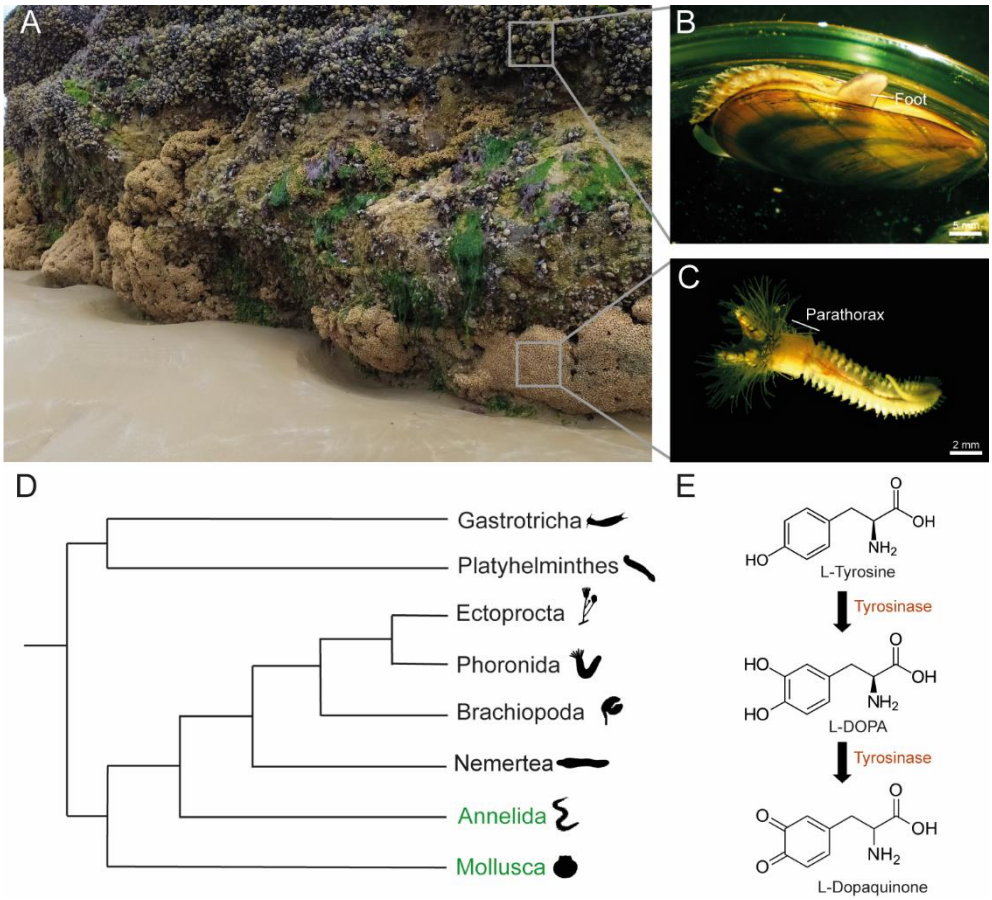


Figure 1. **Mussels and tubeworms share a similar DOPA-based adhesion** (**A**) Intertidal zone with a honeycomb worm reef covered by blue mussels (Douarnenez, France) (Picture courtesy of Alexia Lourtie). (**B, C**) *Mytilus edulis* and *Sabellaria alveolata*, respectively. (**D**) Phylogenetic position of mussels (Mollusca) and tubeworms (Annelida) (in green) within the Spiralia (based on Laumer *et al*., 2015). (E) The possible dual function of tyrosinases in the maturation of adhesive proteins comprising the modification of tyrosine residues into DOPA and the further oxidation of DOPA into dopaquinone (Polivares & Solano, 2009).

2. Methods

2.1. Experimental model and study participant details

Our experimental models included adult individuals of the blue mussel *Mytilus edulis* Linnaeus, 1758 (Mollusca, Bivalvia) and of the honeycomb worm *Sabellaria alveolata* Linnaeus, 1767 (Annelida, Polychaeta). The sex of the individuals was not determined. Neither of these species is threatened and no permits are required for their collection or maintenance. The authors have complied with all ethical standards required for conducting this research and the animals used in the experiments were maintained and treated in compliance with the guidelines specified by the Belgian Ministry of Trade and Agriculture.

2.2. Animal collection and maintenance

Individuals of *Mytilus edulis*, were collected intertidally in Audresselles (Pas-de-Calais, France). Reef fragments of *Sabellaria alveolata,* were collected at low tide on the “plage de Ris" at Douarnenez, France (48°05’34.6’’N 4°17’55.0’’W), or were obtained from the Station Biologique de Roscoff (Finistère, France). All animals were then transported to the Laboratory of Biology of Marine Organisms and Biomimetics (University of Mons, Belgium), where they were kept in a re-circulating aquarium (13°C, 33 psu salinity).

2.3. Transcriptome sequencing and *de novo* assembly

In the present study, two tissue transcriptomes have been generated: one from the foot of the blue mussel *M. edulis* and one from the anterior part of the honeycomb worm *S. alveolata*.

RNA extraction, library construction and sequencing were performed at the GIGA Genomics platform (Liège, Belgium). After dissection, mussel feet and tubeworm anterior parts (corresponding to the head and parathoracic region) were immediately frozen with liquid nitrogen and stored at -80°C until use. Total RNA was extracted from 100 mg of frozen tissue using Trizol (Life Technologies, Carlsbad, CA) and its quality was assessed using the Bioanalyzer 2100 (Agilent). Truseq Stranded mRNA Sample Preparation kit (San Diego, CA) was used to prepare a library from 500 ng of total RNA. Poly-adenylated RNAs were purified with oligo (dT)-coated magnetic beads (Sera-Mag Magnetic Oligo(dT) beads, Illumina) and then chemically fragmented to a length of 100 to 400 nucleotides -with a majority of the fragments at about 200 bp (base pairs)- by using divalent cations at 94 °C for 5 min. These short fragments were used as a template for reverse-transcription using random hexamers to synthesize cDNA, followed by end repair and adaptor ligation according to the manufacturer's protocol (Illumina, San Diego, CA). Finally, the ligated library fragments were purified and

enriched by solid-phase PCR following Illumina's protocol. The library quality was validated on the Bioanalyzer 2100. The high-throughput sequencing was conducted by a HiSeq 2000 platform (Illumina, San Diego, CA) to obtain 2x100-bp paired-end reads according to manufacturer's instructions. Real-time quality control was performed to ensure that most read quality score was higher than 30. The raw sequencing data have been deposited in the NCBI Sequence Read Archive with accession numbers NCBI SRA: SRR29446349 and NCBI SRA: SRR29446350 for *M. edulis* and *S. alveolata*, respectively.

Transcriptome quality was checked using Fast QC software (Babraham Bioinformatics). The Trinity software suite (Grabherr *et al*., 2011) which comprises a quality filtering function was used with default parameters to *de novo* assemble the raw reads with overlapping nucleic acid sequence into contigs. Transcriptome completeness was evaluated using BUSCO (v3.0.2) analyses on assembled transcripts (Waterhouse *et al*., 2018). Scores were calculated using Metazoan_odb10 lineage data.

2.4. *In silico* tyrosinase sequences identification

To find the tyrosinase sequences putatively involved in adhesive protein maturation, tBLASTn searches were performed on the two transcriptomes using a reference dataset (see Supplementary Table 1) consisting of byssus-specific tyrosinase sequences found in *M. coruscus* (Qin *et al*., 2016) and *P. viridis* (Guerette *et al*., 2013) and of a cement tyrosinase sequence from the tubeworm *P. californica* (Wang & Stewart, 2012). Several transcripts encoding tyrosinase-like proteins were obtained for both species. They were subsequently used in a reciprocal tBLASTn search against the NCBI non-redundant protein database and only those having a tyrosinase as the best-hit match were kept for further analyses.

The selected transcripts were translated *in silico* (Expasy, Translate tool) and analyzed by looking for the tyrosinase copper-binding domain (IPR002227) using InterPro (https://www.ebi.ac.uk/interpro/search/sequence/) (Paysan-Lafosse *et al*., 2023) and for the presence of a signal peptide using SignalP-6.0 (https://services.healthtech.dtu.dk/service.php?SignalP) (Teufel *et al*., 2022).

2.5. Protein extraction and mass spectrometry analyses

To assess whether tyrosinase enzymes are secreted in the adhesive materials of mussels and tubeworms, proteomic analyses were conducted on induced byssal threads and reconstructed tubes.

For *M. edulis*, the secretion of fresh byssal threads was induced by injecting a 0.56 M solution of KCl at the base of the foot as described by Tamarin *et al*., 1976. For protein extraction, approximately 30 byssal threads were crushed in 500 µL of a solution containing 5% acetic acid and 8 M urea (Rzepecki *et al*., 1992). The extract was then centrifuged for 20 min at 16,000 x *g* and the supernatant containing the proteins was collected. For *S. alveolata*, single tubes with their dwelling worm were isolated from the reef fragment. The upper third of each tube was then removed and the worms were allowed to reconstruct this missing portion of their tube with glass beads (425-600 µm in diameter; Sigma) (Jensen & Morse, 1988). Approximately 500 mg of freshly rebuilt tube fragments were subjected to protein extraction using a 1.5 M Tris–HCl buffer (pH 8.5) containing 7 M guanidine hydrochloride (GuHCl), 20 mM ethylenediaminetetraacetate (EDTA), and 0.5 M dithiothreitol (DTT) (Tris-GuHCl buffer). This mixture was incubated for 1 hour at 60 °C under agitation and then centrifuged as described above.

To further process proteins, both samples were treated with a 2.5-fold excess (w/w) of iodoacetamide to DTT, for 20 min in the dark at room temperature, to carbamidomethylate the sulfhydryl groups. The reaction was then stopped by adding an equal quantity of β-mercaptoethanol to iodoacetamide. The extract was centrifuged at 13,000 rpm for 15 min at 4 °C and the supernatant was collected. The protein concentration was determined using the Non-Interfering Protein Assay Kit (Calbiochem, Darmstadt, Germany) with bovine serum albumin as a protein standard. For each sample, 50 µg of proteins were then precipitated overnight at -20 °C in 80% acetone. After a 15-min centrifugation at 13,000 rpm and evaporation of acetone, the resulting pellet was subjected to overnight enzymatic digestion using modified porcine trypsin at an enzyme/substrate ratio of 1/50, at 37 °C in 25 mM NH_4HCO_3. The reaction was stopped by adding formic acid to a final concentration of 0.1% (v/v) (Hennebert *et al*., 2015c). Tryptic peptides were analyzed by LC connected to a hybrid quadrupole time-of-flight TripleTOF 6600 mass spectrometer (AB SCIEX, Concord, ON). Byssal threads and reconstructed tubes MS/MS data were searched for protein candidates against a database composed of the six open reading frames (ORFs) of the transcriptome of the foot of *M. edulis* or of the transcriptome of the anterior part of *S. alveolata*, respectively, using the Protein Pilot software (version 5.0.1). The samples with a false discovery rate (FDR) above 1.0% were excluded from subsequent analyses.

A proteomic analysis was also carried out on the mussel foot tissues. Feet were dissected and cut transversely into two halves to separate the proximal part from the distal part.

The samples were homogenized in a Potter-Elvehjem tissue grinder with a 4% solution of SDS in Tris-HCl buffer (pH 7.4), with protease inhibitors (EDTA-free), and 5 units of DNase per 100 mg of foot tissue. The samples were then sonicated three times for 30 s each and left at room temperature for 30 min, followed by overnight incubation at 4 °C. The protein content was quantified using an RCDC kit (Biorad). Aliquots of 15 µg of proteins were reduced and alkylated. The samples were then treated with the 2-D Clean-Up Kit (GE Healthcare) to eliminate impurities not compatible with mass spectrometry analyses and recovered in 50 mM ammonium bicarbonate buffer. Trypsin digestion was carried out for 16 hours at an enzyme/substrate ratio of 1/50 at 37 °C, the reaction was stopped by adding trifluoroacetic acid, and the samples were dried using a speed vac. The tryptic peptides were dissolved in water with 0.1% trifluoroacetic acid and purified using a Zip-Tip C18 High Capacity. They were analyzed by reverse-phase HPLC–ESI-MS/MS using a nano-UPLC (nanoAcquity, Waters) connected to an ESI-Q-Orbitrap mass spectrometer (Q Extractive Thermo) in positive ion mode. MS/MS data were analyzed against a database composed of the six open reading frames (ORFs) of the mussel foot transcriptome using the MaxQuant software (version 1.4.1.2). The peptide mass tolerance was set to ±10 ppm, and fragment mass tolerance was set to ±0.1 Da. The oxidation of the amino acids tyrosine, arginine and proline, as well as the carbamidomethylation of cysteine and the phosphorylation of serine were defined as fixed modifications, and the oxidation of methionine as variable modifications. Protein identifications were considered significant if proteins are identified with at least two peptides per protein taking into account only an FDR<0.01.

2.6. Localization of tyrosinase mRNA

The best tyrosinase sequence candidates found in the *in silico* and proteomic analyses were localized using in *situ* hybridization (ISH) technique to see if they are well expressed in the adhesive glands of both studied species. RNA was extracted from three parathoracic parts of honeycomb worms and three blue mussel feet using TRIzol™ Reagent kit (Thermofisher). The cDNA synthesis from the RNA extracted was done using Reverse transcription kit, Roche. Later, double-stranded DNA templates were amplified by PCR using the Q5 High-Fidelity DNA Polymerase kit method, with primer designed by Open Primer 3 (bioinfo.ut.ee/primer3/) with an optimal probe length between 700 and 900 bp. A second PCR was done with T7 promoter binding site (5'-GGATCCTAATACGACTCACTATAGG-3') added to reverse strand PCR primers. PCR products were purified using the Wizard SV Gel and PCR clean-up system kit (Promega) and used for RNA probe synthesis. Digoxigenin (DIG)-labelled RNA

probes were then synthesized with the kit DIG RNA Labelling Kit (Roche) with T7 RNA polymerase and DIG–dUTP. *In situ* hybridization was performed according to Lengerer *et al.*, 2019. The probes were used on parathoracic sections showing the adhesive glands for the honeycomb worm, and transversal sections of the blue mussel foot, at a concentration of 0.2 ng/µl and detected with antidigoxigenin-AP Fab fragments (Roche) at a dilution of 1:2000. The signal was developed using the NBT/BCIP system (Roche) at a dilution of 1:50 at 37 °C. Sections were observed using a Zeiss Axioscope A1 microscope connected to a Zeiss AxioCam 305 color camera.

To confirm the localization of the candidates in the blue mussel, we conducted a complementary whole mount *in situ* hybridization. For each candidate, the mussel feet were divided into two parts along the frontal axis. The protocol was done in 6-well plates according to Pfister *et al*., 2007, except for these steps; the *in situ* hybridization was carried out on samples re-incubated at 55 °C for 72 hours, and color development was performed in the dark at 37 °C using an NBT/BCIP system (Roth) until a satisfactory precipitate coloration was achieved.

2.7. Phylogenetic analyses and clustering analyses of tyrosinase sequences

To gain a deeper understanding of the relationship between all the tyrosinase enzymes, we conducted a CLANS analysis on a comprehensive dataset. This dataset included lophotrochozoan sequences in FASTA format, each containing a pfam00264 domain. These sequences were retrieved from the NCBI database (retrieved on September 20, 2022). In addition, we also incorporated the previously identified sequences obtained through *in silico* and proteomic analyses. We also incorporated three tyrosinase transcripts from the transcriptome of *Mytilus edulis*, which exhibited higher expression levels in mantle tissues (Matthew J. Harrington personal communication). To expand the diversity of tubeworm tyrosinases within the phylum Annelida, we obtained transcriptomes from *Phragmatopoma caudata* (Buffet *et al*., 2018). We conducted a local tBLASTn search targeting tyrosinase mRNA sequences that may play a role in protein cement maturation, using the previously mentioned *P. californica* sequence. As the number of sequences was relatively limited for Annelida compared with other phyla, we conducted a complementary analysis. In this analysis, we searched for the pfam00264 domain among the sequences found in the transcriptomes of *Capitella teleta* (3 transcripts), *Owenia fusiformis* (13 transcripts), *Oasisia alvinae* (4 transcripts), and *Riftia pachyptila* (3 transcripts), all of which are available on https://github.com/ChemaMD. The CLANS analysis was based on all-against-all sequence

similarity using BLAST searches with the BLOSUM62 matrix (https://toolkit.tuebingen.mpg.de/clans/) (Frickey & Lupas, 2004).

We performed a phylogenetic analysis using the tyrosinase sequences from two closely clustered groups identified in the CLANS analysis, and which contained the sequences from the two studied species. First, we conducted a multiple alignment with all these sequences, totaling 312 sequences, using the MAFFT algorithm (using the automated parameters of Mafft v7.490 implemented on Geneious Prime 2023.2.1 (Katoh *et al*., 2019)). Subsequently, we trimmed this alignment using the online TrimAL tool implemented on Phylemon2 (Capella-Gutierrez *et al*., 2009) with the Automated1 option parameter. We selected two tyrosinase sequences from Mollusca, which contained a hemocyanin domain as outgroup. The construction of the phylogenetic tree was carried out using the IQ-TREE (1.6.12) software, using the maximum likelihood method and performing ultrafast bootstrap (UFBoot) analysis with 1000 replicates (Hoang *et al*., 2018). According to the software's BIC scores, the best-fit model was determined to be the WAG+F+I+G4 model (Trifinopoulos *et al*., 2016). The tree was modified using the software iTOL (version 6.9) (Letunic & Bork, 2021).

3. Results

3.1. Transcriptomic analyses

The initial phase of this study aimed at identifying enzymes containing a tyrosinase domain that might play a role in adhesive protein maturation in both *M. edulis* and *S. alveolata*. To achieve this, a mussel foot transcriptome and a honeycomb worm anterior part transcriptome were obtained. Both transcriptomes were generated from 100 bp reads sequenced using the Illumina platform. The raw data comprises 6.1 Gbp for *S. alveolata* and 3.9 Gbp for *M. edulis*. The *M. edulis* dataset contains 173,165 predicted genes, encompassing a total of 152,956,891 nucleotides. The contig N50 is 1,652 nucleotides, with a maximum contig length of 19,275 nucleotides. The *S. alveolata* dataset includes 414,599 predicted genes, amounting to 332,743,778 nucleotides in total. The contig N50 is 1,409 nucleotides, and the maximum contig length is 31,799 nucleotides. The sequence length distribution of the predicted transcripts is shown in the Supplementary Figure 1A-B. The completeness of both transcriptome datasets was evaluated using BUSCO (Benchmarking Universal Single-copy Orthologue) analyses on assembled transcripts. Scores were evaluated using the predefined lineage data "Metazoan_odb10". For the *M. edulis*, analyses showed that 92.4% complete BUSCO groups were detected in the gene set. In detail, out of the 954 evaluated BUSCOs from the Metazoan

dataset, only 5.2% were fragmented and 2.4% were missing. For *S. alveolata*, the BUSCO analyses indicated that 97.8% complete BUSCO groups were detected with only 2.2% and 0% of fragmented and missing BUSCO groups (Supplementary Fig. 1C). Transcriptome data were then used to search for tyrosinase mRNA sequences using a similarity-based approach. A dataset of tyrosinases comprising mussel byssus (*Mytilus coruscus* (Qin *et al*., 2016), *Perna viridis* (Guerette *et al*., 2013)) and tubeworm cement (*Phragmatopoma californica* (Wang & Stewart, 2012)) sequences (see Supplementary Table 1) was used for local tBLASTn searches in the two transcriptomes to highlight sequences with similarity to these tyrosinases. Candidate matches were then used as queries in a reciprocal BLASTn search against online databases, and only those corresponding to a tyrosinase-like protein as the reciprocal hit were kept as putative candidates. Additionally, after *in silico* translation, short sequences and sequences lacking a tyrosinase domain (cl02830 or pfam00264) were removed.

In the blue mussel, the BLAST searches allowed retrieval of 85 transcripts coding for proteins with tyrosinases as the best reciprocal hit. This list was reduced to 17 candidates when only the longest sequences comprising a tyrosinase domain were considered. Most of the 17 sequences are full-length, except for Medu-TYR2 (comp79852) and Medu-TYR3 (comp74994) which lack the C-terminal region. Eight of these tyrosinases are the closest homologues (BLAST best hits) of the reference tyrosinases from *M. coruscus* and *P. viridis* (Table 1). All these proteins possess a signal peptide indicating they are potentially secreted at the level of the foot, some of them presumably by the byssus forming glands. Yet, the expression level of the tyrosinase encoding transcripts, expressed as Fragments Per Kilobase of transcript per Million mapped reads (FPKM), is on average one to two orders of magnitude lower than that of transcripts coding for mfp-1, mfp-2 and PreCol-NG, three byssal proteins produced by the cuticle, plaque and core glands, respectively (Table 1).

In the honeycomb worm, a total of 86 transcripts encoding tyrosinase-like enzymes were identified, among which 28 were (almost) full-length and, once translated *in silico*, comprised a tyrosinase domain. For *S. alveolata*, an additional filtering step was implemented based on the differential expression data reported in Buffet *et al.* (2018). Only transcripts overexpressed in the parathoracic region of the worms were considered, bringing the number of candidates down to 13 (Table 2). In this species too, most of the sequences were full-length, except for Salv-TYR1 (comp274293) which is incomplete in N-term and therefore lacks a signal peptide. Although the sequence of Salv-TYR3, the closest homologue to the catechol oxidase from *P. californica*, appeared to be complete in our transcriptome, no signal peptide

could be detected. Similar to mussels, the expression level of mRNAs encoding tyrosinases is much lower than that of mRNAs encoding the cement proteins Sa-1, Sa-2, and Sa-3A/B (Table 2).

Table 1. **List of tyrosinase sequences identified in the mussel *Mytilus edulis* after proteo-transcriptomic analyses (see text for details)**. Proteins with names in bold are the closest homologues to reference tyrosinases from *M. coruscus* and *P. viridis*. Indicated are the transcript ID form the foot transcriptome, the normalized expression level of the transcripts in the transcriptome (FPKM), the protein length in amino acid (with asterisks indicating incomplete sequences), the presence of a signal peptide, the peptide coverage of the translated transcripts from the MS-MS analysis (number of detected peptides in induced byssal threads [BT], distal foot tissues [DF] and proximal foot tissues [PF]), the position in the phylogenetic tree (Fig. 3), and top reciprocal BLAST hit. Three byssal proteins representative of the cuticle (Mfp-1), plaque (Mfp-2) and core (PreCol-NG) glands are included for comparison.

Sequence name	Clade in phylogenetic tree	Transcript ID	FPKM	Length (aa)	Signal peptide	Number of peptides identified by proteomic analyses			Reciprocal hit tBLASTn	
						BT	DF	PF	Name	Accession number
Tyrosinases										
Medu-TYR1	I	comp83177_c0_seq2	510.1	702	Y	30	17	17	*M. coruscus* byssal tyrosinase	KX268645
Medu-TYR2	I	comp79852_c1_seq2	59.2	678*	Y	10	10	2	*M. coruscus* byssal tyrosinase	KX268645
Medu-TYR3	I	comp74994_c0_seq1	8.1	489*	Y	1	/	/	*M. coruscus* byssal tyrosinase	KX268645
Medu-TYR4	I	comp80529_c1_seq1	32.9	686	Y	/	/	/	*M. coruscus* byssal tyrosinase	KX268645
Medu-TYR5	I	comp87104_c0_seq11	64.7	805	Y	/	/	3	*Mizuhopecten yessoensis* tyrosinase	XM_021518068
Medu-TYR6	I	comp75651_c0_seq1	17.83	802	Y	/	/	/	*Mytilus californianus* uncharacterized LOC127724782	XM_052231853
Medu-TYR7	II	comp83597_c2_seq1	12.6	580	Y	/	/	/	*M. californianus* uncharacterized LOC127706830	XM_052211537
Medu-TYR8	III	comp66244_c0_seq1	15.0	573	Y	/	/	/	*M. californianus* tyrosinase-like	XM_052233075
Medu-TYR9	III	comp83819_c0_seq1	34.9	628	Y	/	/	/	*M. californianus* uncharacterized LOC127706823	XM_052211524
Medu-TYR10	IV	comp85123_c3_seq1	215.0	696	Y	26	3	12	*M. coruscus* byssal tyrosinase	KP322726
Medu-TYR11	IV	comp76132_c0_seq1	80.9	648	Y	21	2	8	*M. coruscus* byssal tyrosinase	KP322726
Medu-TYR12	IV	comp83122_c1_seq1	29.9	694	Y	20	16	/	*M. coruscus* tyrosinase	KP876481
Medu-TYR13	IV	comp72405_c0_seq1	12.1	439	Y	1	/	/	*M. coruscus* tyrosinase-like	KP757802

Medu-TYR14	IV	comp76510_c0_seq1	1.3	393	Y	/	/	/	*M. californianus* tyrosinase-like	XM_052215114
Medu-TYR15	IV	comp79670_c1_seq2	192.8	561	Y	/	/	/	*M. galloprovincialis* catechol oxidase	MG975894
Medu-TYR16	IV	comp83799_c1_seq1	66.0	490	Y	/	/	/	*M. californianus* tyrosinase-like	XM_052240476
Medu-TYR17	IV	comp84313_c0_seq2	5.0	597	Y	/	/	/	*M. californianus* tyrosinase-like	XM_052207170
Other representative byssal proteins										
Mfp-1		comp48040_c0_seq1	2213.6	100*	N	/	/	/	*M.edulis* gene for polyphenolic adhesive protein	X54422
		comp63881_c0_seq2	0.38	165*	N	/	2	/	Mytilus edulis clone 21 foot protein 1 (fp-1) mRNA, complete cds	AY845258
Mfp-2		comp74249_c0_seq2	4210.2	508	Y	17	19	/	*M. edulis* clone 7 foot protein 2 (fp-2) mRNA	AY845261
PreCol-NG		comp75832_c0_seq1	1594.5	339	Y	12	/	5	*M. edulis* nongradient byssal precursor, mRNA	AF414454

Table 2. **List of tyrosinase sequences identified in the tubeworm *Sabelleria alveolata* after transcriptomic analyses (see text for details).** The protein with the name in bold is the closest homologue to the reference tyrosinase from *P. californica*. Indicated are the transcript ID from the transcriptome of the anterior part of the worm, the normalized expression level of the transcript in the transcriptome (FPKM), the differential expression of the transcript between the parathoracic part of the worm and the rest of its body (Log2FoldChange reported in Buffet *et al*., 2018), the protein length in amino acid (with asterisks indicating incomplete sequences), the presence of a signal peptide, the position in the phylogenetic tree (Fig. 3), and the top reciprocal BLAST hit. Four cement proteins are included for comparison.

Name	Clade in phylogenetic tree	Transcript ID	FPKM	Differential expression	Length (aa)	Signal peptide	Reciprocal hit tBLASTn	
							Name	Accession number
Tyrosinases								
Salv-TYR1	V	Comp274293_c0_seq4	6.1	-7.29	470*	N	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR2	V	Comp275725_c0_seq1	8.1	-11.85	476	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR3	V	Comp278284_c0_seq2	9.2	-8.84	432	N	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR4	V	Comp264814_c0_seq1	1.7	-4.89	434	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR5	V	comp276298_c0_seq1	82.3	-4.43	477	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR6	V	comp277654_c0_seq1	97.5	-4.39	458	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR7	V	comp239180_c0_seq1	130	-4.44	489	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR8	V	comp275276_c0_seq1	27.1	-3.86	515	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR9	V	comp279307_c0_seq1	52	-4.42	471	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR10	V	comp276771_c7_seq1	136.4	-4.21	450	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2
Salv-TYR11	V	comp269290_c0_seq1	24.25	-7.23	515	Y	*Lingula anatina* putative tyrosinase-like	XM_024075990.1
Salv-TYR12	V	comp280500_c0_seq2	8.3	-4.07	504	Y	*Phragmatopma californica* tyrosinase-like protein	JN607213.2

Salv-TYR13	V	comp273963_c1_seq5	141.7	-4.08	402	Y	*Lingula anatina* putative tyrosinase-like	XM_013539498.1
Other representative cement proteins								
Sa-1		comp225468_c0_seq2	30403	-4.25	112	Y	NA	
Sa-2		comp271660_c3_seq1	23866.7	-3.95	102	Y	NA	
Sa-3a		comp271458_c0_seq5	1061.7	-4.52	298	Y	NA	
Sa-3b		comp267107_c0_seq3	1867.6	-4.96	213	Y	*Sabellaria alveolata* mRNA for cement precursor protein 3B	HE599639.1

3.2. Proteomic analyses

We carried out protein analyses on different samples from *M. edulis.* The secretion of byssal threads was induced by injecting KCl at the base of the foot. These freshly secreted threads were collected with fine forceps and the proteins comprising them were extracted with an 8M urea solution in 5% acetic acid and analyzed by de novo peptide sequencing in mass spectrometry (ESI-MS/MS). In addition, mussel feet were dissected and cut in half to separate the foot tip from the basal part. Foot proteins were extracted with 4% sodium dodecyl sulfate (SDS) and were also analyzed by MS/MS after trypsin digestion. By comparing the mass spectrometry results with the foot transcriptome, a total of 7 tyrosinases were identified with at least two peptides, with five detected in both the induced threads and foot samples, with a high peptide coverage, and two detected exclusively in the foot samples, but with only 3 peptides each (Table 1). Medu-TYR1, for example, exhibited the highest peptide coverage and its corresponding transcript was also the most abundant in the transcriptome compared to other candidates (Table 1). However, the correlation between abundance at the transcript and protein levels does not hold true for the other candidates. Among the tyrosinases identified in foot tissues, 2 were expressed exclusively in the tip (Medu-TYR8 and 12), 2 in the basal part (Medu-TYR5 and 10), and 3 in both parts (Medu-TYR1, 2 and 11). As for the byssal proteins used for comparison, two were detected (Mfp-2 and preCol-NG) in all samples and their distribution in the foot corresponded to their expected expression pattern, only in the tip for mfp-2 and in the basal part of the foot for preCol-NG (Table 1). Due to its tandemly repeated sequence making assembly difficult, mfp-1 was represented in our transcriptome with two partial transcripts. Only two peptides corresponding to the sequence encoded by one of these transcripts could be detected by mass spectrometry.

In *S. alveolata,* the proteomic analysis was performed on tube fragments constructed by the worms using glass beads. Indeed, in the laboratory, isolated individuals can rebuild their tubes with different materials, including clean glass beads. Freshly constructed tube fragments were collected, and proteins were extracted from the cement dots present on the surface of glass beads using 7 M guanidine hydrochloride. Despite these highly denaturing conditions, no protein was detected in mass spectrometry, suggesting that the adhesive secretion might be highly crosslinked and therefore posed challenges for protein extraction.

3.3. Phylogenetic analyses

To gain a comprehensive understanding of the relationships between tyrosinases, we investigated protein sequences containing the tyrosinase domain (pfam00264) from various spiralian species. These sequences were retrieved from the NCBI database. They include 30 proteins from Annelida, 8 from Brachiopoda, 375 from Mollusca, and 63 from Platyhelminthes. Additionally, we incorporated sequences obtained through the proteo-transcriptomic analyses of our two model species, specifically 20 predicted proteins from *M. edulis* (3 sequences found in the mantle were added to the 17 previously identified) and 28 from *S. alveolata* (Tables 1 and 2) (Supplementary Table 1). A sequence similarity-based clustering analysis was carried out using CLANS (Frickey & Lupas, 2004) to explore potential similarities between all these tyrosinases. An all-against-all BLASTp was conducted using the scoring matrix BLOSUM62 and linkage clustering was performed with a maximum e-value of $1E^{-20}$ to identify coherent clusters. The clustering was initially performed in 3 dimensions and then projected into 2 dimensions to generate the illustration shown in Fig. 2. The darker connections between the dots indicate higher similarity between the proteins based on the BLASTp e-values (Pearson, 2013).

The CLANS analysis revealed that spiralian tyrosinases form several clearly distinct clusters, each comprising sequences that share a high similarity among them as evidenced by the low e-values associated to the lines connecting them (Fig. 2). All the tyrosinase sequences from each phylum are generally grouped together, although a few isolated dots are noticeable, usually corresponding to very short (partial) sequences or very long sequences (>3000 amino acids) comprising additional domains (e.g., kielin/chordin-like domain). The phylum Mollusca is an exception as its sequences were divided into two clusters: one comprising sequences with only a tyrosinase domain and one consisting of sequences containing both a tyrosinase (pfam00264) and an hemocyanin domain (pfam14830). Hemocyanins are extracellular proteins involved in oxygen transport and are found in the phyla Arthropoda and Mollusca (Van Holde & Miller, 1995; Salomon *et al*., 1996). Previous studies have demonstrated that both hemocyanins and tyrosinases belong to the type 3 copper protein family and share a common ancestor (Drexel *et al*., 1987; Burmester & Schellen, 1996). However, structural modifications at the binuclear copper active site underlie the divergent evolution of tyrosinase and hemocyanin functions (Aguilera *et al*., 2013). This functional divergence may explain why all the Mollusca sequences are not grouped in the same cluster.

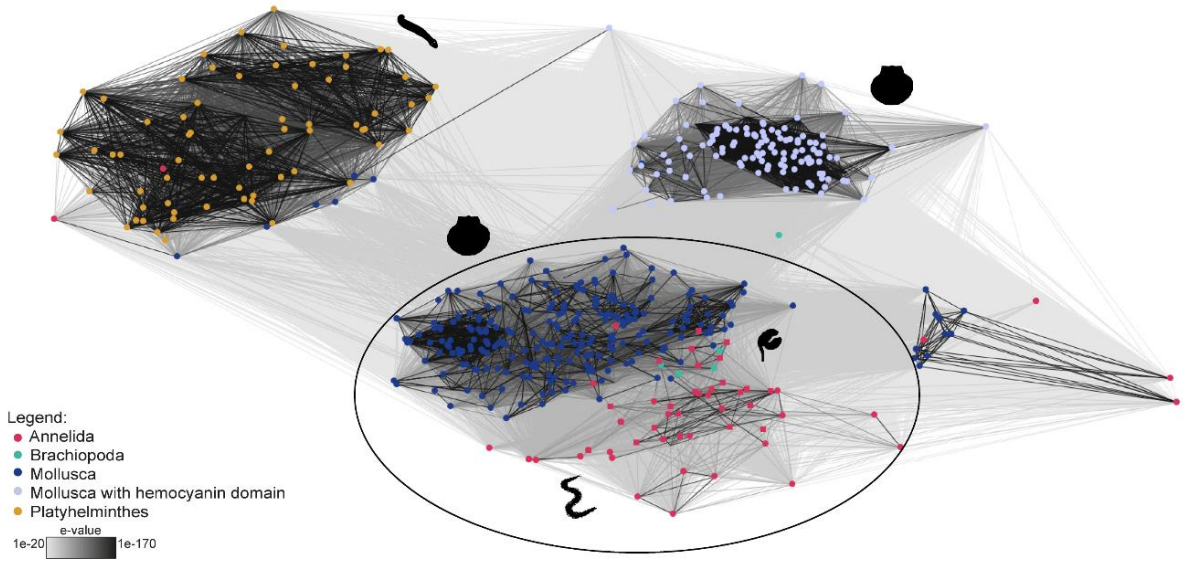


Figure 2. **Cluster analysis of tyrosinase sequences from Spiralia**. Sequence-similarity-based clustering approach based on BLASTp e-values with the tyrosinase sequences identified in the present work and all the spiralian sequences containing a pfam00264 domain retrieved from the NCBI database. This analysis shows that the different phyla tend to cluster together. The sequences circled were selected for phylogenetic analyses.

In the CLANS analysis, tyrosinase sequences from Platyhelminthes form a well-separated cluster whereas those from Lophotrochozoa (Brachiopoda, Annelida and Mollusca) are grouped together (Fig. 2). To delve deeper into their evolutionary relationships, the 312 tyrosinase protein sequences from the lophotrochozoan super-cluster were subjected to a phylogenetic analysis using molluscan hemocyanins as outgroup (Fig. 3, Supplementary Fig. 2). On the phylogenetic tree generated, 5 main clades (numbered from I to V on Fig. 3) can be distinguished. Clades I to IV group together the sequences from molluscs: clades I, II and IV contain only tyrosinases from bivalves, while clade III includes sequences from all classes of molluscs, from gastropods to cephalopods. Clade V contains a few oyster tyrosinase sequences (Bivalvia, Ostreoida), but comprise mostly non-molluscan sequences. Most of the sequences from the phylum Annelida are grouped with those from Brachiopoda in this clade, except some of the sequences from the polychaete *Owenia fusiformis* (Delle Chiaje, 1844) which are in clade III. The reference byssal tyrosinase sequences from the mytilids *M. coruscus* and *P. viridis*, as well as their closest homologues from *M. edulis* (Table 1) are localized either in cluster I (e.g., Medu-TYR1 to 4; Fig. 3) or in cluster IV (Medu-TYR10 to 15; Fig. 3). However, a few of the sequences shortlisted in the blue mussel are also present in clades II and III (Table 1). The reference cement tyrosinase from *P. californica* and all the sequences from *S. alveolata* are in cluster V. It should be noted that the phylogenetic tree generated from these analyses

showed low support for some nodes (Fig. 3), likely due to the high level of conservation of residues surrounding the copper-binding sites and globally the short protein-based alignment. Similar results have been observed in previous studies (Aguilera *et al*., 2014; Buffet *et al*., 2018).

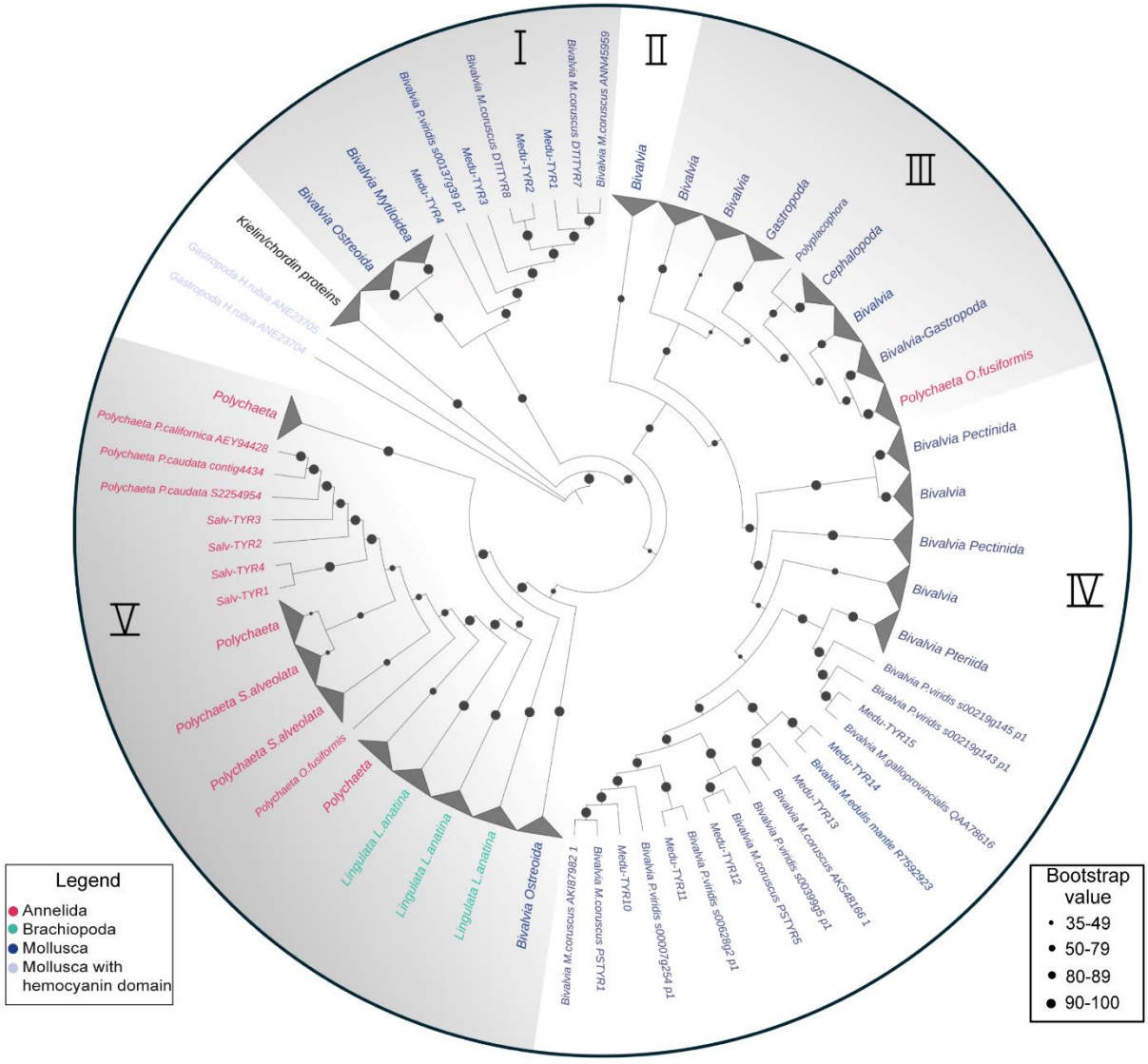


Figure 3. **Evolution of tyrosinases in Lophotrochozoa.** Maximum likelihood phylogenetic tree showing the distribution of tyrosinase sequences, based on circled sequences selected in the CLANS analysis. The sequences are distributed among 5 main clades, numbered I to V. The reference byssal tyrosinase sequences from *M. coruscus* and *P. viridis* and their orthologues in *M. edulis* (Table 1) are localized in clusters I and IV, whereas the reference cement tyrosinase from *P. californica* and all sequences from *S. alveolata* (Table 2) are in cluster V. The bootstrap values are shown in the legend in the box. See also Supplementary Figure 2 for the corresponding detailed tree.

3.4. *In situ* hybridization experiments

To further investigate the role of the identified enzymes in adhesive protein maturation, the localization of their coding mRNAs in the tissues was determined using *in situ* hybridization (ISH). DIG-labelled RNA probes were designed based on sequences retrieved from the transcriptomes and used on tissue sections from both mussels and honeycomb worms, as well as on whole-mount preparations of the mussel's feet, to determine whether the tyrosinases are expressed in the adhesive-producing gland cells or in other cell types. Controls were performed using sense RNA probes, as well as without probes or without antibody (Supplementary Fig. 3).

In the case of the blue mussel, we selected tyrosinase sequences showing the highest similarity to the reference sequences and/or coding for proteins detected with a minimum of two peptides in the induced byssal threads by MS/MS analyses. This corresponds to nine candidates: Medu-TYR1 to 5, Medu-TYR10 to 12, and Medu-TYR15 (Table 1). *In situ* hybridization was performed on the entire foot cut open in two halves along the central frontal plane (whole mount; see Fig. 4) but also on several transverse sections through the foot to make sure to visualize all specific foot glands, namely the core, cuticle, and plaque glands. For each section processed for ISH, a directly consecutive section was stained with Heidenhain's azan to validate the identification of the glands. The results of the ISH experiments are illustrated in Fig. 4 and Supplementary Fig. 4. Among the nine candidates, one was found to be exclusively localized in the plaque gland (Medu-TYR12), one in the core gland (Medu-TYR10), and one in the cuticle gland (Medu-TYR11). On the other hand, five candidates were expressed in at least two glands: both the plaque and the cuticle glands (Medu-TYR4 and 15), both the plaque and core glands (Medu-TYR2 and 5), or both core and cuticle glands (Medu-TYR1). No probe could be produced for the candidate Medu-TYR3.

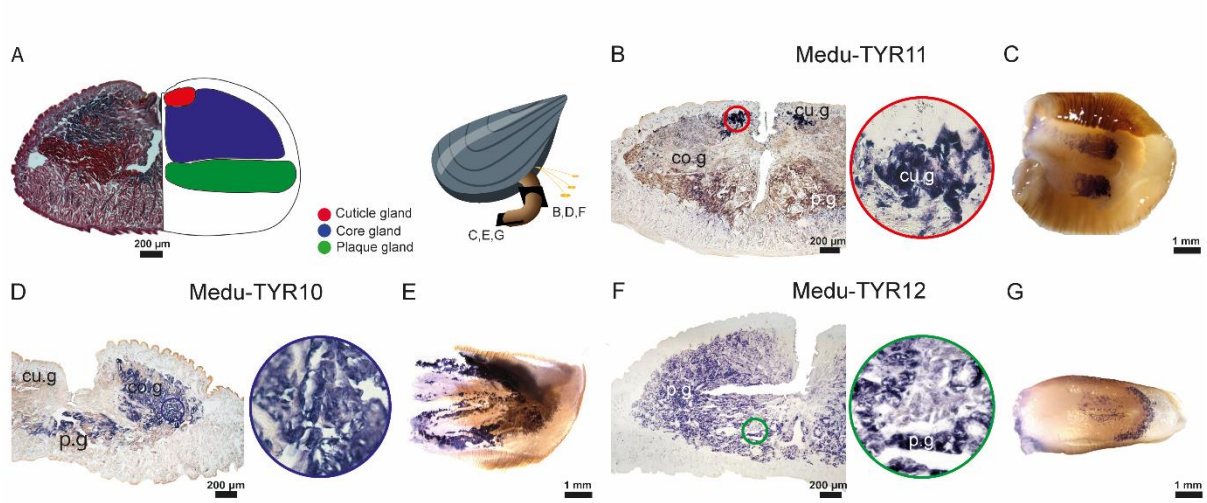


Figure 4. Examples of the expression patterns of selected tyrosinase transcripts in the foot of *Mytilus edulis* (A) Illustration showing both a transverse histological section of a foot stained with Heidenhain's azan stain (left), and a diagram with the arrangement of the different glands in the foot tissues (right). (B-G) Section (B, D, F) and whole mount (C, E, G) *in situ* hybridization of the mRNAs coding for the three tyrosinase candidates. Insets in circles show a zoom on the labeled gland. Abbreviations: cu.g – cuticle gland; co.g – core gland; p.g – plaque gland. See also Supplementary Figure 3 and 4 for controls and other transcripts.

For the honeycomb worm, four transcripts were selected: those exhibiting the highest similarity to the reference sequence from *P. californica* (Fig. 3), among which three were also the most highly differentially expressed in the parathoracic region of the worm (Table 2). The four tyrosinase candidates (Salv-TYR1 to 4) were exclusively expressed in the cement glands (Fig. 5). Moreover, they were all expressed within both cells with heterogeneous granules and cells with homogeneous granules (Fig. 5). The two types of granules can easily be distinguished at high magnification.

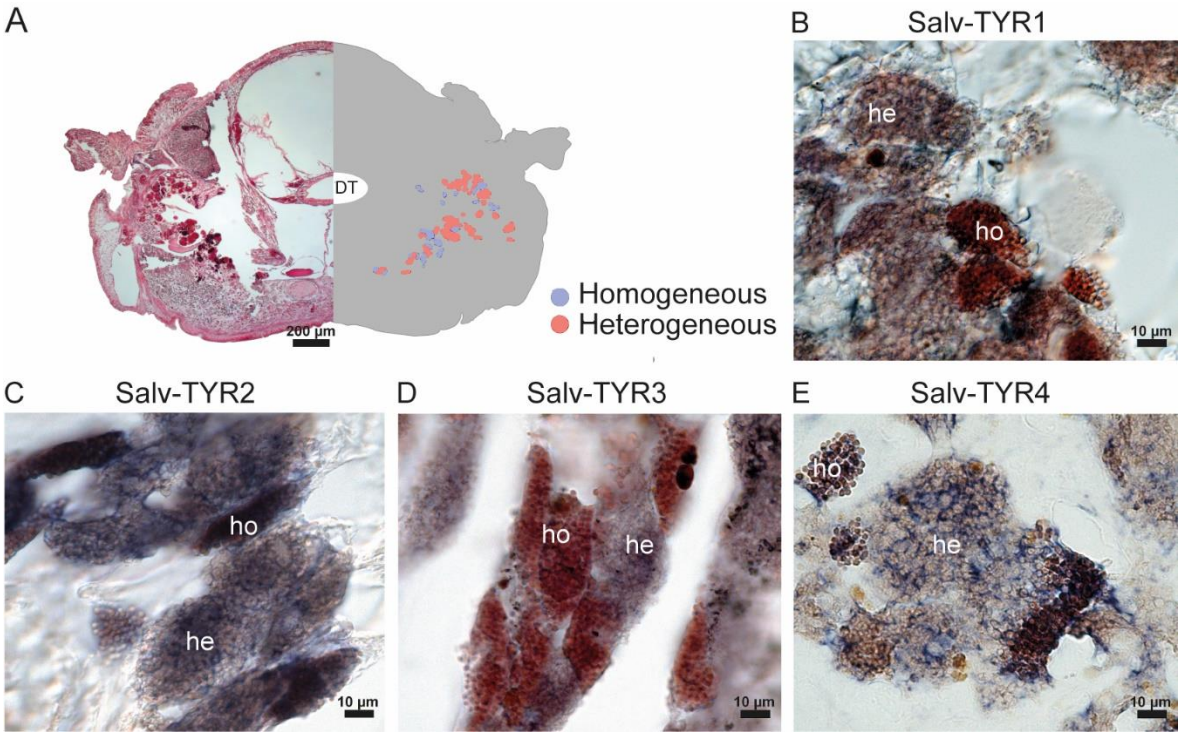


Figure 5. Expression patterns of selected tyrosinase transcripts in the parathoracic region of *Sabellaria alveolata*. (A) Illustration showing both a transverse section of the parathoracic part stained with Heidenhain's azan stain (left), and a diagram showing the arrangement of the cement glands (right). (B-E) *In situ* hybridization of the mRNAs coding for four tyrosinase candidates. Abbreviations: DT - digestive tract; he – cement gland with heterogeneous granules; ho – cement gland with homogeneous granules. See also Supplementary Figure 3 for controls.

4. Discussion

4.1. Diversity of tyrosinases potentially involved in adhesive protein maturation

Various marine invertebrates rely on quinone-tanned biomaterials as durable glues or protective varnishes (Waite, 1990). Two well-studied examples of such materials are the byssus produced by mussels and the cement secreted by tube-building worms (Waite, 1990). These materials contain a catechol known as DOPA, or 3,4-dihydroxyphenylalanine, a modified amino acid which is widely present in nature. DOPA is essential for bonding with the substrate and providing cohesive cross-links within the adhesive materials of these organisms. The enzymes responsible for its production are the tyrosinases which play a crucial role in various biological functions. Although DOPA has been extensively studied in marine adhesion, the specific enzymes responsible for its production are still not well understood. However, several studies have highlighted the fact that not one but several tyrosinases may be involved in the maturation of adhesive proteins in a single organism (Guerette *et al.*, 2013; Qin *et al.*, 2016; Buffet *et al.*, 2018). In this study, our first objective was to identify the tyrosinase sequences

involved in the adhesive systems of mussels and tubeworms by comparing a set of reference sequences known to be present in the adhesive secretions of some species with the transcriptomes of our two model species, the blue mussel *M. edulis* and the honeycomb worm *S. alveolata*. For both species, more than 80 different tyrosinase-like sequences were retrieved by the BLAST searches, but these numbers were greatly reduced when only (almost) full-length proteins with a tyrosinase domain were considered. Moreover, differential expression, mass spectrometry analyses and *in situ* hybridization experiments were used to pinpoint the candidate expressed in adhesive glands and secreted.

In the blue mussel, 17 candidates were identified by the *in silico* analyses, among which 7 were shown to be present in the byssus and/or in the foot tissues by mass spectrometry analyses. In particular, 5 tyrosinases were detected in induced byssal threads with a high peptide coverage (Medu-TYR1 and 2, and Medu-TYR10 to 12). The mRNAs coding for these 5 proteins, but also those coding for 3 other candidates sharing high homology with reference tyrosinases from *M. coruscus* and *P. viridis*, were localized in the foot glands. Two patterns of expression were observed: three candidates (Medu-TYR10 to 12) are each specific in a single gland type while the other five (Medu-TYR1, 2, 4, 5 and 15) have been localized in different glands (Fig. 6). The gland-specific mRNA expression of these different tyrosinases corresponds to the localization of the corresponding protein in the foot (distal half, proximal half, or both) for candidates which have been detected in mass spectrometry (Table 1; Fig. 6). It also matches the distribution of their homologues within the byssus of *M. coruscus* (Table 1). For instance, Medu-TYR12, whose mRNA is specifically expressed in the plaque gland, is the closest homologue to a tyrosinase-like protein (encoded by cDNA KP876481; Table 1) identified in the byssal plaque in *M. coruscus* (Qin *et al*., 2016). It is not surprising to find tyrosinases in the three different foot glands involved in byssal threads synthesis as they all produce DOPA-containing proteins (Fig. 6) (Waite, 2017). The plaque gland is known to contain more than ten different proteins (DeMartini *et al*., 2017), including mfp-3 and -5 which are the byssal proteins with the highest DOPA content (20 and 30 mol%, respectively) (Papov *et al*., 1995; Waite & Qin, 2001). The cuticle gland produces mfp-1, which contains 15 mol% of DOPA, as well as a handful of other, less-characterized proteins (Waite & Tanzer, 1981; DeMartini *et al*., 2017). Finally, the preCols and thread matrix proteins secreted by the core gland also contain DOPA, although in lower amounts (Qin *et al*., 1997; Sagert & Waite, 2009).

In the honeycomb worm, 28 candidate tyrosinase transcripts were retrieved from the transcriptomic analysis. The comparison of these sequences with the differential transcriptome

of Buffet *et al*. (2018) for the same species allowed us to reduce the list to 13 candidates. These transcripts are highly overexpressed in the parathorax, the region of the worm containing the cement glands. A proteomic analysis was conducted on tubes reconstructed by the worm using glass beads but, unfortunately, no tyrosinase or cement protein was detected, suggesting that the proteins could not be extracted from the cement spots binding the beads. Four transcripts, among the most differentially expressed and showing the highest homology with the cDNA encoding the reference tyrosinase from the sandcastle worm *P. californica* (JN607213.2), were selected for localization by *in situ* hybridization. These tyrosinase mRNAs were expressed in both types of unicellular cement glands, the cells with homogeneous granules and the cells with heterogeneous granules, similarly to what was observed in *P. californica* (Wang & Stewart, 2012). There was therefore no gland specificity for the four candidates we selected. In *P. californica*, at least two adhesive proteins, Pc-1 and Pc-2, are known to contain DOPA (about 10 and 7%, respectively (Waite *et al*., 1992)). The former is produced by cement glands with heterogeneous granules and the latter by glands with homogeneous granules (Wang & Stewart, 2012). Their homologues, Sa-1 and Sa-2 have been identified in *S. alveolata* (Becker *et al*., 2012) and preliminary results suggest that their distribution in the cement glands corresponds to the one in *P. californica* (unpublished data). Thus, in the honeycomb worm too, each type of cement gland would produce at least one DOPA-rich protein, which would explain the expression of tyrosinases in both glands.

Although the high diversity of tyrosinase candidates in both model species is suggestive of a variety of functions, linking their sequences with their different functionalities (monooxygenase vs oxidase) may prove difficult. Tyrosinases and catechol oxidases share a very similar active site architecture in which three distinct oxidation states have been identified, oxy-, deoxy-, and met-states, based on the structure of the bicopper structure of the active centre (Ramsden & Riley, 2014). Transitions from one state to another lead to the molecular mechanisms involved in the monophenolase or diphenolase catalytic activity (Ramsden & Riley, 2014). However, some enzymes appear to show only the oxidase activity (Pretzler & Rompel, 2018). Moreover, while the active site of tyrosinases is highly conserved, variations exist in their sequences, size, glycosylation, and activation (Jaenicke & Decker, 2003). To date, the structural motifs responsible for the different activities remain elusive (Pretzler & Rompel, 2018). Only homology with enzymes of known function or expression pattern could therefore be used to investigate the function of the new tyrosinase candidates from *M. edulis* and *S. alveolata*.

In mussels, DOPA plays important interfacial adhesive and bulk cohesive roles during byssus fabrication (Lee *et al*., 2011; Waite, 2017). Priemel *et al.* (2020) have identified two distinct DOPA-based cross-linking pathways, one achieved by oxidative covalent cross-linking and the other by the formation of metal coordination interactions under reducing conditions. The former takes place in the thread core while the latter occurs in the cuticle, plaque foam and plaque-surface interface (Priemel *et al*., 2020). In the core gland, preCol-D was shown to possess several DOPA residues at the N-terminus (Qin *et al*., 1997) which, after secretion, are oxidized into dopaquinone for covalent cross-link formation (Priemel *et al*., 2020). It is therefore tempting to postulate that the core gland-specific tyrosinase, Medu-TYR10, which is also one of the most highly expressed, could have a catechol oxidase activity. A catechol oxidase was extracted from the foot and byssus of *M. edulis* by Waite (1985), but it is not possible to link it to any of our candidates as no sequence was reported. On the other hand, the DOPA-rich proteins from the cuticle and plaque glands should be prevented from oxidation after secretion to promote intermolecular DOPA-metal coordinated bonds (Schmitt *et al*., 2015a; Priemel *et al*., 2020). This is possible thanks to a locally low pH and the presence of reducing conditions, usually provided by the sulfhydryl groups of cysteine-rich proteins such as mfp-6 (Lee *et al*., 2011). Based on this reasoning, Medu-TYR11 (cuticle gland) and Medu-TYR12 (plaque gland) should only display a monophenolase activity. Unexpectedly, Medu-TYR15, whose corresponding transcript is expressed in both the plaque and cuticle glands, is the closest homologue to a catechol oxidase from *M. galloprovincialis* (Table 1). This catechol oxidase, when expressed recombinantly in bacteria, showed no catalytic activity but instead exhibited a thiol-dependent antioxidant activity suppressing DOPA oxidation (Wang *et al*., 2019). However, Medu-TYR15 was not detected in the byssus or the foot by our proteomic analyses.

In tubeworms, it was demonstrated that the tyrosinase identified in *P. californica* catalyzes only catechol oxidation. This observation was made through substrates and inhibitors profiling on the isolated secretory granules and secreted cement (Wang & Stewart, 2013). As Salv-TYR1 to 4 are close homologues to this catechol oxidase and display the same expression pattern in both cement glands, it is likely that they would also show a diphenolase activity. In *P. californica*, the oxidation of DOPA to dopaquinone leads to the formation of cysteinyl-DOPA covalent cross-links within the cement (Zhao *et al*., 2005).

4.2. Evolution of tyrosinases in Lophotrochozoa

The high number and diversity of tyrosinases identified in *M. edulis* and *P. californica* prompted us to examine the phylogenetic relationship they have together and with other lophotrochozoan tyrosinases.

Firstly, we built a cluster map of all spiralian sequences containing a tyrosinase domain retrieved from the NCBI nr database. Tyrosinase sequences from Platyhelminthes formed a cluster clearly separated from another large cluster comprising sequences from Brachiopoda, Annelida, and Mollusca. This reflects the evolution of type-3 copper proteins in animals in which three ancestral subclasses (α, β and γ) have been described, with differential losses or expansions of one or more of these subclasses in specific phyla (Aguilera *et al*., 2013). Platyhelminthes possess γ-subclass, transmembrane tyrosinases whereas the three other phyla possess α-subclass, secreted tyrosinases (Aguilera *et al*., 2013). Molluscan hemocyanins also belong to the α-subclass (Aguilera *et al*., 2013), but form a well-separated cluster compared to tyrosinases.

Secondly, a phylogenetic tree was constructed to examine the relationships between lophotrochozoan tyrosinases, using molluscan hemocyanins as the outgroup. The tyrosinase tree is fairly complex, with five main clades that do not necessarily reflect the main taxonomic groups. Mollusc, annelid and brachiopod tyrosinases appear to be paraphyletic and their sequence variation would not be solely influenced by the evolutionary distance between lineages. Indeed, three of the five clades in the tree we generated (clades I, II and IV) contain exclusively sequences from bivalves while the other two clades (III and V) gather sequences from different phyla or mollusc classes. In their study about the evolution of the tyrosinase gene family in bivalve molluscs, Aguilera *et al*. (2014) reported that tyrosinase sequences were separated into two clades: an ancestral clade (A) comprising tyrosinases from all mollusc classes and a bivalve-specific clade (B). Based on the sequences shared by the two studies, our clades I, II, III and V would correspond to their clade A, and our clade IV to their clade B. The different tree topology we obtained could be linked to the inclusion of annelid and brachiopod sequences in our tree, a different trimming of the sequences, or a different rooting of the tree. It should also be noted some of the basal nodes in our tree are poorly supported.

The high number of transcripts obtained during our *in silico* analyses supports the hypothesis that the tyrosinase gene family has undergone large independent expansions in bivalve molluscs (Aguilera *et al*., 2014) as well as in sabellariid polychaetes (Buffet *et al*.,

2018). This apparently also includes tyrosinases potentially involved in the maturation of adhesive proteins, with 6 and 4 enzymes identified in *M. edulis* and *S. alveolata*, respectively, as shown by proteomic data and/or *in situ* hybridization results (Fig. 6). Mussel byssus and tubeworm cement tyrosinase sequences are separated in distinct clusters. This suggests an independent functional evolution of tyrosinases involved in the maturation of bioadhesives in these two lineages. Interestingly, the mussel tyrosinases are also distributed in two different clusters (I and IV) and enzymes from both clusters are co-expressed in the same foot gland. The identification of byssal tyrosinases in these two different clusters is congruent with the results of Aguilera *et al.* (2014) indicating an early gene duplication in bivalve evolution. These genes would then have undergone further independent duplication and divergence to acquire new gene functions.

4.3. Implication for the development of biomimetic adhesives

Marine adhesives display impressive performances in their natural context and, therewith, the potential to inspire novel adhesives working in fluid environments, including the human body (Almeida *et al*., 2020). Marine adhesive proteins have therefore been exploited to produce increasingly complex biomaterials with added functionalities, such as wearable electronics or drug delivery systems (Sun *et al*., 2022). Many of these materials utilize recombinant proteins, which are generally seen as the closest mimics of marine adhesive proteins and can be synthesized in sufficient quantities for incorporation in biomaterials (Lutz *et al*., 2022). Recombinant DNA technology also comes with the possibility to alter and rearrange protein sequences by genetic engineering to create new truncated or chimeric proteins.

Because the mussel's byssus is the best-characterized marine bioadhesive, it is in these organisms that most of the recombinant adhesive proteins have been produced (Harrington *et al*., 2018; Wang & Scheibel, 2018). Many mfps from the cuticle and plaque glands, as well as the preCols from the core gland, have been produced in heterologous host cells such as bacteria or yeasts (Wang & Scheibel, 2018). However, in all these recombinant proteins, important native post-translational modifications are missing, such as the hydroxylation of tyrosine residues into DOPA. Protocols have thus been developed for the *in vitro* conversion of tyrosine residues to DOPA utilizing a mushroom tyrosinase, but they exhibit low modification yields which can limit underwater adhesion (Kitamura *et al*., 1999; Hwang *et al*., 2004). In another approach a co-expression system was used in bacteria, with the concomitant production of mfps and bacterial or mushroom tyrosinases with a dual vector system (Choi *et al*., 2012; Yao *et al*., 2022). The *in vivo* modification efficiency was higher than that in vitro, leading to an increased

adhesive strength (Choi *et al*., 2012). Under the assumption that tyrosinases produced in specific glands have evolved to specifically modify the DOPA-containing proteins synthesized and stored in that gland, we hypothesize that using the appropriate tyrosinase to target and modify a given recombinant protein might give improved DOPA modification yield and thus materials performance. For example, one could co-express a mussel adhesive protein with the specific mussel tyrosinase expressed in the same gland (e.g., mfp-3 or mfp-5 with Medu-TYR12). Before this can become a reality, however, the tyrosinases identified in the present study need to be produced recombinantly, better characterized and their function confirmed by *in vitro* assays.

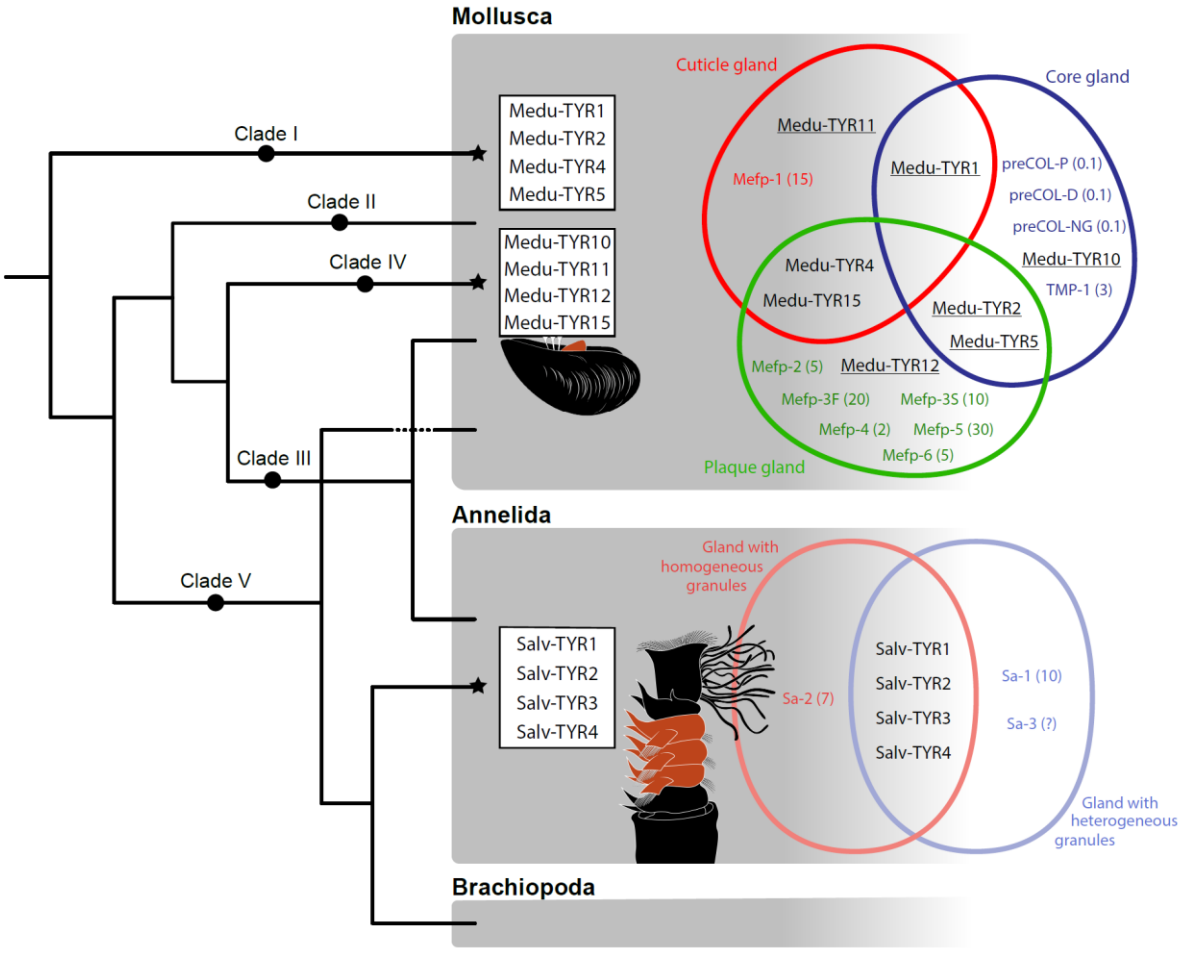


Figure 6. Tyrosinases involved in the adhesive systems of the mussel *Mytilus edulis* and the tubeworm *Sabellaria alveolata*. Simplified tyrosinase phylogenetic tree in which stars indicate the clades comprising enzymes expressed in adhesive glands. The gland specificity of each tyrosinase, as highlighted by *in situ* hybridization experiments, is illustrated on the right. In the mussel, the ISH localization of several candidates (underlined) is supported by mass spectrometry data even though the experimental protocol used (foot distal part vs foot proximal part) does not allow a precise gland assignment. The main adhesive proteins produced by the different glands have been included, as well as their DOPA content in brackets, expressed in mol% (data from different species Waite *et al*., 1992; Waite, 2017).

5. Conclusion

Honeycomb worms and blue mussels are frequently found together in the intertidal zone, where they encounter similar environmental challenges. Both organisms rely on a DOPA-based adhesive system and our study has identified a catalogue of tyrosinase enzymes involved in the maturation of adhesive proteins in *M. edulis* and *S. alveolata*. The diversity of tyrosinases highlighted in the two species suggests the coexistence of different functions (monophenol monooxygenase or catechol oxidase activity) or different substrate specificities. However, the exact role of the different enzymes needs to be further investigated. Phylogenetic analyses support the hypothesis of independent expansion and parallel evolution of tyrosinases involved in adhesive protein maturation in both lineages, supporting the convergent evolution of their DOPA-based adhesion. These results contribute to our understanding of the molecular basis of adhesion mechanisms in marine organisms, but also pave the way towards the use of specific tyrosinases in development of biomimetic adhesives.

Limitations of the study

Our study presents certain limitations that should be highlighted. We did not explore the complete tyrosinase repertoire of the blue mussel (*Mytilus edulis*) and the honeycomb tubeworm (*Sabellaria alveolata*) based, for example, on complete chromosome-scale genomes. Rather, using a set of reference tyrosinase sequences from the literature, we specifically focused on tyrosinase sequences retrieved from a foot transcriptome for the mussel and from a parathoracic region transcriptome for the tubeworm– tissues relevant for the material formation processes in question. While we identified several candidates expressed in adhesive-secreting tissues, we might have missed species-specific or highly derived tyrosinases using this approach. Moreover, additional *in situ* hybridization experiments could be conducted to highlight the expression patterns of other candidates from our list. Although we provided evidence of tyrosinase production and secretion at the protein level through mass spectrometry analyses, the use of specific antibodies could further validate the expression profiles of the identified proteins. Another limitation is that our study does not address the functional roles of the tyrosinases investigated. *In vitro* functional characterization of the tyrosinases of interest will be crucial for validating their specific molecular functions and could potentially lead to the emergence of new avenues in bioinspired research and biomaterial development.

Acknowledgements

We thank Dominique Baiwir from the GIGA Proteomics Facility (ULiège, Belgium) for the mass spectrometry-based analyses conducted on mussel feet, as well as Antoine Flandroit, Nathan Puozzo and Paolo Rosa for their help with microphotography and figures. We also acknowledge the support of the Research Institute for Biosciences of UMONS. This work was supported by the Fund for Scientific Research of Belgium (F.R.S.-FNRS) through (1) a FRIA doctoral fellowship to E.D., (2) a CR postdoctoral fellowship to B.M., and (3) a research project (PDR T.0088.20). PJL would also like to thank the Labex DRIIHM; the French program "Investissements d'Avenir" (ANR-11-LABX-0010), which is managed by the French National Research Agency (ANR). J.D. is financially supported by a F.R.S.-FNRS research project (PDR T.0169.20) granted to UMONS and ULiège. P.F. is Research Director of the F.R.S.-FNRS.

Chapter I

Supplementary materials for

Diversity and evolution of tyrosinase enzymes involved in the adhesive systems of mussels and tubeworms

Emilie Duthoo[1], Jérôme Delroisse[1,2], Barbara Maldonado[1,3], Fabien Sinot[1], Cyril Mascolo[4], Ruddy Wattiez[4], Pascal Jean Lopez[5], Cécile Van de Weerdt[3], Matthew J. Harrington[6] and Patrick Flammang[1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cellular and Molecular Immunology, GIGA, University of Liège, 11 avenue de l'hôpital, 4000 Liège, Belgium

[3] Molecular Biomimetic and Protein Engineering Laboratory, GIGA, University of Liège, 11 avenue de l'hôpital, 4000 Liège, Belgium

[4] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[5] UMR Biologie des Organismes et des Ecosystèmes Aquatiques, MNHN/CNRS-7208 Sorbonne Université/IRD-207/UCN /UA, 43 rue Cuvier, Paris 75005, France

[6] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

This part includes:

Supplementary Figs. 1, 3, 4 and 5

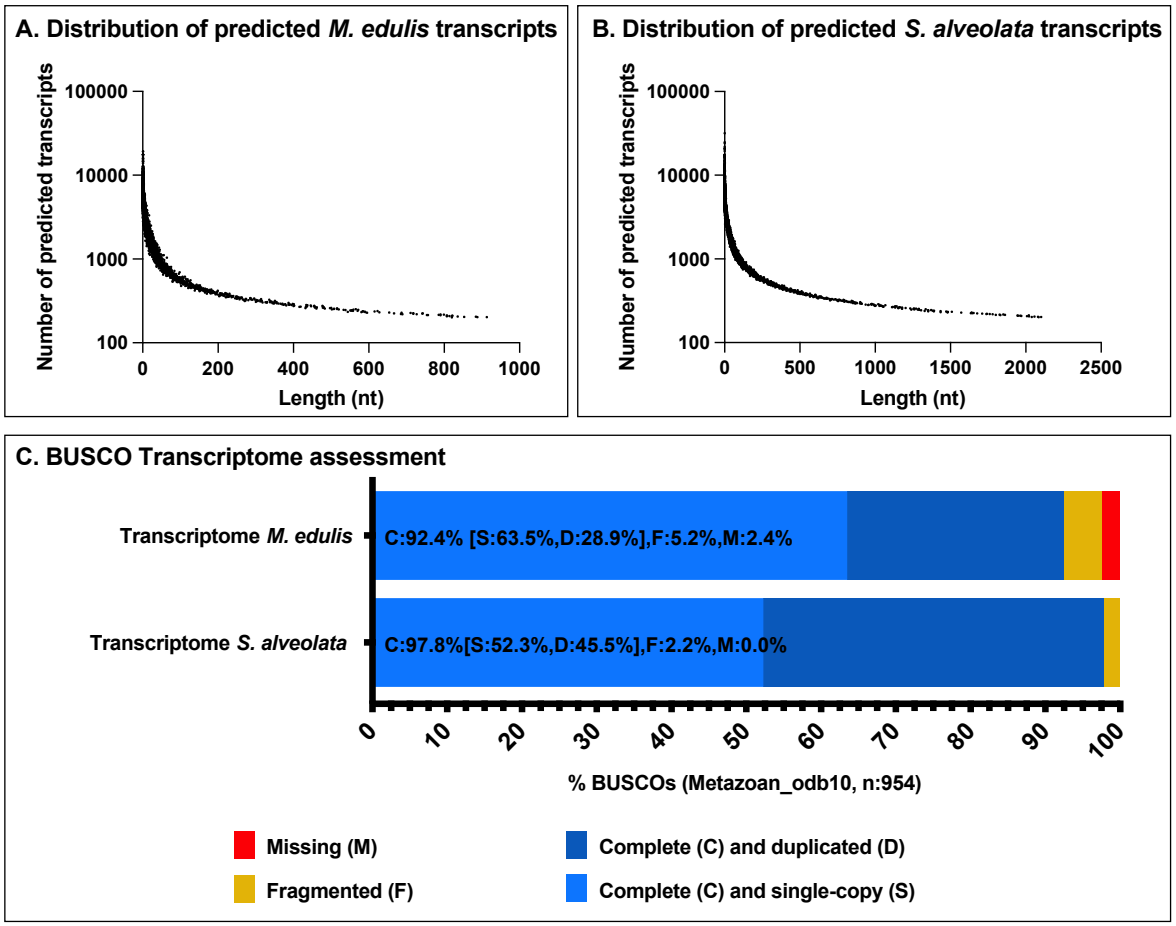
Other Supplementary Materials for this manuscript include the following:

Supplementary Figs. 2. Detailed maximum likelihood phylogenetic tree (corresponding to main text Fig. 3) showing the distribution of tyrosinase sequences in Lophotrochozoa, based on sequences listed in Suppl. Table 2. The bootstrap values are indicated at each node. The scale bar indicates the substitution rate per site.

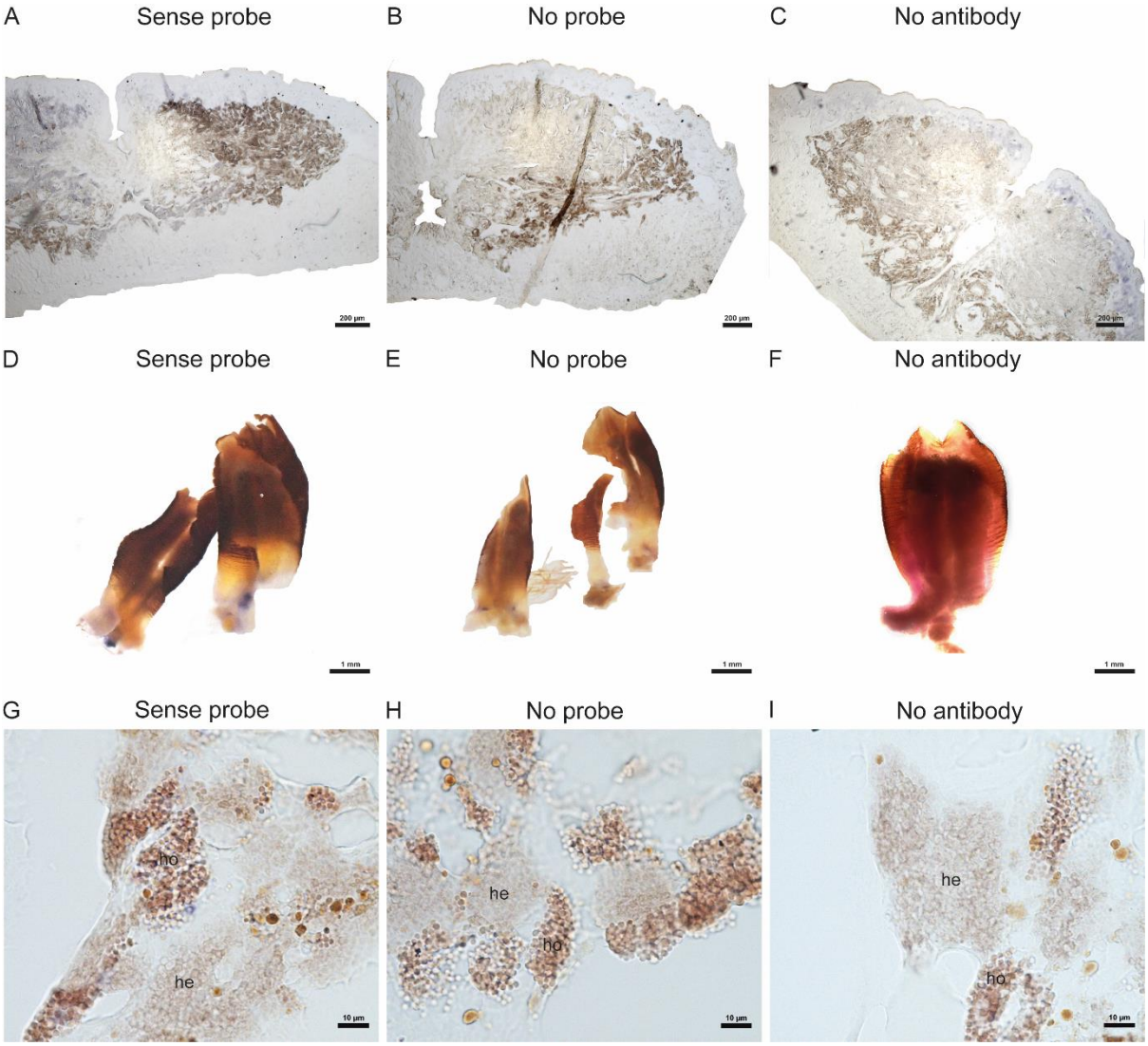
Supplementary Table 1. List of the tyrosinase protein sequences retrieved from the NCBI database and used for the CLANS analysis and phylogenetic analysis. The reference sequences used in the BLAST analyses are highlighted in bold.

These two supplementary materials are available on:

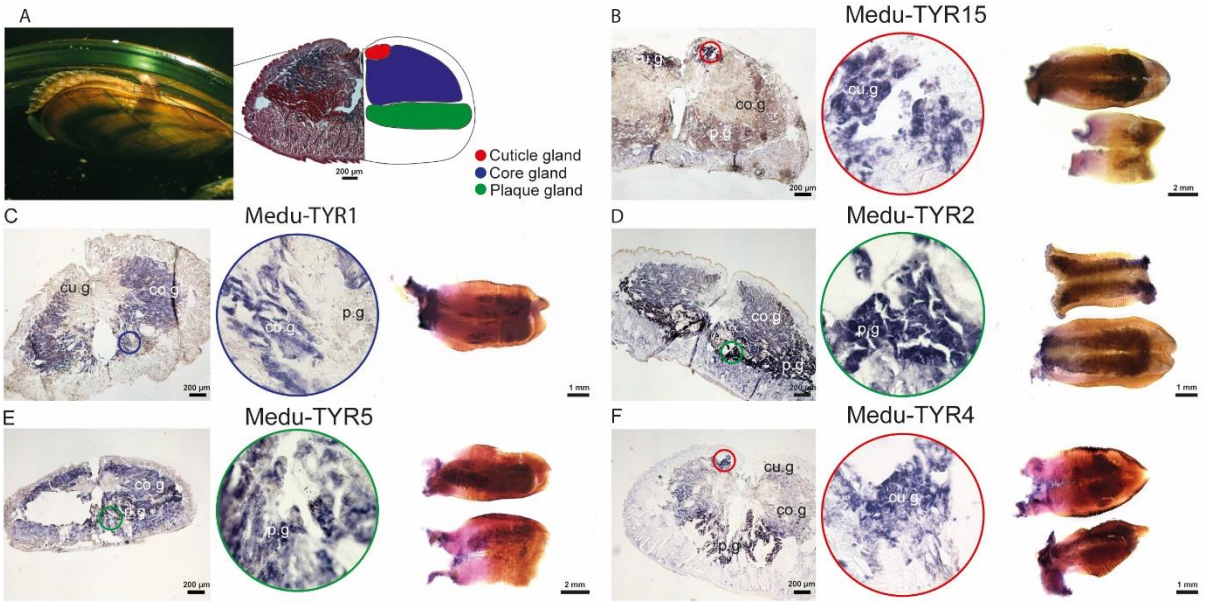
https://hdl.handle.net/20.500.12907/50241



Supplementary Figure 1. Distributions of predicted transcript length and BUSCO analyses of *S. alveolata* and *M. edulis* transcriptomes.



Supplementary Figure 3. *In situ* hybridization (ISH) controls with sense RNA probes, without probes and without antibodies. Section (A-C) and whole mount (D-F) ISH on the foot of the mussel *Mytilus edulis*. Section (G-I) ISH parathoracic cement glands of the tubeworm *Sabellaria alveolata*.



Supplementary Figure 4. Expression patterns of selected tyrosinase transcripts in the foot of *Mytilus edulis* (A) Illustration showing the mussel's foot and a transverse histological section through this organ stained with Heidenhain's azan stain (left), with a diagram with the arrangement of the different glands in the foot tissues (right). (B-F) Section and whole mount *in situ* hybridization of five tyrosinase candidates. Insets in circles show a zoom on the labelled gland. Abbreviations: cu.g – cuticle gland; co.g – core gland; p.g – plaque gland.

Supplementary Figure 5. Nucleotide and protein sequences of the tyrosinases involved in the adhesive systems of the mussel *Mytilus edulis* and the tubeworm *Sabellaria alveolata*.

>Medu-TYR1_comp83177_c0_seq2

GAACAATTCACCGTAAAATCTGGTGAAGGATCAAACTGTTACTCAGTTAAATGAGAAAAAATGTTTAAGTAGGGA
TTTTACGAATCTTTAATATAAAAACATGATTATAGTTAAGTGCTTTTACAAACACAGGTTGACCAGCGAAGGTCT
CAAAATGCAGTGTGCCTTCGCATGCTTGTTTATACTGTCTTACATCAGTATAGGAACATGCTTGATAGAAGAAAT
AGAACTACCGCAACATCTTAAAGAATGCATAATCATGAAAACTAAATTTAATGATGTTACTAAATCCACAAGTGA
AGCTGTATGTAATACATGTGTAACAAGATACATGTGGATGAATGGACCAAATTTACGAGACTGCGGTCATCCTCA
TGACAACACAACAATGTCAGACATTTCAGACTTCGTAGGGAAACTGCTTTGTCCTAAAATATCAGGCAAACGAAA
GAAGCGACAAGCAGGACGCAATTATCGTAAAGAGATCAGAATGTTGACCAAAGGAGAAAGACAACGACTTCGTAA
TGCATGGAGCCAGGCTTATAGAAGTGGGTATTTTGGTTGGATTGCCAGATTTCACAACTCTCAAGTCAGAACTTC
TGCTCACAGTGGACCTGCCTTTCTTGGTTATCACAGATTTCTGATCCTGGTATTAGAACTGGTATTACAGTATTA
CGACCCTTCCGTCACAATGCCATATTGGAACTCTGGTCTTGATGCAAATATGGATACTCCTGCCAATTCTATACT
CTGGACACCCGACTATCTTGGTGAAATGAGTGGAACTGTCCGAACTGGTATTTGTGGAGGCTTTCAGGACGTTAG
AGGCAATGCCGTAATCAGAAATGTTGGTAACGCAGGTTCTCTGTTCACGTCAGCTGAAGTTAGGGATGTAATGAG
GATGCCAAATATTAATCGTCTGACCGAACCTTCTCCATATCCCACCATGGAATCTTATCACGACGATGTTCACAA
TTATATAGGTGGATCAATGGGACCTATCGAATTGGCACCATGGGATTGTATTTTTTGGTTACATCATACTTTTAT
AGACTATTTGTGGGAAGTATGGAGATACAGTCATCAATATAATACAGCTTATCCTCATAAATTTGGACTTCCTGC
TCAGTTACCAGACGTTCCGATGAAGAACATGCCTGTTGTTCAATTTCTTGGACGTGTTCCAACAAATCGTGACGG
GTATGGTGTTCAACTGGCACGAAAAAGTTTTTATCAACTATCGCCTACGTGTTCAAGACAGAACACTTACTGTGG
GTCCCCTGATCTGCAATGTAGAATGGGAACCAACGGAGCAGAATGTGCTGCTGTAGATCGAAATTTTAGAGCAAG
AACGCAGCCGTTAGTTGATGCTAGAAGAAATGGACGAGTCCAGATGAATTTCCGTAACAATCATGGTAAGAGGAG
AAAAAGAGATACCCACGTACACCACACAACTATCAGTCACAGTTATCACATTACTGTTAACTCTCAACAAATAAC
AAATATCTATCTTCCACAGGAACGTGAACCATGCATGGGTAAACCTATACAAAATAACTTCATAGCGGGTTGCGC
AAGTGACGTACGAAAATGGGCATTTATACCAATTAAAGTTGTTCATCTGAGACCTCAGGAGGCTCACTTTGCATC
ACAATCTGACGCCAACGCCAAAAATCCAAGTGGGTATGATATGTACGACGAACACAAATACAAAGGTCTGAAAAA
TGTTGTTAAACCAGGTAATCCAGCAACATATAATGATTGTATAGAAGATGAATCTGGTGCATTCAAAGTAAGACT
GAAGTCGACTGGATTGAGTTACTTCGGAACATACACAGATTATGTATTTGTTGACAACAGACTTCCTATATCCTC
TCATATCGGGTACATTGCTGTTGAAAAACCTACTGCAAATAAACCAACAGAAGTAATGATCACTGCCTCTGATGC
CTGTGGACGACTCTGTAAACCAACGTGCCGAAAATGGGAGAATGGAAATCTGTACTATGTCTCTTGTACGGGAGC
AATAAAAGTCACAGCGCAGAGTCCACTTATGTATGGTAACGACTTTGGTGAAGCTGTTCTATCGATCTGGCAGTT
CTATGGCAAGTTCACACCAATAGAAAGAGAATCAAACATTTACATGGAATTTTTCTGTGATTATTCAAATGAATG
GATATGGAAAGAATGTTATCAGAAAAAAAACTAAAGAAATTATATACTTTTTTCTTCCAAGTTTTTTTTATTTAT
TCTACATGTATGTTACCGAAATAGTTTTGTTTTCTTAGCATCTGTTTGTGACATGAAGCAGAATAAATAAACGGT
GTTAAGTAATAAATTTCAATGATTATTTCACATTATATACCTTATAAAAGAGGCTCAGATGTTGTTGAATCGCTC
ACCTGGAAACAGTTTGCTTTTTTTTATATTTAAGAAACGCACGTCTGGCGTACTAAATTATAATCCTGGTACCTT
TGATAACTATTTACACCACTGGGTCGATGCCACTGCTGGTGGACGTTTCGTCCCCGAGGGTATCACCAGCCCAGT
AGTCAGCACTTCGGTGTTGACATGAATATCAATAATGTGGTCATTTTTATAAATTTCCTGTTTACAAAACTTTGA
ATTTTTCGAAAAACTAAGGATTTTCTTATCCCAGGCATAGATTACCTTAGCCGTATTTGGCACAACTTTTTGGAA
TTTTGGA

MIIVKCFYKHRLTSEGLKMQCAFACLFILSYISIGTCLIEEIELPQHLKECIIMKTKFNDVTKSTSEAVCNTCVT
RYMWMNGPNLRDCGHPHDNTTMSDISDFVGKLLCPKISGKRKKRQAGRNYRKEIRMLTKGERQRLRNAWSQAYRS
GYFGWIARFHNSQVRTSAHSGPAFLGYHRFLILVLELVLQYYDPSVTMPYWNSGLDANMDTPANSILWTPDYLGE
MSGTVRTGICGGFQDVRGNAVIRNVGNAGSLFTSAEVRDVMRMPNINRLTEPSPYPTMESYHDDVHNYIGGSMGP
IELAPWDCIFWLHHTFIDYLWEVWRYSHQYNTAYPHKFGLPAQLPDVPMKNMPVVQFLGRVPTNRDGYGVQLARK
SFYQLSPTCSRQNTYCGSPDLQCRMGTNGAECAAVDRNFRARTQPLVDARRNGRVQMNFRNNHGKRRKRDTHVHH
TTISHSYHITVNSQQITNIYLPQEREPCMGKPIQNNFIAGCASDVRKWAFIPIKVVHLRPQEAHFASQSDANAKN
PSGYDMYDEHKYKGLKNVVKPGNPATYNDCIEDESGAFKVRLKSTGLSYFGTYTDYVFVDNRLPISSHIGYIAVE
KPTANKPTEVMITASDACGRLCKPTCRKWENGNLYYVSCTGAIKVTAQSPLMYGNDFGEAVLSIWQFYGKFTPIE
RESNIYMEFFCDYSNEWIWKECYQKKN

>Medu-TYR2_comp79852_c1_seq2

TAATGCCAGAAGCCTTTAAGTAAGACACGCATATTTAAGTAATTTATTGTGTTCTGAATAACTCTTGATAAAGTA
ACATGTATGATTGAAGTTACATGTATGAGGAGAAAAACAACCCTCAGTTGTAATTATGCCTAATATATTGAGCTT
TGTACTGGTGTTCACTTTTATTAGTCTTGGGAGATGCCTCATAGAACAACTAGAACTTCCACCGTTTTTAAAAGA
ATGTATAGTCATGAAAACACAGCATCAGAATGTGAAAATATCACCAAGTGAATCGGTTTGCAATACTTGTTTAAC
ACGATATATGTGGATAAAAGGACCAAACTTAAAAAATTGTTATAGTGTTAAAGGTAATACATCTATGACAGATAT
TACAAATTTCTTCGGAAAACTTCTCTGCAAAGAAATTTCAGGAAAACGAACAAAGCGATATGCCAACAAAATCCG

ATATCGTAAAGAAATAAGAATGTTAACAAAGATGGAAAGGTTACGATTACGGAGAGCCTGGAACAACGCTACCGA
AAGTGGGTCATATGCATGGCTTGCTAAATTTCACAACGAGCAAATTAGAACTTCTGCTCATCACGGACCTGCTTT
TCTTGGGTACCATAGATTCCTGATACTGATTTTGGAATTGATTTTGCAGCATTTTGATCCTTTGGTGACGCTACC
GTATTGGAACTCAGGTTTGGATGCCAATATGGATAAATCGGAAAATTCTATACTCTGGACGGATGAATATCTTGG
AGAGTTTACTGGAGTAGTTAAAACAGGTCTTTGTGGTGGTGCTAAGGATAACAGTGGCAATCCAGTTATTAGGAA
TGGCGGAAATGCAGGATCATTATTTACTTTCGCTGAAGCAAGGGACGTGATGAATATGCCTGACGTTAACAGTCT
AGTCGAACCGTCTCCGTATAACACTATCGAATCTTATCATGACGATGTGCACAATTACGTAGGCGGATCAATGGC
ACCGATTGAATCAGCACCATGGGATTGTATTTTCTGGCTTCACCACGTCTATGTAGATTATTTATGGGAGGTATG
GAGGTATACTCACAAATACAATACTACCTATCCAACAAAAGCCGGAATGGAAGCTCAAAATCCTAATATCACAAT
AAAGAATATGCCTTTCATAAATTTTCTTGGACGTGTTCCCATGAATAGTGATGGATCTGGTGTACAGCTTGCAAG
ACAAAGCTACTACCATCTTTCACCAACTTGCTCAAAAGACCATACTGATTGTGGTTCGCCTGATTTGCAATGTAA
AATAAAGAAGTCGGGTACAGAATGCGCTTCTGTCGACAGATTTTTAAAAAGACAAACACAGCCTTTTGTTGATGC
CAAAAAACATAAGAAAGTTCAAATAGAATACCGTAATAAACACAAACAACATAAAAAACGCGGGAGAAGAGACAC
TCATTATCATTCTACGGCAATAGCTAATAGTTTTAGTATAACAATAAATGCGGACCAAATTACCAACGTGTACGT
TCCACAGGAACGTGAACCTTGTATGGGGAAGCCTATCCAAAACAATTTTATTGCAGATTGTTCAACTGATTCAAA
GCGTTGGGCATTTATACCTATCAAAGTTGTCCATCTTCGTCCAAGAGAAGTCCACTTTGCATCACAGACTAATGT
TAATAGACGACACAACCGATATGATATGTATGACGAACATAAATACAAATATTTGAGAAAACTCGTACATCCGGG
CAATCCTGCGAAATACCAAGATTGTGTTGAAGACCAATCAGGAGCATTTAAGGTCAGATTGAAGTCTTCTGGGTT
AAGTTACTTTGGTATCTATACAGATTATGTTTTTGTCGATAACAGACTTCCGGTTTCTTCTCATATTGGCTACAT
CGCTATTGAAAAACCAAAAGTGAATAAACCGACGGAAGTCCTAGTAACTGCATCCGACGCATGCGGACGACTGTG
TAAACCGACCTGCCGCAGACAAGAAGGGAGACATGTCTTTTATGTCCCTTGTTCTGGAGCCATCAAAGTAACCAA
ACAGAGTCCACATATGTATGGTAACGACTTTGGCGAGGCTGTATTATCAATTTGGCGATTCCAAGGCACATATAC
ACCTAGAGATAGAGAGTCAAATATTTATATAGAATTTTTTTGTGATTATTCAAATGAATGGATA

MPNILSFVLVFTFISLGRCLIEQLELPPFLKECIVMKTQHQNVKISPSESVCNTCLTRYMWIKGPNLKNCYSVKG
NTSMTDITNFFGKLLCKEISGKRTKRYANKIRYRKEIRMLTKMERLRLRRAWNNATESGSYAWLAKFHNEQIRTS
AHHGPAFLGYHRFLILILELILQHFDPLVTLPYWNSGLDANMDKSENSILWTDEYLGEFTGVVKTGLCGGAKDNS
GNPVIRNGGNAGSLFTFAEARDVMNMPDVNSLVEPSPYNTIESYHDDVHNYVGGSMAPIESAPWDCIFWLHHVYV
DYLWEVWRYTHKYNTTYPTKAGMEAQNPNITIKNMPFINFLGRVPMNSDGSGVQLARQSYYHLSPTCSKDHTDCG
SPDLQCKIKKSGTECASVDRFLKRQTQPFVDAKKHKKVQIEYRNKHKQHKKRGRRDTHYHSTAIANSFSITINAD
QITNVYVPQEREPCMGKPIQNNFIADCSTDSKRWAFIPIKVVHLRPREVHFASQTNVNRRHNRYDMYDEHKYKYL
RKLVHPGNPAKYQDCVEDQSGAFKVRLKSSGLSYFGIYTDYVFVDNRLPVSSHIGYIAIEKPKVNKPTEVLVTAS
DACGRLCKPTCRRQEGRHVFYVPCSGAIKVTKQSPHMYGNDFGEAVLSIWRFQGTYTPRDRESNIYIEFFCDYSN
EWI

>Medu-TYR3_comp74994_c0_seq1

ATCGAGATCACATTTCTTTAAATATCATTAGGCAACAACGAACATAATTCTTTGGATCGGAGGGGAGGTAAAAAG
TGGTGTTTTACCATGCTGTGTTTGCGTATTTGGCTTCTGCTGTCTTTGGTCAGTTTAGGAAGATGTTTGATAGAA
GAAATACAACTGCCGCCATCTTTAAAAGAATGTTTAATGATGAAGAGTAAATACCAAGACGTTACTAAATCACCC
AGTGAGTCCGTCTGTAATAATTGTATAACAAGATTTTTGTGGAAGGAAGGACCGAACTTGCAAAAATGTGACCAC
CCGAACGATAATATAACTGTATCGGACATTAGTAACTATTTTGGAAAACTTCTACATGATGAAATAAAAGGAAAG
CGAACACAACGTCAAGCAGGGAGTGTCCTTTATCGTAAAGAAATTAGAATGTTAACACGTGCCGAAAGACTCCGA
TTACAAAGAGCCTGGTCAAAAGCCTACAAAAGTGGATATTTTGGCTGGCTTGTTAGATTTCACAACGACAACAGC
AGAAATGTTGCACAAAGCGGACCTGCCTTTCCTGGTTACTACAGATTTTTATTACTGATATTAGAACGAATTTTA
CAGAGTTTTGATCCAGAAGTATCTCTCCCATATTGGGACACCACTGTCGAGGCTAATATGGATAAACCTGAAAAT
TCCATACTCTGGAATAATGAATACCTAGGGGAAATAAACGGATTCGTCAACACTGGTATTTGTGGGGGATTTAGG
GACCTTAGGGGTAACAAAATCATAAGAAATGGTGGGAATGATGGTGCATTATTTACACATACTAATTTAACGTAT
TTATTGAATAAACAAACCATTGACGATTTAGTTGAAGATTCTAGTTACGATACAATGGAGTCGCTTCATGGCAAT
GTTCACAATTATATAGGGGGTACAATGTCCTTACTTGAATTAGCACCATGGGATTGCGTTTTCTGGTTTCAACAT
GCTTATATAGATTATACTTGGGAATTATGGAGATATACCCATGGGTACAACATTGACTATCCATTTAAACCCAAC
AAAAGAATACAAGAGGCTGAAGTCCCAATGGAAAATATGCCTTATATAGATTTTCTTGGACATATTCCGACCAAC
AAGGACGGATATGGTCATCAGCTCGCATCTTTGACCCAATATCAAATGTCACCTACTTGTTCACAACATAACCTG
TATTGTGGATCACCTGATTTGCAATGCCGAAGAATTGATGGCTCCTTTAGATGTACTGCTGTCGATAGAATATTT
AGACACAGTGTAGCGCCATATTTAAAAGCTGCGCAAGGTGGAAAAGTCAAGCTTGATTTAGATGACCACCATGAA
GGAAAAAGGACAAAAAGAGACACGACTAATATCCATCTGACTACTATCAAAAACACTGTATCTATTGTAATTCAA
TCAACCGAAATAACAAATATCTACTTTCCACAAGAACATGAACCATGTATGGGAAAGCCAATTCAAAATAATTTT
GTAGCAGGTTGTACAAAAGCAGACGTGAAGAGATGGGCATTTATACCAATTAAAGT

MLCLRIWLLLSLVSLGRCLIEEIQLPPSLKECLMMKSKYQDVTKSPSESVCNNCITRFLWKEGPNLQKCDHPNDN
ITVSDISNYFGKLLHDEIKGKRTQRQAGSVLYRKEIRMLTRAERLRLQRAWSKAYKSGYFGWLVRFHNDNSRNVA
QSGPAFPGYYRFLLLILERILQSFDPEVSLPYWDTTVEANMDKPENSILWNNEYLGEINGFVNTGICGGFRDLRG

NKIIRNGGNDGALFTHTNLTYLLNKQTIDDLVEDSSYDTMESLHGNVHNYIGGTMSLLELAPWDCVFWFQHAYID
YTWELWRYTHGYNIDYPFKPNKRIQEAEVPMENMPYIDFLGHIPTNKDGYGHQLASLTQYQMSPTCSQHNLYCGS
PDLQCRRIDGSFRCTAVDRIFRHSVAPYLKAAQGGKVKLDLDDHHEGKRTKRDTTNIHLTTIKNTVSIVIQSTEI
TNIYFPQEHEPCMGKPIQNNFVAGCTKADVKRWAFIPIK

>Medu-TYR4_comp80529_c1_seq1

TTCCTATTGGAGAAATACAAATAAGAAAGACTTAAGTATTTTATACGGTATTAATCGGCTGTTAAACTTAAGTTT
GTTTTATTGCTCAAAATATTATCCCCATTCGCCCATTTGCAACTAGAATTATGTACGTAACAAAGACGCGCACGT
GAACTAACAACTAAAAGGACGTTACTTCGCTGTTTCAGACTTCATTTTGAACGGATTAGGTGACTCTTTCATGTT
GTTTCAAAATTTCCGAATAATGATGTAATTGGTTCATTTATGTGTTTAAACCTGTTGCTTTACAAGATTAATACC
GTGTTTGCATACATAGTCTCAATGGAAAATGAATATATATGAGAAATAAAAAAGACAACCTAGAAACATAGTTTT
CATTAAACACAGCAATATAGTTTTGTACTTTTATATCAGTGTAATTAGGTTGAAGTTTTTAAAATCTTTATTCTA
ATAAATAGTTCACGTAACTCAAGGATTAGTATTATCAGAAGTGATCTTAACTAAAAAGCACACCTACAAAAACCC
ATAAAGGATACGACTAATCTGGATTACCCATTATTGTAAACATGAAACGCATCACCATAGTTTTGGTTTTTGGAG
TTGGGTGTATTCTATCACTAGTTAGAGGTATGATTGATCCCGGAGACTTACCAGATGGTCTGAAAAAATGTTTCG
AAATGAAGACTAACAATAGTAACGTCAAAGAGGTTCCGAGTGAAGACATCTGTATAACGTGTATAACAGATTATC
TATGGAAACAGGGAAAACACGTTCATTGCTTCAATAAAACAGAACATGATGTAAAATGGGTAGCGGAATTATTGG
CAGAACTTCATCCCGCTTCAATGAGAACTAAACGAAGTGCACCGAGACGATGTCGAAGGAAAGAATACAGAATGT
TGACAGACGATGAAAGAGAAGATTACCATCGTGCTGTAAATCTAATGAAAAAAGATAAGAGTATATGGCCGAATC
GCTACGATGGCTTTGTGATCTTCCACAGCGGAATGTCTGGTAAATCAGCTCACAGGGGACCTAACTTTTATGGAT
GGCATAGAATTTTATTGATCATGTACGAATATTCGTTACGAAAACTTGTACCGGGTGTTTGTCTACCATATTGGG
ACTCTACTTTAGATTCTGCCATGTCAACAATGCCTGAAAATTCTATCATATGGAGCAAGCATTTTATGGGAAATG
GAGATGGACAAGTACAAACTGGTCCATTCAGACGTTGGATGAATATAGATGGCAGTCCACTTATGAGAAATATTG
GAAGTACTGGTCAATTATTGACAAAACGTACTGTCTTAAATATCCTTTCAAAGCGTTATAATCACCAAATATGTG
AGCCTTCGCAGAATGTGATGCATAGTTTGGAAGGGTTACACGATCAGGTTCACATGTGGATCGGTGGTAAAATGT
TTGATATCGCAACGTCACCAATGGATCCAATTTTCTTTATGCACCATGCCTTTGTTGACTGTCTGTGGGAAATAT
TCAGGAAACAACAAATTGGATATAAGATAGATCCCAGTAAGGACTATCCCAATACCGGAGACCTCTTACAACACC
CAGATAGACCAATGGACAGTCTTCCCCCGATACCAGGATTTGGAAGACTAAGGAATAAGGATGGTTATCTCTCTT
ATTGGACGGATAAATTTTACTTTTATGACCCTCCTCCAACATGCAGTTTCCCAACTTTAGACTGTGGATCAAAGT
GGCTAGAATGTGACACAAGAAGGAAAATTTGTGTTTCCAAAAGTAACCAAGCAATGTCTTTTCAATATCCAACTT
TTAACTCATTTAATCCCTGGAATTTCTACTTAAAAAGAGGTAAACGAGACACATATAAGGTTTTGTTAGGTAAGC
TCGGCGAAACTGAATCGCCTGATCAAAGCCAAATGAGTTATGAGAAGGTTAAGCTAAACGTTCCAAAGGAGGTTG
CTTTAGGATCACCAATACAAAATGCATTTGACATAGATTGCGTTCAAGATATAAGAGCATGGGCATTTATGGCTA
TCAAAGTAATCCACTTAAGACCTCAAGAAGTTAAAATGTCAGCGCATATCATTGAAAATGGCAAAATACAAGATA
CAGACATGTATAACGAGAGTAACTACAGAAGCTTAAAAACACATATAAGACCCGGTTCTCCAGCCATGTATGATA
CTTGTATGGAAGACAATTCTGGCGCTTTCAGGATCAATCTCGTTTCACGTGGACTTAGTTACAATGGTTGGTATG
AAGACTATGTTATGGTAGACAATCGCTTACCCATCTCGTCACATGTCGGGTATATTGCATTCAAAAAGCCAGGTC
CAAACGATACAGAACCAGTTACATCATTTATTTCAGCAACAGATACATGTGGTCGGATGTGCAAACCCAAGTGTC
TGAAAAAGGGATCAGGTTCAAAGCCGACATATGTTCCATGTGATGGAGTAATTAAGGCTACAGCACAATCTCCTT
TGATGTATGGAGAGACGTATGGGGATGCTGCAGGTGACATTTGGGACTTCTCTGATCCAATGGTTCCTGTAAGCA
AGGAGAGTGGTATCTTTATGATATTCTATTGTGATTATTCGGATACGTGGCCTTGGGAGAGTTGCAATAAAAAAT
AATTCAATTTGCAATCCTTC

MKRITIVLVFGVGCILSLVRGMIDPGDLPDGLKKCFEMKTNNSNVKEVPSEDICITCITDYLWKQGKHVHCFNKT
EHDVKWVAELLAELHPASMRTKRSAPRRCRRKEYRMLTDDEREDYHRAVNLMKKDKSIWPNRYDGFVIFHSGMSG
KSAHRGPNFYGWHRILLIMYEYSLRKLVPGVCLPYWDSTLDSAMSTMPENSIIWSKHFMGNGDGQVQTGPFRRWM
NIDGSPLMRNIGSTGQLLTKRTVLNILSKRYNHQICEPSQNVMHSLEGLHDQVHMWIGGKMFDIATSPMDPIFFM
HHAFVDCLWEIFRKQQIGYKIDPSKDYPNTGDLLQHPDRPMDSLPPIPGFGRLRNKDGYLSYWTDKFYFYDPPPT
CSFPTLDCGSKWLECDTRRKICVSKSNQAMSFQYPTFNSFNPWNFYLKRGKRDTYKVLLGKLGETESPDQSQMSY
EKVKLNVPKEVALGSPIQNAFDIDCVQDIRAWAFMAIKVIHLRPQEVKMSAHIIENGKIQDTDMYNESNYRSLKT
HIRPGSPAMYDTCMEDNSGAFRINLVSRGLSYNGWYEDYVMVDNRLPISSHVGYIAFKKPGPNDTEPVTSFISAT
DTCGRMCKPKCLKKGSGSKPTYVPCDGVIKATAQSPLMYGETYGDAAGDIWDFSDPMVPVSKESGIFMIFYCDYS
DTWPWESCNKK

>Medu-TYR5_comp87104_c0_seq11

CAGTGTAATACTTACGCCTGATAGATATAAATTGAACATTTTTCTTATTTCTTTCTTATACTACGGCAGATGAGC
TATGAATATAAAGATATATTTCATTCTGTTCTCACTATTGGTTTGTGTTATATGTGATATTAAGCGAACACCAAT
CCCAACAAGTCTAAGACGATGTTTACATACAAGAGATCGATACTCGTCTATGACAGATGAACAAATAAGCACGAC
TTGCTTGTTAGAATTTCTGTGGTTGAATCGTCAAATTTGTGACACTATATCTAAACAAACATATGATTGGCTTGT
GTCACTTGTTGAAAAATCACAGCAGCGATATAAGTTGACACCTTTACATACAAGTGTTAATTCAATTACAAGATC
TGAACCACAAACGACAATTATGTCACAACCAGGTACTTCAACACCTACACCACTTAAAAGATATTTAAGACACAG

Chapter I

GAAAAAAAGACAAATATTAACGACCACATCTCTACCGCATTTATCAACCAAACCAAGATTATCAACCACAACCAC
TCAACCACAATTTTCAACAACAACGTCACGACAACAAACAACAAGATCGCAACTAAGACAGCGTAAAGAATACAG
AATGCTTACTGACCAGGAAAGGGATGATTATCATCAGGCCATTAACGACTTAAAGCAAGACACTAGCGTATCGCC
AAATAAATATGATGCACTCGCTGAACATCACCAGAAAACATCAGAGACGGCACATGGTGGAGTAGCTTTTGGCGG
TTGGCACAGATATTACCTTCTTCTGTATGAACTCGCATTGCAAGAAAAGAACCCGAATGTCATGTTACCATACTG
GGATTCCACATTAGACAGTGCGATGGACACTCCTACAGATACAGTTATTTTTACAGATAAGTTTCTTGGAAATCC
AAATGGAAATGTAACATCAGGACCATTTGGACATTGGGAACGAAAAATACAGCGATCCCTTCACAAGGATACATC
GTTGTTAATCACAAAAAATGAATTAAGTTTGTTAATCCATAAGAGTAAGGACCAACGAGAATTCATGAGTACTTT
AGAAATGCACCATAACGGCCCACATGTATGGATAGGTGGTACTATGGTAGATGTAGACAGTGCTCCTTTTGATCC
CGTTTTTTACATGCATCATGCATTTATTGATTGTATATGGGAGCGTTATAGAGTTCTTGAGAAACAAAGGGGCCA
AAACCCTGAAATCTACCCAAGCTTAAAATCTGATAGTCCAGCGCATCAACCCTCTACTGTATTGATAAATCTACC
ACCTTTCAATGGTGTTAATTTTACGAATGCTGATGGATATAAAAACTTTTGGACAGAAACTTATTATTCTTGTGC
ACCTCAACCATCGTGTAGCGTCCATTCACCAGAATGTGGATCAAAGTGGCTTGAATGCGACATATTTACAGAAAG
ATGTGTTTCTAAAGGATCAGAAGCTGCTTTACCTCGTGTTACCCCTTCAGATGTGTTCCGAGTAAAGAACGCTAG
AGATCTAACATCTCCAACCATTCACAGTTTAACGACAACTCCACCAGCCACACAGTTTGGTCCTATATCAAGTCA
GCCAACAAGCAAAAATAAAAAGATCACAGCGGAGGACTTTTTCTCCACACTAAAATCATTGAATAAGCTACCTAC
CAATGTTCCAAAAACTCCAAGTCCTCGAATCAGACGTAGTATGATGAAAGGGAAAACAGGAAAAAACGGTGGCAC
AGGAATGCTTGGAAAACTTAGTTCTCCCTCTATCCTAGTTACTCAACAGCAACAACATAGAATGCTTTCTATTGA
TGGATTGAAAATACCAAAATTTAAGAAACCATGTTGTGGACATCCAATACAGAACACTTTTAGAATAGATTGCAA
AGAAGGAACAGATTTATGGGTATTTGTTCCAATCAAAGTTGTCAATCTACGACATGGTGATATAGTGTATCCGTC
ATATCCAGTTAATTCATATGGTCGTAGATTGAATAGTTCCGATGTGTTTTCGGTTGGACAACAATTATCATTTGC
ATCTAATTTACAATCCAAAGCACGTTCTAAACATAAGCCATGTAGCTACGACAAATCTGGAATGTTAAAAGTTAA
AATATCCTCTTACGGCTTGAATTACTATGGTATTTACGAAGATACAGTCTTTTTGGACAATCGACGAACAGTTTC
AACTGAAATTTCATATATTGGCATAAGAAATCCAAAACAAAGGGAAACAAAGGTCTTCGTCACTGCTGTTGATGA
ATGTGGAATTGCTTGTCAACCGATGATTTTAGCACGGGGTTATCGAACTCATCATTTGCGATATAGAAAAAATAT
AGGTGCTATTGAAATAACAACAAAGGGACATCAGTATCACAGTAATACATATAGAGGTGCTGAGTCATTGGTTTG
GAAAGCAGATAAATTTTCTGTTCCTCAACCAGATGAAAGCCAGATTCCGATTGTATTTGTTTGTGATTATAGAAG
TAAAAATCCATGGTCTTGAAAAAAGATAGTTAAGTGGGTATGGAACTTTTTTTCACTTTCCTTTAAAGACCTTTC
CAAAAATTATAATTATTTTAGTGAGCACGTTTGGGATGTTCTATGTACAACAATTTTTTATTATTAATGTTCAGA
AGTCAGCAGCTCTGTAATGACATGAAATATATATTCAGTGCTGTCATTTTCTGTGTTTACTCTTCATTAACTATT
ATATCTTTTGAAACTTGTTAATTTCAACTATATTCAGTCTCTGGAGGATATTTGTTTCATACCACGTCATACGTT
TCATCTTAATTGAGACAAATTAGTGCTGACAGAAAACGAAAACAACACTTTAAAAGTTAAAGCGTCCGAAGAGCA
TTTCCCAATTTTTTTCAGAACGCTTAAAAACGAATATATTAAGGCCAAACATGTATAAAAATAAAAACAGTTTGC
AGAACTTATGACTAAAAAAATACATAAAAGTAGCTAAATCAATTCAAAGGAGTTTTTTCAAAAGAAAATTGAAAC
CTTATTTTCTTCAAATTCATAATTTTATAAACGGACAACGTCAAAGTTTGTTATAAAATAACTACTTCAAGATTT
CACAAAGAAAGCTATAGAAATGCCAAGTAATATAATATTGTAAAAATTCAGTCCAAGAAGTCGAAAACAACGCTA
TTTGGTGCCTCAGAAACCTATGTAAATCTATCAGTGTGATATTGTAAACGATTTACACAATCTTGGAGGAGCATT
TACACTATAGATGCGAATACCATGAGCTGATCTCGAATGTATACCTAGGAAAATACAGACATTTGAATAACCTCT
AACAGAGACAAAGATTTTATGAATATAAAATTCCAAATCTTGATTGTCATTTAACACTGAAACAATAGTTTACTT
TATCATCCGGAGCTCTTGATATCCTCCCGGTCTTTGGTGGGGTCGTGTTTCACAGTCTTTGGTCCTCTATTTTGC
CGTTTGCATACTCTTGTTTGTCACTTAGTATTTTCTTTTGTTTGCTATTGCTTAGTAAGTGTTTTTAGACTCTTT
TTTGGAGGGGGAGAAGGATATCTTTTTTCCTTTTATTATTAATCATAGTAAACATGTATGAACAAAATTGACAAT
GGAATTTGGGGATGTGTTAAAGAGACAAAAACCTGATCAAAGAGAAAAAAAAACAGATTTTTTTCGAGACAAATA
GGGTAAAGCATCATATAAAAGCACTATATTTTTTATCATTTTGACTGAACAGATACCAACTAACAAGAAGTATTC
TCTTTATTTGTTACATTCTAAGTTAGACTTATCTAACAAAAAAAAAACAACAAATGCGTTTGAATAAATAAAGGC
AACTGTACTATACCGGTGTTCAAAAGTCATAAATCGACT

MNIKIYFILFSLLVCVICDIKRTPIPTSLRRCLHTRDRYSSMTDEQISTTCLLEFLWLNRQICDTISKQTYDWLV
SLVEKSQQRYKLTPLHTSVNSITRSEPQTTIMSQPGTSTPTPLKRYLRHRKKRQILTTTSLPHLSTKPRLSTTTT
QPQFSTTTSRQQTTRSQLRQRKEYRMLTDQERDDYHQAINDLKQDTSVSPNKYDALAEHHQKTSETAHGGVAFGG
WHRYYLLLYELALQEKNPNVMLPYWDSTLDSAMDTPTDTVIFTDKFLGNPNGNVTSGPFGHWERKIQRSLHKDTS
LLITKNELSLLIHKSKDQREFMSTLEMHHNGPHVWIGGTMVDVDSAPFDPVFYMHHAFIDCIWERYRVLEKQRGQ
NPEIYPSLKSDSPAHQPSTVLINLPPFNGVNFTNADGYKNFWTETYYSCAPQPSCSVHSPECGSKWLECDIFTER
CVSKGSEAALPRVTPSDVFRVKNARDLTSPTIHSLTTTPPATQFGPISSQPTSKNKKITAEDFFSTLKSLNKLPT
NVPKTPSPRIRRSMMKGKTGKNGGTGMLGKLSSPSILVTQQQQHRMLSIDGLKIPKFKKPCCGHPIQNTFRIDCK
EGTDLWVFVPIKVVNLRHGDIVYPSYPVNSYGRRLNSSDVFSVGQQLSFASNLQSKARSKHKPCSYDKSGMLKVK
ISSYGLNYYGIYEDTVFLDNRRTVSTEISYIGIRNPKQRETKVFVTAVDECGIACQPMILARGYRTHHLRYRKNI
GAIEITTKGHQYHSNTYRGAESLVWKADKFSVPQPDESQIPIVFVCDYRSKNPWS

Chapter I

>Medu-TYR10_comp85123_c3_seq1

CTAAAAGCCGGTGAAATTTCATTTTTACTTCATCAATCCCACTCTCAACTGTGTATTTAGTGAAAACAGTTTTAA
ATTTATATTTATTGTTGAACTATTATAGAACGAAGTGTTTTGTTTTTTTGGATCTCAAGATTCAAAAAAGTATAC
AACGACGAATGGTTCAACATAACATCTAGTTCGAAATGAAAGCCGTCTTGTACTTAATCCCTTTCGTAAGTTCTG
TGTTTAGCATGATGACAGAAAATGATATGCCAGAAATACTGTGTAAATGTTTAAGTGAAAAGGCGGGTTCGATAA
CCATAATTCCAGCCGACCTAATTTTATACACATGTGTAAATAAGTATCTTGCAGAAACCTGGTTGGAACGCTATC
AACCGCCAAGACTGATTGACCCTGACACTATAAACTGGTTACGATCACTTGGTCGTAAGGTCAGAGGACAAGAAC
ACAACATTCGAACAAAAAGACAAGCAGTCAGAAGAGTAAGACGGGAAATCAGAACACTTACAGATACACAAAGGA
GAAGATTTATTCGAGCTGTTCAAAAGTTAAAAGATAAAAAAATTGGATCTACCAACAGATATGATGCTTTAGCTA
GTGCACACGATGGAAAGGCTCTTGGGTCGGCGCATTCTGGACCAAACTTTCCATCATGGCATCGGGAATACCTCA
AAATATTTGAAGCTGCTCTACAGGAAATTGATAATCGAGTTACCTTGCCATACTTTGATAGTAGATTGGATTATA
ACTTGCGTGACCCAAAAGAATCAAATTTTTGGGGTGATAATTTCATGGGCGATCACCAAGGTGTGGTTTCAAGTG
GACCATTTAGACTTTGGCGTCAGCGTAACGGAGCATATTTAGAAAGAAACGGTGGAGCATCTGGTTCAATGATTC
CACCAAGGGGTATAAGACATGTACTATCCCAAAGAAGAAATATTGATATATTGTCACCTTGGGCAGCAAATGGCA
GATATGGTCTTGAAAACTACCACAATGCTGTGCACGTTTGGGTTGGTGGTTCAATGCTTGGTGTAAATACCGCTC
CATCAGACCCAGTATTCTTTCTGCATCACTGTTTTATAGACTATGTTTGGGAGCGATTTCGAGAAAGACAAACTG
TCGATCCTGAGAAGGACTATCCATTTGATAGTCGTGCTCCACCATCCCAGGCACCGAACCGAATTATGGACAACC
TTGAAACTGGGAAGAGGAATATTGAGGGGTATAGTAATGATTACACAGATCAATATTACAGATATGCGAAAAGTC
CAAGAACATGTCGTCAATGTAGAAACAAATATCCCTATCTTAAATGTGGAGATTGGTTCAGAGGACTCTGTTCTG
CTACAACCAGACGGACTCCACCAGGAAAACCAGGACCACCAGGCGCACCAGGGATACCAACGCGTCGTTCTAACA
ATAATCTAGGGATTACAAAAAAATTTTCAACTACAATTAATGACTCCTCAAGAAACAAAACATCACGTAAAAGAA
GGTCAACAGATGATTGGCTTGTCTACGAACCTTTCATAGAGTTAACCAAAGTCAGTAAAATTGGACGAACGAAAA
GATCGATGAGTTCAAAGAGATACATTCAGAAAAAACTAACAACTCACATAACAGACATGGACTCTTCTATGCAAA
ACTCCTTTTTACTGGATGGCGAGCCTGATATTAAGCGTTGGGTATTTATACCGGTTAAAATTATTCATAGACGTC
CATCTGGCTTACTCTTTGGCACAAAAGTGATAAGGAATAAAAATATTGTAGATGGAGTTGATGTATATTCATATA
ACAAATATATGAACTATACAGATGAAGAGCAGCAAGCATCATATTCTCATAGTTTTACGTCTGGGTCAGGGGCAG
ACAAAGTATATGTTCAGGTTGATGGTATGTCATACAATGGAAAGTACTTAGATTACACTATAGTTGATTCCAGAC
TCCCTGTATCCGAATCTGTTGCTTTCGTTGCTGTTAAAAATCCGAAACTTAATTCTAGCATTTCTTATATATCTG
CTTACGACTCACATGGTCGGATTTGTCGTCCATATTGTCATTCATATGGATCTTATTCAAATCAGTATAAAACAT
GTTCCGGGGTGATAAATTTGACTGAAGAAAAACCTAAGCTGTACGGTGAAGATGTCGGAGAGGTAACTAGACTAC
TGTATAAGTCTGGAAAAGGGAAAGGTTTTTCCAGAAGTGATAAAAATATATTTCTCAGTTTCTACTGCGACTTCC
ATGAAATTCCATGGGGAAAATGCTAATTAATAAATTTGTGACGACGCTTGGCTTACATTGCATATATATGTACAC
ATATACTTTATTCTTGCACTTGAGAAATAGTGACGAAAACCCATGCATATGTAAAAATTTGGATCTAAAAGGTGA
AAACACCAAAAGTTTTTAAAAAAATAAATGAAATAGTCAAAAAGATTTCGGATAACACAAATCATTTTAAAGATA
TTGGAACAGCAAATCACTTCATTTCAATCAAGGAGATGTGTAAAAGTACACATTCCCTCACAAAAATAATGCCTT
TTATGTCATAACAAACAGATTTGCAGAAATGAAACTTTAAGTGAAGGAGAAATATATGTAAACCATTTTAGTCTA
ATTTTAAATGTCCATGAATTTGTATTAAAAATTGAGGATTAATGCACA

MKAVLYLIPFVSSVFSMMTENDMPEILCKCLSEKAGSITIIPADLILYTCVNKYLAETWLERYQPPRLIDPDTIN
WLRSLGRKVRGQEHNIRTKRQAVRRVRREIRTLTDTQRRRFIRAVQKLKDKKIGSTNRYDALASAHDGKALGSAH
SGPNFPSWHREYLKIFEAALQEIDNRVTLPYFDSRLDYNLRDPKESNFWGDNFMGDHQGVVSSGPFRLWRQRNGA
YLERNGGASGSMIPPRGIRHVLSQRRNIDILSPWAANGRYGLENYHNAVHVWVGGSMLGVNTAPSDPVFFLHHCF
IDYVWERFRERQTVDPEKDYPFDSRAPPSQAPNRIMDNLETGKRNIEGYSNDYTDQYYRYAKSPRTCRQCRNKYP
YLKCGDWFRGLCSATTRRTPPGKPGPPGAPGIPTRRSNNNLGITKKFSTTINDSSRNKTSRKRRSTDDWLVYEPF
IELTKVSKIGRTKRSMSSKRYIQKKLTTHITDMDSSMQNSFLLDGEPDIKRWVFIPVKIIHRRPSGLLFGTKVIR
NKNIVDGVDVYSYNKYMNYTDEEQQASYSHSFTSGSGADKVYVQVDGMSYNGKYLDYTIVDSRLPVSESVAFVAV
KNPKLNSSISYISAYDSHGRICRPYCHSYGSYSNQYKTCSGVINLTEEKPKLYGEDVGEVTRLLYKSGKGKGFSR
SDKNIFLSFYCDFHEIPWGKC

>Medu-TYR11_comp76132_c0_seq1

CTTACTTACTAATCGATGACGCCACATGGAACCGACCAATCATCTGGGGGCATACACATGTTTCATTGTGCTGTC
GTTCACGTACGCAAAATCCTCTTATAAACCTGTAAATTCGTTTCATTCTTATGAAATTATTTGAATGTGCATTGG
AGGAGTACAATATCCAGGGGACAAATTAAACACAAAGTGAATACATTTTTTCTCATCATAGTTTCCACAATGAAA
TGGGTCCTCGAGATCCTCGTCTTTTTGCCTTTGATCACAATATCAAACTGTAATATTGAGGAACAGCCGATGCCA
AGTATACTATGTGACTGCTTCTCTCGAAACAAGTGTGACATCATTAAGGAAAAGGCAGACGTTATTCACTTCAAA
TGCATCAATTATTACCTTGCTCATACTTACCAAGACAGATGGTACAAGAATATTTCAGGGGATGCCTTGAACTAT
ATATTGTCGCTCCAAAGAGAGGCTAATCAAGCCATTGAATCCGTCGGCAGAAAAAAGCGACAGGCAAATGGCAAT
TTGTTCCATGGTATACGTAAAGAGTTGCGCACGCTCAGTCGTGAGGAGAGGACTCGTTTTTACGCAGCAGTAAAC
AGTCTAAAAAACAACAGAATTGGTAATACAAACTCATACGAAGCGCTCGCAAGTATTCACAACACCAATGCTCTA
AATGCAGCTCATTTTGGTGTTGCCTTTCCAGGTTGGCATAGATATTTTCTATTCTTATTTGAACAAGCACTACGA

CGTTTTGATCGAACTGTTACCCTGCCCTATATCGACACCACAATGGAAAACAGTTTACCGAATCCATGGATGTCC
AATTTATGGTCTGCTGAGGGTATTGGTAATATTGACGGGGGCCAGGTTAGGGTGGGACCATTTGCGAACTGGAGA
TATCAAACACAGGACCGACGGTATGTCCCACTCACAAGAAATGGCGGGTCAGGAAGAGACTACTTTCCGGCAGAT
TGTTACAGAAACATAATACGTGAAAGCAGAAACTCTAGAATATTGGAACCATTGGCTGGTACTAATCGTAACTTA
GAGGCGTGTCATAATTATGCACACGCCACTATTGGTGGCACGATGAACGATATTGATGTTTCTCCAAATGACCCG
GTGTTTTACTTGCATCACTGTTATGTAGACAAAGTCTGGCAAGATTTCAGAGACAATGCAAAATCTGTTATAGGT
AATGACTTCGAATTTGATTATCCAGTTACTGCTGATATATTCCATCAGCCGAACCGAGCAATGGGTAATTTAAAT
GAATTCTCAAATAGCATGGGATATCTTAAAATTTTCGATGAAAGAATAACCAATTATGCGCCATCTCCTGGAGAA
ATCACGTGCACACGTAACTCTGACTGTCAAAGTCCTGAATTTTTGAGATGCTCATCTGGACGATGTATACCAATT
TTGCGTAGTTCGTCATCACCCTTTGGGAGAAGAAAAAAACGTGACGTAGAATACGAAAATGACGAGTATAAATCT
GTTATGGACAGTTCATATCAGAACAATTTTGTGATTGATGGAGAAGCTGATGTAAGTCAATGGGCCTTCATTCCA
GTCACGTTGATGTATATCAGACCGATAGGACAACATTTCGGTTGCAATCCTGTACGAAATGGTAGTATAGATGAC
AACGATGATATTTATTCAAACGACAAAACCTCAAAATTATGGGATTATTTTAAACCAGGAGTAGATAAAAGACGG
GTTTCCGATTCTCAATCTGGAGCAACCAAGTTTTTCGTACAATCCGATGGTATATCGTACAAAGGTCGATACATC
GATTATGGTATCATCGATACACGGCAGATGGTATACGAAACAGTCGCTTATGTTGGTGTAAAAAATCCTCGGAGT
GGATCTGCAACATCTTATGTGAGTGTTTATGATCTTCACGGCCATGTCTGTCAGCCTCGGTGTATAGACAGATCG
TCCAAAACACTTTTTTATAAAAAGTGTTCCGGTGTTATCAAAGTGACGAGAGAAGAACCCCAGATGTATGGTGAC
GACATTGCCGAAGCAGTCAGGTACAGATATACGTACAACTTTGACGGATACAAGCCTACAAGTAATTATCGAGAT
ATTTTTCTTCAATTCGTTTGTGAGTATGGAGCAGAGTATCCATGGAAAGACTGTAGTAGCGTTTAAATTTATGTA
AATTGTAAAAGAGAAATATAAATAAATGAAAGTAACAGCAGAAAAAAAAAAAAA

MKWVLEILVFLPLITISNCNIEEQPMPSILCDCFSRNKCDIIKEKADVIHFKCINYYLAHTYQDRWYKNISGDAL
NYILSLQREANQAIESVGRKKRQANGNLFHGIRKELRTLSREERTRFYAAVNSLKNNRIGNTNSYEALASIHNTN
ALNAAHFGVAFPGWHRYFLFLFEQALRRFDRTVTLPYIDTTMENSLPNPWMSNLWSAEGIGNIDGGQVRVGPFAN
WRYQTQDRRYVPLTRNGGSGRDYFPADCYRNIIRESRNSRILEPLAGTNRNLEACHNYAHATIGGTMNDIDVSPN
DPVFYLHHCYVDKVWQDFRDNAKSVIGNDFEFDYPVTADIFHQPNRAMGNLNEFSNSMGYLKIFDERITNYAPSP
GEITCTRNSDCQSPEFLRCSSGRCIPILRSSSSPFGRRKKRDVEYENDEYKSVMDSSYQNNFVIDGEADVSQWAF
IPVTLMYIRPIGQHFGCNPVRNGSIDDNDDIYSNDKTSKLWDYFKPGVDKRRVSDSQSGATKFFVQSDGISYKGR
YIDYGIIDTRQMVYETVAYVGVKNPRSGSATSYVSVYDLHGHVCQPRCIDRSSKTLFYKKCSGVIKVTREEPQMY
GDDIAEAVRYRYTYNFDGYKPTSNYRDIFLQFVCEYGAEYPWKDCSSV

>Medu-TYR12_comp83122_c1_seq1

CTTTCCTCTGATCAATTAGAATCTTAGTTAACTCACTTCTTAAATCCTGGCAGGAGAAAGATGAATTCGTACTGT
GCTCTTCTCCTAATTATAATAAACGGAGTCGAAGCTCTCATCTACAAGGACTGGTTCCCTGGTCCAATGGAGAAA
TGCTTAATGGATCGTAGCCGTGGAGTTTCACCGAGAAGGATACCAGCGTTTGATATCCTGTTCGAGTGTAAAAAT
TACCAGGTTGCTTATAATAATGTGGATAATGATGTGGTCAGTCCGGTTACCGAAGACAACGAACGATACTTTAAA
CATCTTGGTCGTCGTCTGCAGGGACTTGAGTCCGAATACAAACGGAGGAAAAGAAGTGCTAAATGGAAGTGGAAT
AATCAAAGAAAGGAAATTCGGACGTTGACAGACAAACAACGTTCTGATTATTTTGCTGCTCTTAATGCATTAAAG
AAAGATGGATCCTACGATGCCATTACAAGACTTCATCAACAAACAGCCATAATGGGTGCACACTTTGGACCTGGG
TTTCTAGGATGGCATAGAATATATAATTTAGTGTTACAGCTAGCAATCTGGGATAAGAACCCACGAGTCATGTTA
CCATACTGTGACACAACATTGGACCACAACATGATAGACCCTAAAAAAAGCATTTTATTTAGTAAAAGTTATTTT
GGAAATCTTCGTGGAATTGTTGTTCATGGGCCGTTCGTAGGTTGGAATACGACGGCAAACGTACTATTACAACGA
GATGGCGGACGTTTTGGATCGTATATGACAGATGAAGGATATAAACAAACTATGTCTCTTAAACATATGCGAGAG
GTCACTTTAGGAACATCTCTGGGTCAAATTGAGAGACAACACAACAACCCTCACGTGTATATTGGGGGTCTTATG
AGTGATCTCAATTATGCACCATCTGACCCATTTTTCTTTTGTCATCATAATTACATTGATTATGTTTGGGAAAAG
TTTAGAGAAAATCAACGGAGAAAAAATATCGACCCAGGTAGTGATTACCCTAGCAATACTATGAATTTACATGAA
CCAGGCAGGCGCCTTCCTATGATAGCAGATTATTTGAAAGTTAATCTTACACAAGCACAATGCAGTCATAATATG
TTCTCCGATAAATTTTTAATAAATTTTGCAAATTCACCAAAATATCCAAATTGTGGAAACTATCCACATATTAAG
AGAAATGATACGAGGAAGGTATGTTACTCTGCTGAATATCCTTCGTTCGAAGATGATAATGTGGTAGCTCTTCAA
ACACGTGTTCGTTCAAACAAAGATGCAACGAAATTTAAAGTAACAAAAGGTTTTGTAGCTTCTTCTAATGACAAA
AGAAACAAACCTAAAAAAGTTTCTAAAAGATCAGCAGATGTTCACAAATTGTCTGCTAGTGAATCAAATTCTGTG
TCATCTAGAAACCAACCGATCCAGAATGACTTCCTGTTAAATGGGATTGTTTCTGCAAAAAATTGGGCATACATA
CCAGTACGAATAGTATACAAGCGTCCACCTAATTGCCGTTTTAATATTAGAACGGCTAAGTACCAAAGAGTAATG
GAAGAAAAGGCGGAACATAAAAGACAAGAACGTTTGCATTTTGAAGATGGCGATTTGTACGATCCACGGGCGTAT
CCAAGACTTTGGGAAAAAATTCGAGAGGGTAATCCTGCTTCATTTGGTGAATTCAAAGTAAGGGGATCTGGAGAA
TCTAAAATATTTGTCGAAACCACTGGGTTATCCTACAAAGGAGAATATAAAGAGTATGCTATAGTTGACGAACGG
CAACCAATCAGTGAAGCTGTAGCATATGTTGCATTAAAGCGCCCAAGTGATGGTGAAACAGTCCAAACATTTGTA
ACAGCTTTTGATCCCAATGGACGTGTTTGTGAGGCCCATTGCCATGTGAAAGGCTCAAACCCACCAAAATATGAA
CGCTGTTCTGGTGTTGTAAATATTAACTCTAAAGAACCAAAAATGTATGGGGATAATTTGGGTGATGCTTATTTG
CTAAGGTACCGTTTCCAAGGCGACAATCTACCTAGTAGTCATGATGGAGACGTCTTCTTGATGTTTTTCTGTAAC
TATGGGAAACGATTTCCATGGGAAAAGACGTATAATAAAAAATAAATACATTTTGTATATTATACATGCCGATTT

TTCTTTACAGAATGTACAGAAATTCATTTAATGATTTTTCTTTTTCTTTATGTCATAAACGTTTACCAAACAAAT
TTCCAAGCAATTGGCTATGACTGAAAATGCCATTCCATTTGTTCAGTAACATTAAGCGTTCGATGTGAGGCTTCT
TCACAGGACTGACCTACAGCAGCTAATATTTAAGTCATATAACAATAGTTTATTAATGTGGTGCAAATCACAGAA
ATCGATTAAAAAGATACAATACAAGCTGACTCGCAAAAACTTCGGAAAATGCACGATCTCCAAACAGAACTAGAA
CCAGGCGTGTCTCCTGTTATTCATGCTACATCCACAGTGTGAATCAATACTGTCGTTCTAATTGTTACTGTCATA
AATGGAAAACCGTCCAGAGTTTTTTGCTAATGAACTTGAATTTGAGGCTTCATGTCACGTGTAAGATTACAAACA
ATAGCTTTCCCGCCTTCATGATGAGCTTTTATCGTTATAAAGTCTTTTTTTTATTAACGATTAATGTTCCTATAA
TATGCAAAAACATAGTTCTTTAGTAGCAGAGCACCTGGTTTCAAACTTCAATCAAGTTTATATTAAATGGTTTGT
CTGTGTATATTCGTCGAATAAAACTTGATATAAACATTTAAATGGCAGTGGTTGATCGTTTTGTACACCTTTCAA
AAACAACGTCTCAAAACTCCATATTTGAAAATTGGTCAATGTGTTTAAAAGTTGCCTCTTTTGACAGGTTTTCGC
CTTTGCCTTCGACGAAAGTCACTTCACCTATTTTCAAGATTCAGTTACTGCTTAAAAAGCCCATTTTTATTCAAT
TGATTTCTCCATTTTGATGATATCATAGCCGG

MNSYCALLLIIINGVEALIYKDWFPGPMEKCLMDRSRGVSPRRIPAFDILFECKNYQVAYNNVDNDVVSPVTEDN
ERYFKHLGRRLQGLESEYKRRKRSAKWKWNNQRKEIRTLTDKQRSDYFAALNALKKDGSYDAITRLHQQTAIMGA
HFGPGFLGWHRIYNLVLQLAIWDKNPRVMLPYCDTTLDHNMIDPKKSILFSKSYFGNLRGIVVHGPFVGWNTTAN
VLLQRDGGRFGSYMTDEGYKQTMSLKHMREVTLGTSLGQIERQHNNPHVYIGGLMSDLNYAPSDPFFFCHHNYID
YVWEKFRENQRRKNIDPGSDYPSNTMNLHEPGRRLPMIADYLKVNLTQAQCSHNMFSDKFLINFANSPKYPNCGN
YPHIKRNDTRKVCYSAEYPSFEDDNVVALQTRVRSNKDATKFKVTKGFVASSNDKRNKPKKVSKRSADVHKLSAS
ESNSVSSRNQPIQNDFLLNGIVSAKNWAYIPVRIVYKRPPNCRFNIRTAKYQRVMEEKAEHKRQERLHFEDGDLY
DPRAYPRLWEKIREGNPASFGEFKVRGSGESKIFVETTGLSYKGEYKEYAIVDERQPISEAVAYVALKRPSDGET
VQTFVTAFDPNGRVCEAHCHVKGSNPPKYERCSGVVNINSKEPKMYGDNLGDAYLLRYRFQGDNLPSSHDGDVFL
MFFCNYGKRFPWEKTYNKK

>Medu-TYR15_comp79670_c1_seq2

GTAAAACAAGAGTAACATCTGATAATTATTGATCGATTAAATTACACTTGTAACATCTATATCAAAATAATACTT
GCTATTACATGTATTCAATAGAGAGATAAAATAATACGCCGGCTTTTCATTGTACCCATGAAGCCTACATCACAA
CTACTGCTATTGCTCCTATGGCATTCCGTATTATTATCCGTCTGTGACGTCATAACTAACCAGTATCCTGCTTTG
GTTTCCGAATGTAAAGATCATGACTCACAGAAGCAATATAAGTGTATACTGTCAAATCTTGACAAATCAGATTGG
TCACGTGACGATGCAGATGGCTTAAAACTGAGATTACGACAAGACTTTTACTGGCATCCAACAAAAAGGATACGA
CGAGAATGTAGAGCACTAACGAGAGAAGAATTCAAAGAAATGGTTGATGCTATAAATGCTTTAAAGAAAGATAAA
AGTATGTCTCCTAATGTTTATGACTACATTGCTGATTACCATACAAACGAAGCAGTTAATTCGGTTCACTTTGGT
TCCAACTTTTTTGGCTGGCATAAAATGTATATTCTAAAATTTGAAGAGCTGCTACGAAAAGTCAACCCAAATGTA
ACATTATGCTACTGGGATTCAAGATTAGACCACCACATGAAAGACCCCAAGAAATCCATGATGTTTACATCAGAG
TTTTTCGGTGATGCTAGAGGTCCAATCCGAACAGGACCTTTTGCACGTTGGAAAACAATAAGAAATAAACTTCTC
TTTCGAGACTTAGGAAAAGAAGGAAATCCCATTTCTTATAAAGATATAGAAACCGTTTTATCAAAAAAGCACCAC
GTCCAAATAACGAACCCGACATCGGATCCTAACAGTGAAGTCGAAATGATGCATAATATGGTCCATGCCTTCGTC
GGTGGTCAAATGAATGATTTTAATACTAGTTCACAAGATCCATTTTTCTGGATTCATCATGCTTTCATTGATGCA
GTTTGGACCGTTTTTTGTCGGCAGTTACGTCATCGTAATATAGATCCACAAGACGATTACGTCATCGTCGATAAT
GAAATGCACAAACCCGAGAGATATATGGATCATCTATTTCCTATGAAGAATATAGATGGATATAGTGATTACTTT
GCCAATAACATATATTCGTATGGTGATTTTGCAAAATGTCCTACTTGTTTAAACTCTCCATACCTAAAATGCGAC
CATGCATCAAACAAGTGTGTAGGCATCGAGCTTCACGAACCACGTAGAAGGGTGGAGACCCCACCAAAAAGAATA
AAAACAAACCATTTATCTACACCTCAGAATGATTTTACCGTACTGGCAGGAGATGATAAAAACTGGGTCTATATC
CCGGTTAAAATTGTTGTTCGGAATGCACTAGAAAAATCTCTAAAGAAAAATGTCATAGGTGGATGTGAATATTTA
GGAGCCGGATATGATGATTTATGTTCAAAGTCTGTAGCCATAGTTCAATCCAATGGAATTACATACCACGGAACA
TATAGAAATTATATTGTAAACGATGCTTCTATTCCTCAGTGGGTATACGCATATGTTGGAGTTAAAGATCCAGGA
ACAGGGTCCAGTCAGGCTTTCATCAGTATAACCGACAAAAACCACAAGCCATGTGATGCATTTTGTCTTGATCTA
AAAGCGAAAAAGTATCGGTCATGTCCAGGAGTAATTTCCGTAACAAATAAGAAACCAAGGCTGTACTATAACACC
CTTAAGGAAGCAGAGGAAAATGGGTATCCATTTAATCCTTTAGAGCCGGATACTCTTGATCCGCGAATTAAAATG
GCGTTCCTTTGCTATTAAATCTGATGGAAGGCATGAATCTTGTTTTATTTTACATTGGATTTCCTTTAAATATAA
AGAGAATCAGTACTACTTTTAGATTTATGCAAACGAGCATAATATTGAAATTTGTTATTAAAGCTATTTATTGGG
AAGGTTTTACATATGTTTCCTTATTCATGTTAATAGTTATTGGCATGGAATTGAAAGAAAAAATAAAAAAAACGT
GGA

MKPTSQLLLLLLWHSVLLSVCDVITNQYPALVSECKDHDSQKQYKCILSNLDKSDWSRDDADGLKLRLRQDFYWH
PTKRIRRECRALTREEFKEMVDAINALKKDKSMSPNVYDYIADYHTNEAVNSVHFGSNFFGWHKMYILKFEELLR
KVNPNVTLCYWDSRLDHHMKDPKKSMMFTSEFFGDARGPIRTGPFARWKTIRNKLLFRDLGKEGNPISYKDIETV
LSKKHHVQITNPTSDPNSEVEMMHNMVHAFVGGQMNDFNTSSQDPFFWIHHAFIDAVWTVFCRQLRHRNIDPQDD
YVIVDNEMHKPERYMDHLFPMKNIDGYSDYFANNIYSYGDFAKCPTCLNSPYLKCDHASNKCVGIELHEPRRRVE
TPPKRIKTNHLSTPQNDFTVLAGDDKNWVYIPVKIVVRNALEKSLKKNVIGGCEYLGAGYDDLCSKSVAIVQSNG

ITYHGTYRNYIVNDASIPQWVYAYVGVKDPGTGSSQAFISITDKNHKPCDAFCLDLKAKKYRSCPGVISVTNKKP
RLYYNTLKEAEENGYPFNPLEPDTLDPRIKMAFLCY

>Salv-TYR1_comp274293_c0_seq4

TTACTTTCACACTATGTAGTTGGGGTGAAAAGCAAGCATCATAAATACGAGTATGACGACGACGATGAAGATGCT
GAAGACAGTATAGTCATCCCAGAGAAAGAAGAAAAAGAAATTAAAGATATGAATGACAAGTGGAAAATTAGTAAA
GAATTTTTACCGTTAACGCGTCGGGAGCGAAAATTACTTAACAGGATTTCGAGGGGTAGAACGCATTGGAAACGG
AAGAACAAATGGGGAAATTACTACACTGGACACAACGGGGGACATGATAGAAAAGGAAAAGGGCGTGGACATGGG
CATGGTGGAGGCCATAGGCATGGTAGAGGTCATGGTCATAGACGATCGAAAAGACAAAGTGGTTCTCTTAAAGTC
CGCCGGGAGTACCGTAATTTGACTGACTCTGAAAGACAAGCTTTTCACGTAGCTGTACAACAACTAAAAACTAGG
ATGGTGGATGACATAAGCATGTATGATATATTTGTTGATATGCATAACGCTGTGTCTGCTCCTGGCGCTCATGGG
GGTGCTGCGTTTCTCCCATGGCATAGAGAATTCCTCTACAGGTTCGAGAAAAAGCTACAAGAAATCAACCCAGAC
GTTACTCTACCCTATATCGATACCTGTAATGAACGTTATATCTTTAATCGTTTTCCTGACTCAGCGCTATTCGAT
GATGACATGATAGGCAACGCAGATGGACCAGTAACATCAGGTCCATTTAGTCATTTTAGTATCGAGGGTGCAACA
TGTAAAGCCGTCGGACAAACACTTAGGCGTAATCTTAGGGTGGATATCAACAACCTCTTTAACAAGGCGAAGGTG
AATTTCATAACTACTCGCAATTCATATGATTACCTTACAAAACCAATGCAACAGGCCGCTAAAACGGTTGAGGCT
TATCATGGGGGATTTCACGTTGGTTTAGGAGGACATATGGGATGGCTGGAGTGTGCACCATTTGATCCTGTCTTC
TTTATGCACCATTGTTTTGTTGATTACTTGTGGCATCGTCAGCAACAGCATCAACAACAAATGAATGTTCCTTTG
GAGTATCCTGACGTAGCTTCCCCAGATAGAGCTGATGACTTGATGAAACCATTCGAGGAGACTGATGACGATGGG
AATACTACACCAATGTACAACAGGGACGGCTTCTCCGAGGAGTATATGGAGGATTATACATACCAGGTGGCACCC
TGTGAAATAGATTGCGTTAACGATGAGGACTGTGGTGCAAGTCGTCTTTATTGCTATAAAGGTCGCCGCCACGCG
CACGGAACATGCCGTGCTAAAGTGCGGGCTGGTGGTAAATGCCAAGGCAACATTTCCAGAAGTTGTTTCTGCGCT
ACTGGAACAGCAAATTGTAATACAGTCGGTAACTTGAGAGGCTTCTATTGTACATGTACTTAAACAAACTCAACA
CCGGTGTCTATTTTCTTCATTTTGTCC

LLSHYVVGVKSKHHKYEYDDDDEDAEDSIVIPEKEEKEIKDMNDKWKISKEFLPLTRRERKLLNRISRGRTHWKR
KNKWGNYYTGHNGGHDRKGKGRGHGHGGGHRHGRGHGHRRSKRQSGSLKVRREYRNLTDSERQAFHVAVQQLKTR
MVDDISMYDIFVDMHNAVSAPGAHGGAAFLPWHREFLYRFEKKLQEINPDVTLPYIDTCNERYIFNRFPDSALFD
DDMIGNADGPVTSGPFSHFSIEGATCKAVGQTLRRNLRVDINNLFNKAKVNFITTRNSYDYLTKPMQQAAKTVEA
YHGGFHVGLGGHMGWLECAPFDPVFFMHHCFVDYLWHRQQQHQQQMNVPLEYPDVASPDRADDLMKPFEETDDDG
NTTPMYNRDGFSEEYMEDYTYQVAPCEIDCVNDEDCGASRLYCYKGRRHAHGTCRAKVRAGGKCQGNISRSCFCA
TGTANCNTVGNLRGFYCTCT

>Salv-TYR2_comp275725_c0_seq1

CGAGAATTGATTTCTTTCATTCGTATTTTGGGAAGAAACACGTGAACACTAGAATGGGATCAGAAAAAAATTTCT
GGTTCATTTGTGCAATTTTATGTTGTACAATTATCAATATTCAGAGTCTGACATACGAAAATTATGATGGTGATA
TTAATGAAGGAAAAGAATATGAAGGTCTGACCAAGGAGGAAATTGAAAAAATTCAAAAGATGGAAGGAGGCTGGA
AATATTCAATACCATTTGAAGAGCTCACAGACAAAGATAGAGAATTTTTGAGAAAAATTTCAAAGAAACATGGTC
ACAAATCATGCGGCAGAAAACACGAGCATAACAAAGGGAAAGACAGAGGCGTCGGCAAAGCCAAAGGAAAGGGCA
AGGCAAGAGGACATGGCAGAAGGAAAGGAAAGGGCAAAGGAAAAGACAAAAGACGTCATAAAAGACAAACTGGGG
GCACGAACATACGTGTTCGCAAGTCTTACCACAAGTTGACAGAACAGGAAAGGACTGATTTTCACAATGCTATTA
TAGCTCTGAAGAATACACTAACAGATAATTCTAACTGGTTTGATGTGCTGGTTGACTTACATAACAGCCTGAATG
CTCCTGGTGCTCATAGAGGTCATGCTTTTCTCCCGTGGCACAGAGAATATCTTTTAATGATGGAAAACGCACTGC
AAACAATCAATGAAGATGTGACGTTACCGTACTGGGACAGCTGTATGGAGGCTAACATGGGAACTCGCATGAAGG
ATTCAGCATTATTTGACGATGACACCATGGGTAATGCTGATGGAGTAGTTACCACAGGTCCCTTTGCTAACTTCA
CTGTCCAAGGGGATAGGTGTATCAATTTATCGCCGACTCTACTGCGGGATTTAGACTCCAATCCACATCTTTATA
TAAACCGTACGAAAGTTGATCGTGTCAAGGGATGGACTTGGTACGGGGAAACTGGAGTACAACTGGAATTTATTC
ATAATCCTGTGCACGTGGGTATCGGCGGACATTTGCCAGATTTGGATTGTGCACCCTTCGATCCATTTTTCTGGA
TGCATCATTGTTACATTGACTTCCTGTGGGAGAGAGCAAGGGAGAACCAACAAACATCTACAGAGTCTGATTATC
CCCCAGACTGGTATATAGATAGCCGGCCTGGACGTTTTAACCAACTCAGTAAGGCATACGATTGGATGCAACCTT
TTTATAGGACTGATGAAAATGGAAACCAGCAACCACTGCGAAATATTGATGGTTTGAGCAACGGTTACACAGCGG
ATTGGTACACGTATGAACAGGCGCCTTGTGAAATACAATGCCAAACGGACTCTGACTGTGGAGCCCTTCGTCTGT
ACTGCCAAACCAACCACCAGCGATGTGCGGCTAAAGTAAGGATGGGTGGTGAATGTGAGGGTAATATCGCCAACA
GCTGCTTCTGTTCAACAGGCACACCAAATTGTGACACAGCTATTTGTCAGTGTGATTAGAAGATGGCTATTTGGA
TGAACTTCTGTCATGATTTCAATAAAGATAGTTTCCTTGTACATTGTTCAAAGAATATATCAG

RIDFFHSYFGKKHVNTRMGSEKNFWFICAILCCTIINIQSLTYENYDGDINEGKEYEGLTKEEIEKIQKMEGGWK
YSIPFEELTDKDREFLRKISKKHGHKSCGRKHEHNKGKDRGVGKAKGKGKARGHGRRKGKGKGKDKRRHKRQTGG
TNIRVRKSYHKLTEQERTDFHNAIIALKNTLTDNSNWFDVLVDLHNSLNAPGAHRGHAFLPWHREYLLMMENALQ
TINEDVTLPYWDSCMEANMGTRMKDSALFDDDTMGNADGVVTTGPFANFTVQGDRCINLSPTLLRDLDSNPHLYI
NRTKVDRVKGWTWYGETGVQLEFIHNPVHVGIGGHLPDLDCAPFDPFFWMHHCYIDFLWERARENQQTSTESDYP

PDWYIDSRPGRFNQLSKAYDWMQPFYRTDENGNQQPLRNIDGLSNGYTADWYTYEQAPCEIQCQTDSDCGALRLY
CQTNHQRCAAKVRMGGECEGNIANSCFCSTGTPNCDTAICQCD

>Salv-TYR3_comp278284_c0_seq2

CCTGAGTAAAATATTCGTGTAAATGAATGTGTAAAATGTTCGTACAACGTTGCAGTTTTGGTGAAATAAGAATTC
CGTTATAAAGTGTGGTTTATCAACGATAAAACGATGAAACAAAAAATTAAACTAAACTAAATCCGGAAATTGGAA
GTTATACGTTGTAGCAAAATTAATTTCATAAATTATCCGTCGGCCCATATGATACAAAGGCACATAAGAATTTGA
ATTAAGAACAAGCAGTCAAATCACCAGCTTGTTTTATATTTTTATCTGGTTAAGACGTTTACGCTAAAATGGAAC
AACAGATAAAACTTTTATTTCTTTGCGTGATTTTTACGACTTACAAGTATGAGGCATTTGGAAGATTAACGGCAA
AGGAAAAGGAAGAAATACGACAAATGCAAGACGAATGGAAGCTTACCAAACCTTTCCTTATAATAACACCTAACG
AAGAAACATTTTTGAAAAACATTTCGAAAATTCCTATGCATATCCCAAAAGGAGACGTAGAAGAGAGGGAACGTA
ACCAAAGTAACGAGAGTGGCAAAAAGCGCAAGAAACGACAAATAGAAATTCCACAACGTAGAGAATATAGGGAGT
TAAGTGATGACGAAAGAAATAGATTTCACAATGCTTTCATTGGACTGAGAAATAGTATAATTGATGATTCGAATG
AGTTGGATCTGCTAGTTGGTTTGCACAACAATTTTTTTGCTCCTGGTGCTCATCGGGGTCCTGCGTTTCCTGTTT
GGCATAGACACTACTTGTACATTTTCGAGAAAGCTTTGCAGACGATAGATCCTAGTGTAATGTTGCCATACTACG
ACAGCAGTCCCGAACGTAACATAGGTCGTCGCATGATTGATTCGGCAATATTTGACAACGACACCATGGGCAACG
CTGCAGGTCAAGTGGTCACAGGTCCATTTGCATATTTCGACGTCGTAGATAATTGTAAAAGGATGAGAGGACGAA
CTTATGAAAGTCTCACTCGTAATTTGAACACAGACAGTAATTTATATTTCGATAGTTCTAAAGTGCGCACAATAA
AAAGATGGGGCAGGTTTTCTGACACATTGCAGTTAGAGGGGTACCATAATGGTGCTCACATTGGTATTGGGGGAC
ACATGGGCAACCTGGGGTGTGCACCGTTCGACCCTTTCTTCTGGATGCACCATTGTTTTGTTGATCTACTGTGGG
AGTTAGTACGGGAGAACCAAATAGACGACGGCATTGATCCAGAAAAGGATTACGTAAGAAACGCTCCTAGAAACC
AGGGGTTCGACGATCCAATGGGACCTTTCACTACAACTGATGAGAACAATAATATAGTACCACTTATTGTGGCAG
ACGGCATGCTAAACTGGTATACTGATGACGCGTATGGGTACACGGTGGCGCCAAGTGAAAAACGATGCAGCCGAC
GTATAAAATGTGGAGCAAGCCGCCTGTACTGTCACAGAGGAACCTGCAGGGCTAAAGTGAAAGAAGGGGGCGAAT
GCCATGGCCGAAAGCGTAAAAGTTGCTTCTGTTCTTCTGGACGGCCACGGTGTGTGCGAAACGGTAGACCATATT
ACTGTACCTGTGTATAACGTGATAACGTCGATGTTTTGATGCAGAGTCTGGTAGAGGCAATAAATAGATAGCATA
AAAATAGTTTTAGTTGACGGTTATGTGAATTCTTTAGTGGCAGTCTCACATATTGTCAGTGCATACTAAGATCGT
AATTGATTTTTGTATACTGATTAAAAATTGAAAATTTATTGTTTATTTATTATTAAAGAATGTAAAATGAAAA

MEQQIKLLFLCVIFTTYKYEAFGRLTAKEKEEIRQMQDEWKLTKPFLIITPNEETFLKNISKIPMHIPKGDVEER
ERNQSNESGKKRKKRQIEIPQRREYRELSDDERNRFHNAFIGLRNSIIDDSNELDLLVGLHNNFFAPGAHRGPAF
PVWHRHYLYIFEKALQTIDPSVMLPYYDSSPERNIGRRMIDSAIFDNDTMGNAAGQVVTGPFAYFDVVDNCKRMR
GRTYESLTRNLNTDSNLYFDSSKVRTIKRWGRFSDTLQLEGYHNGAHIGIGGHMGNLGCAPFDPFFWMHHCFVDL
LWELVRENQIDDGIDPEKDYVRNAPRNQGFDDPMGPFTTTDENNNIVPLIVADGMLNWYTDDAYGYTVAPSEKRC
SRRIKCGASRLYCHRGTCRAKVKEGGECHGRKRKSCFCSSGRPRCVRNGRPYYCTCV

>Salv-TYR4_comp264814_c0_seq1

TGTTCTTCCGCATTCTTATTCCGGACTGGGAACAACACACACCGCGAATATGCAACACAAACTTATGATTTTGGG
TGTTTTTGTGATTTGTATAAGTTGTGCTGTTTTGGCTTCTGATGATGATAGGGACCCCTCAATGCCGCCTGAAGA
GGCAGCGGGGATAGAAGAGATGTATAAAGCATGGCCGAGAAGCAAGCCCTTCAAGAAGCAAACTGAAAAGCAAAG
GAAATATTTGGCAAAACTTTCAGGTGGCCCTCACAATAAAAGAAACCGTGGACGTAAAAACAAAAATGGACGACC
GAAAAGACAAACTACTGCACCTAAAGTGCGTCGAGAGTATAGGAACTTGACCGACAGTGAGAGAATCGCTTTTCA
TGTGGCTGTTCAAACATTAAAGAATACAGTAGTAGACGGTGAAAGCCAGTTTGATGCGTTAGTTAATATCCACAA
CTTCCAGGATGCTCCGGGTGCTCACCGTGGGGCTTCCTTCCTTCCCTGGCATAGACACTTTCTCTACATGTTCGA
GAAAGCTTTACAAGAAATAAACGAGAATGTAACTTTGCCTTACTGGGACAGCTGTAATGAAAAGAACATAGGGGA
ACGACACATTTACTCAGCGCTTTACGATTCTGATACAGTCGGAAACACCGTGGGCCGAGTAACCAGTGGCCCATT
TGCTGATTTCCCTATCCAGGGTGAATGTATGGCAGACGGTGGAGATTATTTAACTCGCGGTTTACGACAGGATGG
CCCCTACTACTTTCGCAAAGCTGTTGATGTGGAATTAGTAACTCAATTCACGTCGTTTAGTGATATTGCACATGG
ACCTGTGTCAATAGAACGAAAGCATGGTGGAGCCCATTACGCTATGGGCGGACATATGGGATGGTTGCATTGCTC
GCCATTTGACCCGATGTTCTGGATGCATCATTGCTTTGTGGATTATTTATGGGAGAGTATGAGAGAAAATCAATT
CAACAATAATGTTGATTCTGAGACTGACTACCCATCAACAGACTATCCTTACGGAGCATCAGACCCGATGAAACC
GTTTACTGGTATCGATGAGAACAATAACGAAATAGAACTAACCAATAGTGATGGTATGCAGGATTGGTATACTGA
AGAGGTATATACTTACCAGCGTGCACCCTGCGAAATAACATGCAGCAATGATGAGGAGTGTGGTGGCAGCCGTCT
ATATTGTGAATCCAGTGAATGCCGTGCAAAAGTGAGAGAAGGCGGGCAGTGTGGCTCCTCAGACGCCAATCGTTG
TTTTTGTGCAACTGGTACACCTAGCTGTAACTGGTACACACAATCTGGAGAAAGCGACTACTGGTGTACGTGTTC
CTAATTTGAGTTCGCCTGGGTTGTGCCGACAATTTTCTCGCTGTTTGAACATGTATACTTTGCAACTTTGAGAAG
GTGAAATATTTTCGTCTGGAATAAACAAAAAAAAAATCGTGAAAATGTGGACTAAACATTCACACTTGGCAGTAT
TCGGAAAATTTGATTTCGGAATT

VLPHSYSGLGTTHTANMQHKLMILGVFVICISCAVLASDDDRDPSMPPEEAAGIEEMYKAWPRSKPFKKQTEKQR
KYLAKLSGGPHNKRNRGRKNKNGRPKRQTTAPKVRREYRNLTDSERIAFHVAVQTLKNTVVDGESQFDALVNIHN

Chapter I

FQDAPGAHRGASFLPWHRHFLYMFEKALQEINENVTLPYWDSCNEKNIGERHIYSALYDSDTVGNTVGRVTSGPF
ADFPIQGECMADGGDYLTRGLRQDGPYYFRKAVDVELVTQFTSFSDIAHGPVSIERKHGGAHYAMGGHMGWLHCS
PFDPMFWMHHCFVDYLWESMRENQFNNNVDSETDYPSTDYPYGASDPMKPFTGIDENNNEIELTNSDGMQDWYTE
EVYTYQRAPCEITCSNDEECGGSRLYCESSECRAKVREGGQCGSSDANRCFCATGTPSCNWYTQSGESDYWCTCS

Chapter II

The role of tyrosinase in mussel adhesion: Characterization of a recombinant mussel plaque-specific tyrosinase

Emilie Duthoo, Elise Hennebert, Mathieu Rivard, Alessandra Whaite, Matthew J. Harrington and Patrick Flammang

Contribution: Elise Hennebert designed the tyrosinase sequence to be inserted into the plasmid, as well as the protocol. I produced and purified the recombinant protein and performed the SDS-PAGE gel and western blot. The results were interpreted by Alessandra Whaite, Elise Hennebert, Patrick Flammang, and myself.

Chapter II

The Role of Tyrosinase in Mussel Adhesion: Characterization of a Recombinant Mussel Plaque-Specific Tyrosinase

Emilie Duthoo[1], Elise Hennebert[2], Mathieu Rivard[3], Alessandra Whaite[1], Matthew James Harrington[3] and Patrick Flammang[1]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cell Biology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[3] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

Abstract

The blue mussel *Mytilus edulis* produces a protein-based adhesive attachment called byssus, composed of threads ending in a plaque that attach the organism to the surface. The byssal threads are formed by the self-assembly of DOPA-containing adhesive proteins secreted by specialized glands in the mussel's foot, with the plaque gland forming the plaque. DOPA, which is formed by the post-translational modification of the amino acid tyrosine by an enzyme called tyrosinase, has been shown to facilitate substrate adhesion and enhance the cohesive strength of byssal material through a variety of molecular interactions. In this study, we focused on a specific tyrosinase (Medu-TYR12), which was found expressed in the plaque gland and is therefore possibly involved in the maturation of adhesive proteins. Recombinant protein technology was used to produce this protein, abbreviated rMT12, with the objective of characterizing its enzymatic activity.

Chapter II

1. Introduction

Many marine invertebrate organisms have evolved adhesive strategies to cope with the challenges posed by the hydrodynamic forces of intertidal waves. These organisms exhibit the ability to attach themselves permanently or temporarily to surfaces using underwater adhesives, which have proven to be stronger and more durable than conventional synthetic glues (Lee *et al*., 2011; Almeida *et al*., 2020; Li *et al*., 2021). Among these organisms, the blue mussel *Mytilus edulis* Linnaeus, 1758 has been extensively studied. The blue mussel secretes a protein holdfast known as the byssus. This structure comprises numerous load-bearing byssal threads covered by a protective layer called the cuticle, each ending in a flattened plaque that allows adhesion to the substrate (Waite, 1983a). Byssal threads are made one at a time by rapid injection moulding, each taking between 30 seconds to 8 minutes (Martinez Rodriguez *et al*., 2015), and are formed by the self-assembly of a dozen of proteins from three distinct glands inside the mussel's foot: the plaque gland, the core gland, and the cuticle gland (Priemel *et al*., 2017). The glands secrete their content and are organized around the ventral groove and the distal depression of the foot. They contain spatially organized vesicles holding byssal thread precursor proteins. During secretion, the contents of the vesicles spontaneously coalesce and self-assemble into biomolecular micro- and nano-architectures, resulting in the formed byssal thread (Priemel *et al*., 2017).

Thanks to recent omics research, more than 20 different adhesive proteins were identified (DeMartini *et al*., 2017). The plaque gland secretes the mussel foot proteins (Mfps) that are destined for the formation of the adhesive plaque. These mfps contain many post-translationally modified amino acids (Waite, 2017). These modifications are important as the functional groups they bring influence the physicochemical properties of the cohesive and adhesive proteins, enabling interactions with the substrate and curing of the glue. One of the key modified amino acids is DOPA (3,4-dihydroxyphenylalanine), which has been extensively investigated and appears to play a crucial role in the adhesion of mussels (Waite, 1983b; Waite, 2017). DOPA is formed through the post-translational hydroxylation of tyrosine residues via a tyrosinase enzyme (Ramsden & Riley, 2014). This modified amino acid can bind to different surfaces through hydrogen bonds, covalent cross-linking, electrostatic interactions, or by forming coordination complexes with metal ions and metal oxides. Additionally, it contributes to the cohesive strength of the adhesive material (Lee *et al*., 2011; Petrone, 2013; Waite, 2017; Priemel *et al*., 2020).

Our previous study identified and characterized several tyrosinases from the byssus of *M. edulis*. One of these candidates, Medu-TYR12, was localized only in the plaque gland of the foot with *in situ* hybridization experiments (Chapter I). Additionally, this tyrosinase was identified in induced byssal threads (Chapter I) as well as in the vesicles present in the plaque gland (Mathieu Rivard, unpublished data) by mass spectrometry analyses, reinforcing its potential role in adhesive protein maturation. Consequently, we selected this specific candidate for production using recombinant protein technology, in order to test its enzymatic activity via UV-Vis spectrophotometry.

2. Materials and Methods

2.1. *In silico* analyses of Medu-TYR12

The computational analysis of the tyrosinase sequence Medu-TYR12 (encoded by transcript comp83122_c1_seq1 obtained from the transcriptome of the foot of *M. edulis* (NCBI SRA: SRR29446349) was performed using various software tools. The amino acid length, composition, molecular weight, theoretical pI, and extinction coefficients were obtained via ProtParam - Expasy (Gasteiger *et al*., 2005). Domain searches were conducted using InterPro (https://www.ebi.ac.uk/interpro/search/sequence/) (Paysan-Lafosse *et al*., 2023), while the presence of a signal peptide was assessed with SignalP-6.0 (https://services.healthtech.dtu.dk/service.php?SignalP) (Teufel *et al*., 2022). The three-dimensional structure of the sequence was modelled using the AlphaFold Server, powered by AlphaFold3, using default parameters (Abramson *et al*., 2024), the protein's 3D structure was then modified using the software ChimeraX (Version 1.9rc202411272016 (2024-11-27)) (Meng *et al*., 2023).

2.2. Production and purification of recombinant tyrosinase

The sequence of Medu-TYR12 was codon optimized using the Codon Optimization Tool on the Integrated DNA Technologies website (https://eu.idtdna.com/pages/tools/codon-optimization-tool?returnurl=%2FCodonOpt) and prepared by gBlocks Gene Fragments (Integrated DNA Technologies). The signal peptide and the stop codon were removed from the sequence for expression in *Escherichia coli*. The enzyme restriction sites NdeI (5'-CATATG-3') and XhoI site (5'-CTCGAG-3') were added at the beginning and the end of the sequence, respectively. The full-length sequence was amplified by PCR using Q5 High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA). The obtained PCR products were then cleaved by NdeI and XhoI (Thermo Fisher Scientific) and ligated using T4 DNA ligase

(Thermo Fisher Scientific) into the pET-28a (+) expression plasmid in frame with a C-terminal 6 x His-tag coding sequence. The sequence of the inserts was confirmed by Sanger sequencing (Eurofins Genomics). The corresponding recombinant proteins were named rMT12.

The plasmid was first transformed in the *E. coli* strain DH5 alpha (New England Biolabs) prior to cloning before being transformed in the *E. coli* C2566 strain for protein expression (New England Biolabs). A single colony from a freshly prepared agar plate with 50 µg/mL kanamycin was grown in Luria-Bertani broth (LB) containing 50 µg/mL kanamycin at 37 °C with shaking at 250 rpm until the OD600 reached 0.6. Protein expression was then induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Three hours after induction, cells were harvested by centrifugation at 5,000 x *g* for 10 min at 4 °C, and the pellet was stored at -80 °C until further processing.

Bacterial pellets were then resuspended in 20 mM Tris, 300 mM NaCl buffer, pH 8.0 containing 1 mg/mL of lysozyme and kept on ice for 30 min. Cells were lysed by ultrasound using an IKA U50, IKA Laboratory (10 cycles of 10 sec at 40% amplitude, with 30 sec intervals on ice). Inclusion bodies were separated by centrifugation at 16,000 x *g* for 10 min at 4 °C and resuspended in a 25 mM Tris, 150 mM NaCl buffer supplemented with 1 mM dithiothreitol (DTT), 8 M urea and 5% glycerol pH 8.0 (buffer A), rocking overnight at room temperature (RT) at 250 rpm. The extract was centrifuged again at 15,000 x *g* for 20 min to remove insoluble debris. The supernatants were then loaded onto a His GraviTrap Ni Sepharose 6 Fast Flow column (GE Healthcare) pre-equilibrated with buffer A. Unbound proteins were washed away (x 2) with buffer A containing 30 mM imidazole, and bound proteins were eluted from the column with buffer A containing 500 mM imidazole (pH 8.0).

To progressively refold rMT12 and remove the denaturing agent, the eluate was slowly dialyzed against buffer A containing 0.3 mM of $CuSO_4$ (Wang *et al*., 2020) and decreasing molarities of urea, from 8 M to 0 M. It was then slowly dialyzed again to remove the DTT. Protein concentrations were determined after centrifugation by UV spectroscopy using the calculated extinction coefficients from the ProtParam website. At each step of the above experiments, a protein aliquot was collected and run on a 12% sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) gel which was then stained with Coomassie Brilliant Blue R-250 to check the presence of our protein.

2.3. Antibody production

The peptide CSNDKRNKPKKVSKR from Medu-TYR12 was used for polyclonal antibody production in two New Zealand rabbits via the PolyExpress immunization strategy (Genscript). This peptide was selected based on its potential for successful synthesis and immunogenicity (analyses carried out with the Optimum Antigen Design program of GenScript). The resultant anti-CSNDKRNKPKKVSKR antibody is named anti-rMT12 hereafter.

2.4. Western blot

Proteins collected at different steps of the recombinant tyrosinase production and purification were subjected to SDS-PAGE (12% gel), and the proteins were then transferred onto polyvinylidene difluoride (PVDF) membranes (GE Healthcare) using a transfer buffer consisting of 25 mM Tris, 192 mM glycine, 0.05% SDS, and 20% methanol. The protein transfer was run for 1 hour at 260 mA at 4 °C before washing the membrane three times in a phosphate-buffered saline with 0.05% Tween (PBS-T). The blocking step was performed with 5% bovine serum albumin (BSA) in PBS-T (PBS-T-BSA) for 2 hours at RT before adding the anti-rMT12 antibody or a commercial anti-tyrosinase antibody GTX16389 (Genetex, Bio-Connect B.V., Netherlands) diluted at 1:1000 in PBS-T-BSA, at 4 °C overnight. The membrane was washed five times in PBS-T before adding goat anti-rabbit IgG (H+L) secondary antibody conjugated to horseradish peroxidase (HRP) (Thermo Fisher Scientific, USA), diluted 1:750 for 1 hour at RT. The Lumi-Light Western Blotting Substrate kit (Roche) was used on the membrane for 5 min, followed by rinsing in PBS before visualization using FUSION FX imaging system (Vilber).

2.5. Measurement of tyrosinase catalytic activity

Mushroom tyrosinase (6540 units/mg, Sigma) was added to different substrates: 50 mM L-tyrosine, 50 mM L-DOPA, or 50 mM 4-methylcatechol, which were prepared in phosphate buffers of different pH levels (2, 4, 6, and 8) (Qin *et al*., 2014). These different solutions were placed in 96-well plates with a flat transparent bottom. The absorbance was measured using a plate reader UV-Vis spectrophotometer for 120 min, with three measurements taken every 15 min at RT. Parameters were calculated to determine catalytic activity (Supplementary Document 1).

3. Results

3.1. Characterization of the mussel plaque-specific tyrosinase

The sequence Medu-TYR12 consists of 694 amino acids. It contains both a tyrosinase domain (pfam00264) from amino acids 133 to 308, and three copper-binding domain signatures (from amino acids 151 to 168, 268 to 279, and 288 to 306) (Fig. 1A). Medu-TYR12 also includes a signal peptide, indicating that it is potentially secreted with the adhesive to form the byssus. Its molecular weight is 80.39 kDa, with an isoelectric point (pI) of 9.38. The three-dimensional structure of Medu-TYR12 was analyzed using AlphaFold (Fig. 1B). The results show that Medu-TYR12 is mainly composed of helicoidal structures, but there are also regions with a disordered structure predicted with very low confidence (this may indicate either that AlphaFold is not sufficiently accurate at predicting this part of the protein, or that there is no defined structure to predict).

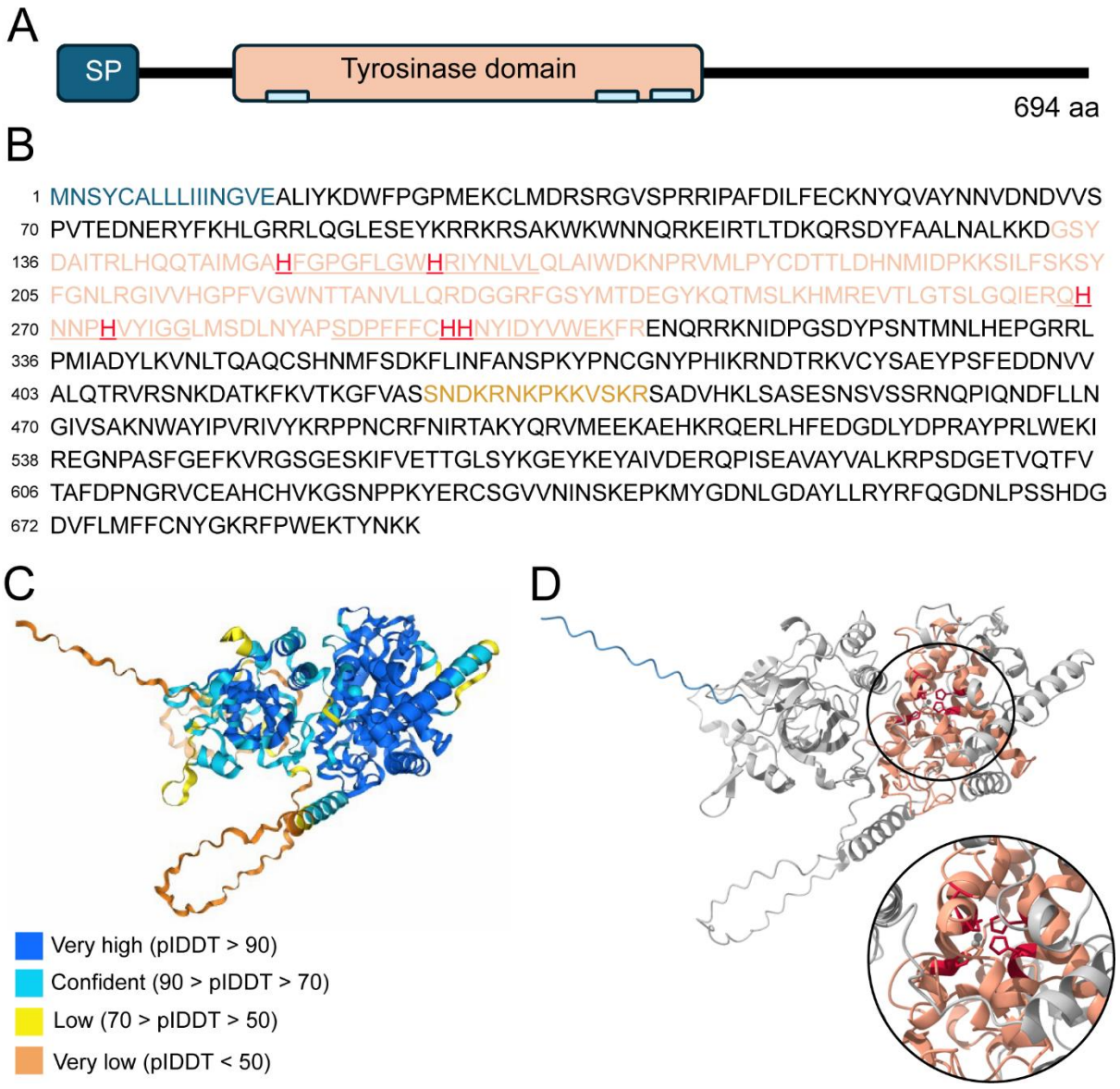


Figure 1. Conserved tyrosinase domain in Medu-TYR12 with the signal peptide (SP) and the three signatures of the copper-binding domain in light blue (A). Amino acid sequence of Medu-TYR12 with the signal peptide highlighted in blue, the tyrosinase domain in beige and the three copper binding sites underlined, with two times three histidine residues coordinating the oxygen in the active site in red (Olivares *et al*., 2002). The peptide sequence selected for the antibody is shown in yellow (B). AlphaFold result of Medu-TYR12 containing two copper ions with confidence on the structure (C), and its 3D structure highlighting the signal peptide in blue, the tyrosinase domain in beige, and the histidine residues in the active site in red (D).

3.2. Recombinant tyrosinase production

To test the enzymatic activity of the newly found mussel plaque-specific tyrosinase, Medu-TYR12, we produced it using recombinant DNA technology. After removing the sequence coding for signal peptide and optimizing the codons based on *E. coli* codon bias, the Medu-TYR12 coding sequence was recombinantly expressed in *E. coli*. Following three hours of

induction, a protein with an apparent molecular weight of ~80 kDa, as predicted by ProtParam, was observed in the bacterial extract after SDS-PAGE. The rMT12 was found in inclusion bodies (insoluble fraction), and after purification, a clear band was observable on the SDS-PAGE gel at the correct molecular weight, as well as other lighter bands at lower molecular weights (~65, ~55, ~50 and ~35 kDa) (Fig. 2). During the refolding process of rMT12, a noticeable precipitate formed in the dialysis bag. To prevent this precipitation, various dialysis methods were tested: a slow or direct dialysis from 8 M urea to 0 M urea, and different concentration in salt (tested with 0.5 M, 0.3 M, and 0.1 M NaCl). Under each condition tested, rMT12 still precipitated.

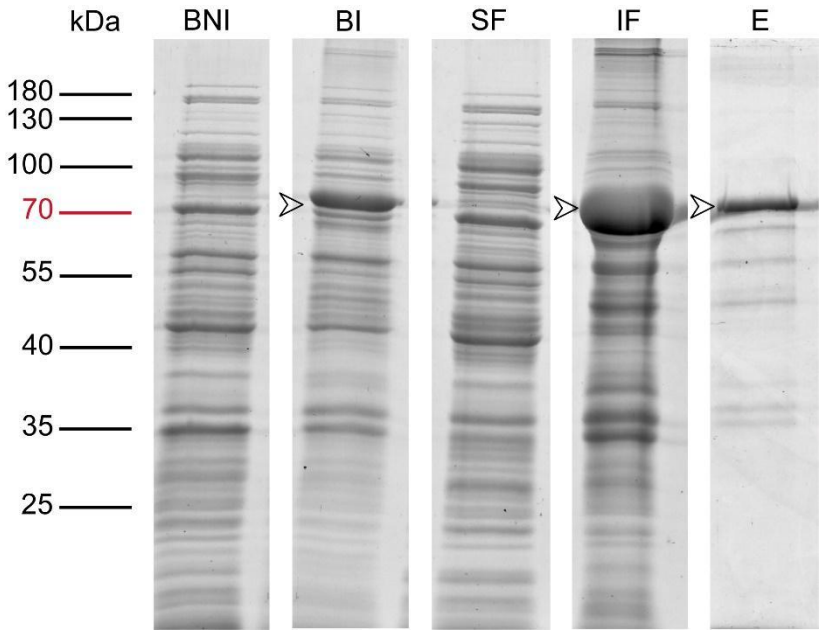


Figure 2. Expression and purification of rMT12 in *Escherichia coli* C2566 strain on a 12% SDS-PAGE gel: BNI - non-induced bacteria, BI - induced bacteria, SF - soluble fraction after lysis, IF - insoluble fraction after lysis, E - eluted fraction after purification on a HisTrap HP column. Arrowheads indicate the protein bands corresponding to the produced recombinant proteins, which are at approximately 80 kDa.

To confirm the identity of rMT12, Western blot assays were conducted using two types of antibodies: a custom-designed anti- rMT12 antibody (Fig. 3B) and a commercial anti-tyrosinase antibody (Fig. 3C). As observed on the SDS-PAGE gel results, several bands were visible in the purified fraction, suggesting that rMT12 may undergo cleavage during the process. A Western blot was therefore performed directly on the induced bacterial fraction, which also revealed multiple bands, indicating that rMT12 is already fragmented during the production step (Fig. 3A).

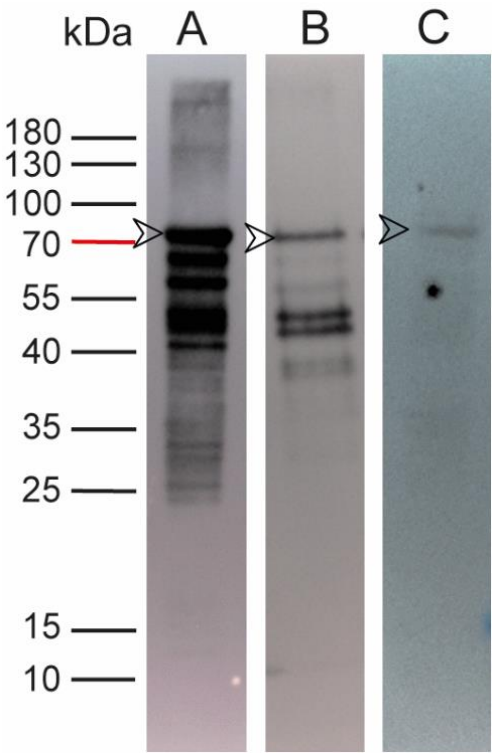


Figure 3. Confirmation of Medu-TYR12 identity by Western blot: on induced bacteria fraction using anti-rMT12 (A); on the eluted fraction of the recombinant protein using anti-rMT12 (B) and the commercial anti-tyrosinase antibody (C). Arrowheads indicate the protein bands corresponding to the full-length rMT12 sequence, which is at approximately 80 kDa.

3.3. Enzymatic catalytic activity

The enzymatic catalytic activity can be characterized by different parameters such as the Michaelis-Menten constant and Km and Kcat (Michaelis & Menten, 1913; Johnson & Goody, 2011; see Supplementary Material 1). Tyrosinase exhibits two distinct activities: monophenolase catalytic activity, which hydroxylates the tyrosine residue into DOPA, and diphenolase activity, which oxidizes DOPA to Dopaquinone (Fig. 4) (Ramsden & Riley, 2014). The absorbance of tyrosine residues and DOPA is recorded at 280 nm (Burzio & Waite, 2000; Tajima *et al*., 2011). The absorbance of 4-methyl-o-benzoquinone, resulting from the oxidation of 4-methylcatechol, a molecule similar and often compared to DOPA, is measured at 400 nm (Cabanes *et al*., 1987; Waite, 1985). While the absorbance of its oxidation product, Dopaquinone, is observed at 395 nm (Rodríguez-López *et al*., 1991). The activities of both monophenolase and diphenolase activity can also be determined by measuring the rate of dopachrome formation at 475 nm, which is the oxidized form of dopaquinone (Qin *et al*., 2014). These measurements allow for monitoring both monophenolase and diphenolase activities with various substrates.

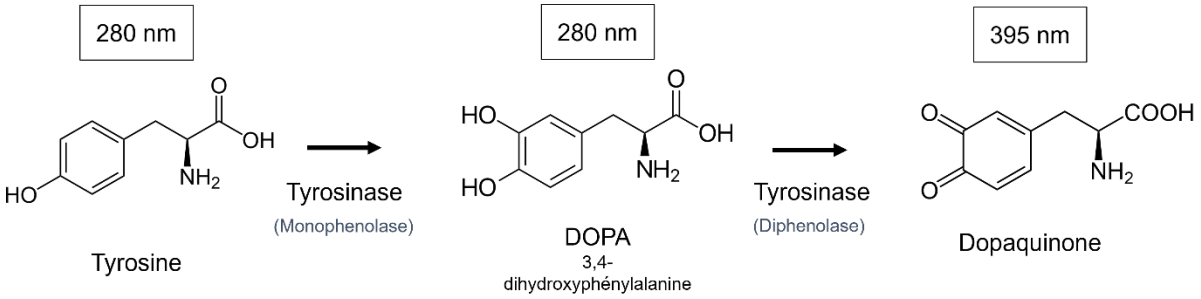


Figure 4. Reaction of the monophenolase and diphenolase activities of the tyrosinase enzyme with the absorbance (inspired by Agarwal *et al*., 2019).

Unfortunately, we were unable to test the catalytic activity of the recombinant tyrosinase as it precipitated during the refolding process. Results obtained with the commercial mushroom tyrosinase are summarized in the Supplementary materials (Supplementary Document 2).

4. Discussion

Tyrosinases are part of the type-3 copper protein family (Aguilera *et al*., 2013). These oxygen-transferring copper metalloproteins catalyze the *o*-hydroxylation of monophenols (e.g., tyrosine) into o-diphenols (e.g., DOPA) followed by the oxidation of *o*-diphenols to form *o*-quinones (Ullrich & Hofrichter, 2007; Aguilera *et al*., 2013). As a result, tyrosinase displays both cresolase activity (monophenol monooxygenase) and catecholase activity (catechol oxidase) (Ramsden & Riley, 2014). The architecture of the tyrosinase active site is very similar between the two tyrosinase functions (monophenolase and diphenolase activity), and three distinct oxy-, deoxy-, and meta-oxidation states can be distinguished, based on the bi-copper structure of the active site (Ramsden & Riley, 2014). However, it has been shown that there are some functional differences between tyrosinase enzymes, e.g. some of them appear to show only diphenolase activity (Pretzler & Rompel, 2018), and variations exist in their sequences, size, glycosylation, and activation (Jaenicke & Decker, 2003).

Tyrosinase plays a crucial role in the proper functioning of the adhesive system in *M. edulis* by allowing the production of DOPA from tyrosine residues in mussel foot proteins. Mussels being model organisms for the design of bio-inspired underwater adhesives, many of their adhesive proteins (mfps) have been produced recombinantly in bacteria or yeasts (Wang and Scheibel, 2018). Although they can produce large quantities of mfps, these heterologous hosts cannot perform the post-translational modifications. Because of the important roles of DOPA in mussel adhesion, many studies have therefore used mushroom tyrosinase for the *in vitro* post-translational conversion of tyrosine to DOPA in their recombinant mussel foot

protein with variable yields of conversion (e.g. in Kitamura *et al*., 1999; Hwang *et al*., 2005; Lee *et al*., 2008; Pilakka Veedu *et al*., 2023). Due to the differences between tyrosinase enzymes, as previously mentioned, it would be preferable to use a mussel foot-specific tyrosinase to improve the conversion of tyrosine to DOPA in the recombinant mussel foot proteins.

The study of tyrosinases in mussels has progressed significantly in these last decades. Waite (1985) and Hellio *et al*. (2000) purified and characterized tyrosinases with a catechol oxidase activity from the foot and byssus of *Mytilus edulis*. Recent omics studies have identified additional tyrosinase isoforms: Guerette *et al*. (2013) found five isoforms in the foot of *Perna viridis* (Guerette *et al*., 2013), and Qin *et al*. (2016) discovered six isoforms in the byssus of *Mytilus coruscus* (Qin *et al*., 2016). Wang *et al*. (2019) were able to recombinantly produce a tyrosinase (called polyphenol oxidase-like protein (PPOL) in their study) from the species *Mytilus galloprovincialis* in *E. coli*, and its enzymatic activity was tested, showing no mono- or diphenolase activity but rather a suppression of DOPA oxidation (Wang *et al*., 2019). In our previous work, we identified several tyrosinase transcripts potentially involved in the maturation of mussel foot proteins in *M. edulis* (Chapter I). This large diversity of tyrosinases probably presumably reflects specialized roles or substrates depending on their distribution in the foot or byssus. One of these candidates, Medu-TYR12, was found to be expressed only in the plaque gland and was detected in the proteome of the vesicles extracted from the plaque gland. Medu-TYR12 is a close homolog of a tyrosinase-like protein identified in the byssal plaque of *M. coruscus* (GenBank number KP876481; Qin *et al*., 2016). The plaque gland is known to contain more than ten different mussel foot proteins (DeMartini *et al*., 2017), including mfp-3 and -5 which are the most DOPA-rich byssal proteins (Papov *et al*., 1995; Waite & Qin, 2001). It was therefore proposed that Medu-TYR12 may have monophenolase activity (Chapter I) because the DOPA-based adhesion and crosslinking pathways in the plaque are achieved by the formation of metal coordination bonds under reducing conditions to prevent the oxidation of DOPA to dopaquinone. The formation of dopaquinone, on the other hand, is a prerequisite for the formation of covalent cross-links in the threads (Priemel *et al*., 2020). Medu-TYR12 was therefore selected for recombinant protein production to better understand its function. The catalytic activity of this recombinant protein can be assessed by UV-vis measurements. These measurements are useful for assessing enzymatic activity, as they reveal an increase in absorbance correlated with the conversion of substrate to product. As the enzyme catalyzes this transformation, the amount of product in the solution increases, resulting in

higher absorbance at a specific wavelength (Lerch & Ettlinger, 1972). However, rMT12 precipitated during the refolding process, and Western blot analysis revealed that it was fragmented during its production in *E. coli*. It was therefore not possible to obtain a soluble fraction that could be used in enzymatic activity tests.

In general, the production of recombinant proteins using bacterial strains is rapid because bacteria divide every 20 min under optimized conditions (Clark & Maaløe, 1967). It is also high yielding, economical and has become an essential aspect of the 21st century biotechnology industry (Mahmoud, 2007). Future efforts might enable expression of folded soluble tyrosinase, but this requires further troubleshooting. Before exploring how to improve rMT12 production, it is important to note that the choice of bacterial strain for expression is crucial. For example, the low or absent expression of a protein in the BL21(DE3) strain may result from various factors, including differences in metabolic compartmentalization and environmental conditions between the source organism (in this case, the blue mussel) and the bacterial host. Additional factors could include the lack of appropriate chaperones for protein refolding, post-translational modifications that do not occur in bacteria, or the formation of disulfide bridges (whose presence has been suggested in the Medu-TYR12 sequence because it contains several cysteines) (Duong *et al*., 2014). Other bacterial strains could therefore also be tested. For example, the Origami cell line may be used to produce disulfide bond-containing proteins as their cytoplasm is sufficiently oxidizing to allow the efficient formation of native disulphide bonds (Bessette *et al*., 1999; Dutta *et al*., 2001). Non-bacterial expression systems such as *P. pastoris*, baculovirus/insect cells or mammalian systems, which have been shown to promote correct protein folding and post-translational modifications, may also be considered. (Brondyk, 2009).

We encountered several issues during the production of rMT12. Firstly, we found evidence of protein degradation upon expression. To address this problem, protease inhibitors could be added to the LB culture medium and the lysis buffer (here NaCl-Tris-lysozyme buffer) to prevent degradation (Duong *et al*., 2014). rMT12 formed inclusion bodies (IBs), meaning that rMT12 aggregates and accumulates in cytosol of the bacteria. Protein characteristics, overexpression, and environmental conditions are some of the factors that can lead to the formation of IBs (Bhatwa *et al*., 2021). Several approaches can reduce IB formation, such as lowering the growth rate of cells by decreasing incubation temperature below 37 °C (Cabilly, 1989), briefly increasing the temperature under specific conditions to enhance protein solubility (Oganesyan *et al*., 2007), or optimizing the IPTG induction process (Donovan *et al*.,

1996). Tandem mass spectrometry analysis could also be carried out on the different bands observed in the SDS-PAGE gel to confirm and see which part of the rMT12 is degraded.

Another issue encountered during rMT12 production was protein precipitation during the refolding process from inclusion bodies, caused by protein aggregation. In this case, the refolding was performed by dialysis, but other methods could have been used, such as chromatographic (Machold *et al*., 2005) or non-chromatographic strategies (Gautam *et al*., 2012). Some compounds can be added to stabilize the native protein state, such as protein stabilizers. Glycerol is a protein stabilizer that was used in this study and enhances hydrophobic interactions (Yamaguchi & Miyazaki, 2014). But other compounds could have been used like sucrose (Akbari *et al*., 2010) or polyethylene glycol (Nian *et al*., 2009). Additionally, protein stabilizers can be combined with aggregation inhibitors like arginine (Rudolph & Fischer, 1990; Reddy K. *et al*., 2009), proline (Samuel *et al*., 2000), or glycine amide (Ito *et al*., 2011), which act as moderate chaotropic agents by forming weaker hydrogen bonds with the denatured protein and water, thereby improving protein solubility and inhibiting aggregation (Yamaguchi & Miyazaki, 2014).

5. Conclusion

In this study, we successfully induced *E. coli* to produce a recombinant plaque gland-specific tyrosinase identified in the blue mussel, *M. edulis*. However, rMT12 was degraded during the expression process and precipitation occurred upon refolding. Herein, we propose several strategies to optimize future rMT12 expression and refolding. Given that different tyrosinases appear to be involved at various compartments and stages of byssus production, our Medu-TYR12 merits further investigation. The success of this technique could lead to a better understanding of DOPA synthesis in the blue mussel plaque gland and how this modified amino acid is regulated during the formation of the plaque. These findings could inspire the design of new DOPA-based adhesion materials and facilitate the production of DOPA in recombinantly produced mussel foot proteins.

Chapter II

Supplementary materials for

The Role of Tyrosinase in Mussel Adhesion: Characterization of a Recombinant Mussel Plaque-Specific Tyrosinase

Emilie Duthoo[1], Elise Hennebert[2], Mathieu Rivera[3], Alessandra Whaite[1], Matthew J. Harrington[3] and Patrick Flammang[1]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cell Biology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[3] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

This part includes:

Supplementary Document 1 and 2

Supplementary Document 1

Calculation of the enzymatic activity

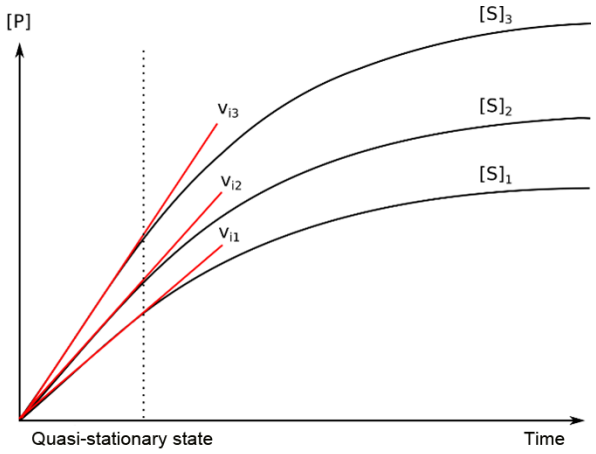
1. Use the Beer-Lambert law to calculate the concentration of product at each time point using the following equation:

$$A = ecl$$

where A is the absorbance of the product at the specified wavelength, e is the molar extinction coefficient of the product (3700 cm^{-1}), c is the concentration of the product in moles per liter, and l is the path length of the cuvette (usually 1 cm).

Note: To calculate the concentration of product in millimoles per liter (mM), multiply the concentration in moles per liter by 1000.

2. Plot the concentration of product (in mM) versus time (in min) to generate a standard curve.

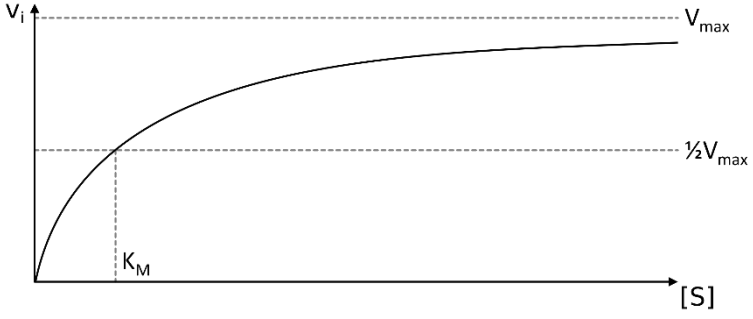


3. Calculate the initial velocity (Vi) of the reaction by determining the slope of the tangent line to the standard curve at time zero (t = 0).

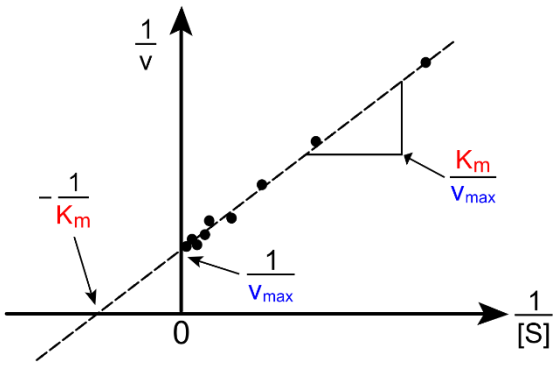
Note: Vi is the rate of the reaction at the beginning of the reaction, when the substrate concentration is highest, and the enzyme is not yet saturated.

4. Plot Vi (in mM/min) versus substrate concentration ([S]) (in mM) to generate a Michaelis-Menten plot.
5. Determine the KM and Vmax values of the enzyme from the Michaelis-Menten plot.

Note: KM is the substrate concentration at which the reaction rate is half of Vmax, and Vmax is the maximum velocity of the reaction when the enzyme is saturated with substrate.



6. To determine more precisely the enzyme kinetic parameters, plot 1/Vi against 1/[S].



This plot will generate a straight line in which you can calculate Km/Vmax.

Chapter II

Supplementary Document 2

Before I could test the enzymatic activity of rMT12, I carried out this experiment with mushroom tyrosinase.

Materials and Methods

Mushroom tyrosinase (6540 units/mg, Sigma) was added to different substrates: 10 mM L-Tyrosine (cell culture reagent powder, Alfa Aesar), 10 mM L-DOPA, 10 mM 4-methylcatechol, and 0.5 mg of recombinant Mfp-1 protein (Mesko *et al*., 2021) (provided by the Harrington lab, McGill University, Montréal, Canada). These substrates were diluted in a phosphate buffer at different pH levels. The tyrosinase was added at a concentration of 250 units/µmol of tyrosine in rMfp-1. The different solutions were placed in 96-well plates with a flat, transparent bottom. Absorbance was measured using a UV-Vis spectrophotometer plate reader (Tecan Life Sciences) for 120 min, with three measurements taken every 15 min at room temperature.

Results and discussion

Mushroom tyrosinase (6540 units/mg, Sigma) was added to different substrates: 10 mM of L-Tyrosine (cell culture reagent powder, Alfa Aesar), 10 mM of L-DOPA, 10 mM of 4-methylcatechol, and 0.5 mg of rMfp-1, which were diluted in a phosphate buffer at different pH. The amount of tyrosinase added was at a concentration of 250 units/µmol of tyrosine in rMpf-1. These different solutions were placed on 96-well plates with a flat transparent bottom. The absorbance was measured using a plate reader UV-Vis spectrophotometer (Tecan Life Sciences) for 120 min, with three measurements taken every 15 min at room temperature.

We examined the absorbance changes at different pH levels for four distinct substrates: tyrosine, DOPA, 4-methylcatechol, and rMfp-1. The blank corresponded to the phosphate buffer used for diluting all the substrates. The initial absorbance measured at 280 nm corresponded to both tyrosine and DOPA. At pH 2, no changes in absorbance were observed. However, starting from pH 4, the absorbance significantly increased, surpassing the spectrophotometer's maximum absorbance limit of 4 for most substrates. Regarding the absorbance linked to the diphenolase catalytic activity, no noteworthy change was observed at pH 2. However, from pH 4, the absorbance exhibited a slight increase for all substrates, except for rMfp-1, where the absorbance began to rise at pH 6, following a similar trend to that of tyrosine. At pH 8, rMfp-1 and 4-methylcatechol followed the same trend, suggesting that the enzyme reacts similarly with both substrates (Fig. 1).

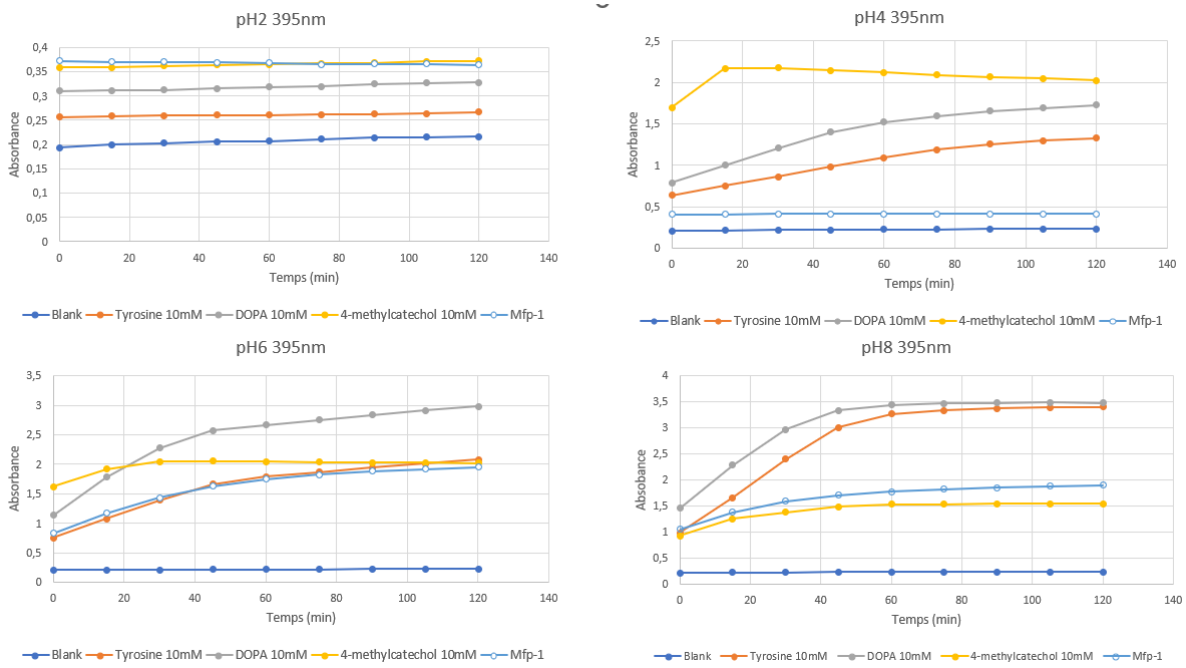


Figure 1. Absorbance spectrum at 395 nm of L-dopa, L-Tyrosine, 4-methylcatechol and rMfp-1 in phosphate buffer from pH 2 to 8 for 2 hours.

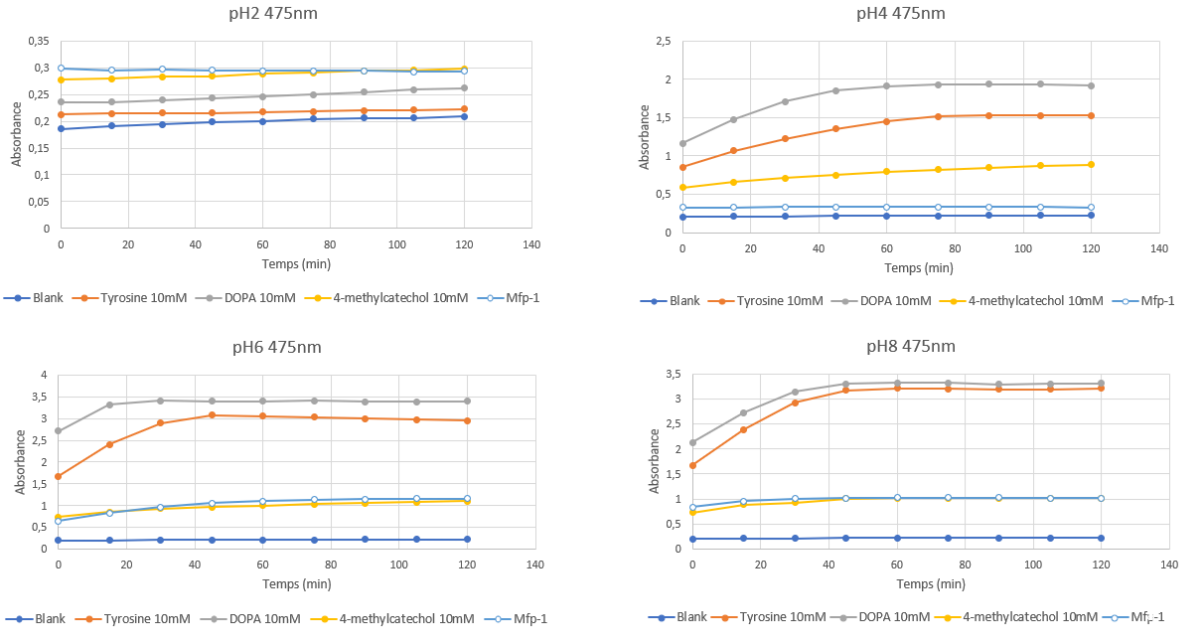


Figure 2. Absorbance spectrum at 475 nm of L-dopa, L-Tyrosine, 4-methylcatechol and rMfp-1 in phosphate buffer from pH 2 to 8 for 2 hours.

The absorbance at 475 nm is often used to study the enzymatic activity of tyrosinase because it corresponds to the final product of its activity, the dopachrome. At pH 2, there is no change in absorbance, demonstrating that a low pH inhibits enzyme activity. At pH 4, a change in absorbance was observed for all substrates except rMfp-1. At pH 6 and 8, all the substrates showed an increase in absorbance, particularly tyrosine and DOPA, whose absorbance almost

doubled compared with pH 4. In addition, rMfp-1 followed the same trend as the substrate 4-methylcatechol (Fig. 2). These tests were supposed to be redone with rMT12, as well as measuring the absorbance of DOPA and tyrosine alone, since they also show an increase in absorbance without the addition of enzymes, which was observable by their color change over time (results not shown).

Acknowledgements

I would like to express my gratitude to the members of the Harrington lab for their invaluable assistance during these experiments. A special thank you goes to Professor Harrington and Dr. Hamideh Rezvani Alanagh for providing invaluable guidance and mentorship. I would also like to thank the Thibodeaux lab (Chemistry Department, McGill University) for providing access to their materials for recombinant protein production and the Dorval lab (Faculty of Engineering, McGill University) for the use of their plate reader spectrophotometer. Additionally, I am deeply appreciative of the generous support from the FNRS, who granted me a 'stay abroad' [V3] scholarship to carry out these tests in the Harrington lab at McGill University (Montréal, Canada).

Chapter II

Chapter III

Characterization of the adhesive system in the honeycomb worm, a marine tube building polychaete

Emilie Duthoo, Carla Pugliese, Leandro Smacchia, Aurélie Lambert, Jean-Marc Baele, Ruddy Wattiez, Matthew J. Harrington and Patrick Flammang

Contribution: Carla Pugliese and I carried out the extraction of granules from the cement gland and the proteomic experiment. Leandro Smacchia and I analyzed and interpreted the proteomic analyses. The elemental composition analyzes were carried out by Carla Pugliese and myself, and the ultrastructure study was carried out by Aurélie Lambert. I did the *in silico* analyses and performed the *in situ* hybridization experiments for the FAM20C and adhesive protein candidates. Patrick Flammang, Aurélie Lambert and I interpreted the results of the study.

Chapter III

Characterization of the adhesive system in the honeycomb worm, a marine tube building polychaete

Emilie Duthoo[1], Carla Pugliese[1], Leandro Smacchia[1], Aurélie Lambert[1], Jean-Marc Baele[2], Ruddy Wattiez[3], Matthew J. Harrington[4] and Patrick Flammang[1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Department of Geology, Faculty of Engineering, University of Mons, 7000 Mons, Belgium

[3] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[4] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

Abstract

Adhesives produced by marine organisms offer remarkable performance and serve as a major source of inspiration for developing biomimetic adhesives. However, a thorough understanding of their composition and operating mechanism is essential for advancing such applications. The sandcastle worm, *Phragmatopoma californica*, is a model species in bioadhesion research and its adhesive secretion has been characterized in several studies. In contrast, little is known about the adhesive composition in the closely related European species, *Sabellaria alveolata*, although several differences have already been pointed out between these two species. This study aims to investigate the adhesive system of *S. alveolata* through proteomic analyses of granules extracted from its cement glands, along with morphological and elemental composition studies. Phosphoserine has been identified as one of the main modified amino acids in tubeworm cement. Here we hypothesized that the phosphorylation of serine residues in adhesive proteins could be catalyzed by Fam20C kinase enzymes. We identified and characterized these enzymes through *in silico* analyses and confirmed their expression in the adhesive glands via *in situ* hybridization experiments.

1. Introduction

Many invertebrate marine organisms have adhesive mechanisms that allow them to firmly attach to various substrates in a wet and salty environment (Hennebert *et al*., 2015a; Delroisse *et al*., 2023). This remarkable ability has raised the interest of scientists in developing bio-inspired underwater adhesive materials for various applications, particularly in the industrial and biomedical fields (Modaresifar *et al*., 2016; Almeida *et al.*, 2020). Tubeworms of the family Sabellariidae are one of the model organisms that have been studied extensively for their adhesion, notably the species *Phragmatopoma californica* Krøyer in Mörch, 1863 and *Sabellaria alveolata* (Linnaeus, 1767). Within a settlement, the tubes of different individuals are joined together, resulting in reef-like formations. These impressive biogenic constructions, such as the largest reef in Europe located in the bay of Mont Saint-Michel in France, have fascinated researchers since the 18th century (Réaumur, 1711; Noernberg *et al*., 2010).

S. alveolata is a gregarious species found on the west coast of Europe (Gruet, 1972). This worm builds its tube using a building organ, located near the mouth, that manipulates and binds together sand grains, shell fragments, or sea urchin spine fragments with a strong proteinaceous cement (Hennebert *et al*., 2015b). Within the tube wall, particles are held together by several dots of cement (Vovelle, 1965). The building organ is the external part of an extended glandular system, and it encloses long cell processes originating from two types of cement glands located around the digestive tract and in the parapodia in the parathoracic part of the worm. These two gland cells can be distinguished by the morphology of their secretory granules: homogeneous or heterogeneous containing inclusions (Vovelle, 1965; Gruet *et al*., 1987). In the adhesive secretion, three adhesive proteins have been identified, Sa-1, -2, and -3, presenting highly repetitive and blocky primary structures and post-translationally modified amino acids (Zhao *et al*., 2005; Becker *et al*., 2012). Two of these modified amino acids, DOPA (3,4-dihydroxyphenylalanine) and phosphoserine (pSer), presumably play key roles in the proper functioning of the adhesive system as they contribute to both the adhesion and cohesion of the adhesive proteins (Zhao *et al*., 2005; Sun *et al*., 2007). The adhesive proteins are condensed into secretory granules through a process called complex coacervation, which involves the coalescence of oppositely charged proteins and divalent cations (Ca^{2+} and Mg^{2+}) (Stewart *et al*., 2004). Secretion is accompanied by a change in pH from 5 (pH of the cells) to 8 (pH of seawater), which can cause the bonding of the pSer with calcium cations to change from electrostatic interactions to stronger, more specific ionic interactions, forming intermolecular bridges (Sun *et al*., 2007; Tagliabracci *et al*., 2012). This modified amino acid

can also condense with histidine to form histidinoalanine cross-links through the loss of phosphate in the reaction (Flammang *et al*., 2009). This results in the adhesive secretion hardening into a cement dot (Flammang *et al*., 2009).

Despite the crucial role of pSer in the honeycomb worm adhesive system, the identity and function of the kinase involved in the maturation of adhesive proteins is not well understood. FAM20C is a secreted protein that has been identified as the kinase responsible for phosphorylating S-x-E/pS motifs within proteins in the secretory pathway (Tagliabracci *et al.*, 2012). This protein is part of the FAM20 family, which also includes FAM20A, FAM20B, and is found in both vertebrates and invertebrates with elevated protein sequence homology across different species (Tagliabracci *et al.*, 2013a; Tagliabracci *et al.*, 2013b; Du *et al*., 2018). Previous research has shown that FAM20C is involved in various biological processes, including myeloid differentiation in mouse hematopoietic stem cells and mineral formation (Nalbant *et al*., 2005; Hao *et al*., 2007), but also the formation of the pearl oyster shell (Du *et al*., 2018). This enzyme could therefore be a candidate kinase for the modification of adhesive proteins in *S. alveolata*.

This study aims at a better characterization of the adhesive system of *S. alveolata* thanks to the morphological and molecular characterization of the two types of granules present in the cement glands. Its goal is also to address the gap in knowledge about adhesive protein maturation by identifying, characterizing, and localizing putative kinases using *in silico* analyses and *in situ* hybridization techniques. The results may provide new insights into the composition and biosynthesis of the adhesive secretion, which is crucial to the honeycomb worm's survival.

2. Methods

2.1. Collection of honeycomb worms and samples preparation

Fragments of reefs of *Sabellaria alveolata* were obtained from the Biological Sample Collection Service of the “Station Biologique de Roscoff” in Brittany, France. Animals were transported to the laboratory of Biology of Marine Organisms and Biomimetics (University of Mons, Belgium), where they were kept in a re-circulating aquarium chilled at 13 °C and filled with artificial seawater of 33 psu salinity.

Individual tubes containing a worm were isolated from the reef fragment and placed on a petri dish. The distal third of each tube was then sectioned, fixed with 4% paraformaldehyde,

rinsed and air-dried. Worms were left in the remaining proximal part of the tube and were able to reconstruct this missing part of the tube with glass beads (425-600 µm in diameter; Sigma) (see Supplementary Fig. 1) (Jensen & Morse, 1988).

A few individuals were retrieved from their tubes and their anterior part was dissected and fixed in a 4% paraformaldehyde solution in phosphate-buffered saline (pH 7.4). The samples were then dehydrated through graded ethanol series and embedded in paraffin wax. Sections of 14 µm in thickness were cut with a Microm HM 340 E microtome and collected on Superfrost Ultra Plus (Thermo Scientific) microscope slides using a Milli-Q water drop.

2.2. Transmission Electron Microscopy

The anterior parts of *S.alveolata* and small fragments of tube reconstructed with glass beads were fixed for 3 hours at 4 °C in a solution of 3% glutaraldehyde in cacodylate buffer (0.1 M, pH 7.8; osmolarity adjusted to 1030 mOsm/l with NaCl). They were then rinsed three times for 10 min in a solution of cacodylate buffer (0.2 M, pH 7.8, adjusted to 1030 mOsm/l), and post-fixed for 1 hour in 1% osmium tetroxide in cacodylate buffer (0.1 M, pH 7.8, adjusted to 1030 mOsm/l) in the dark. After a final rinse in cacodylate buffer, the glass beads tubes were decalcified for 24 hours in a 10% EDTA solution (pH ~8). All the samples were then dehydrated in a series of ethanol baths of increasing strength (25%, 50%, 70%, 90%, 100%) and embedded in Spurr resin. Semi-thin sections of 1 µm thickness were cut using a Reichert Om U2 ultramicrotome. Ultra-thin sections 70 nm thick were then obtained using a Leica Ultracut UCT ultramicrotome fitted with a diamond knife. These sections were contrasted with uranyl acetate and lead citrate and observed using a Zeiss LEO 906E transmission electron microscope.

2.3. Scanning electron microscopy and elemental composition analyses

To observe the cement dots binding mineral particles together, the air-dried tube fragments were impregnated with epoxy resin, cured for 1 hour at 80 °C and polished. The epoxy resin embedding technique provided excellent preservation of the cement dot structure and allowed us to examine them with high resolution using SEM (JEOL JSM-7200F). Honeycomb worms embedded in Spurr resin (TEM samples) were used for the observation of cement gland secretory granules. The SEM images were acquired in low vacuum mode (50 Pa), with the backscattered electron detector at 15 kV with a 10 mm working distance. A transverse section through the parathoracic part of a paraformaldehyde-fixed worm was also observed. It was first dried by the critical point method and coated with gold-palladium in a sputter coater.

The elemental composition of cement dots from tubes embedded in epoxy resin, of both types of cement glands from worms embedded in Spurr resin, as well of freshly extracted granules (see below) was analyzed using energy dispersive X-ray spectroscopy (EDX). The EDX data were acquired with an 80 mm^2 silicon drift detector with a 10 s live time acquisition at approximately 30–40% dead time, using the Aztec software (Oxford Instrument). Specifically, X-ray microanalysis and elemental mapping of the general composition were produced for each structure analyzed.

2.4. Extraction of the granules present in the cement glands

Seven honeycomb worms were anesthetized for 20 min in a solution of 7.5% $MgCl_2$ (Messenger *et al.*, 1985) mixed with filtered seawater before the parathoracic part was dissected. The extraction of the granules present in the cement glands of *S. alveolata* was done by dissecting the parathoracic parts and, after rinsing with filtered artificial seawater, pressing them through a 38 µm-meshed filter. The resulting cell material was either used directly for protein extraction (see below) or was fixed in 4% paraformaldehyde in artificial filtered seawater for one hour at 4 °C before being centrifuged at 5000 x *g* for 10 min and resuspended with filtered seawater three times to remove the PAF solution (protocol modified from Wang & Stewart, 2013).

2.5. Protein extraction from the cement granules and mass spectrometry analysis

The freshly extracted granules were centrifuged at 18,000 x *g* for 15 min to concentrate them into a pellet. The pellet was resuspended in a solution of 5% acetic acid and 8 M urea and left for one hour (Rzepecki *et al.*, 1992). Ultrasonication (Amplitude 20X, cycle 1) for 3X10 sec was done followed by centrifugation for 15 min at 18,000 x *g*. The protein concentration in the supernatant was estimated using a Bradford assay with Bovine Gamma Globuline (BGG) as a standard and adjusted to 50 µg/µl for further processing. The samples were reduced by adding dithioerythritol (DTE) at a final concentration of 12.5 mM (1 µl/20 µl sample) for 25 min at 56 °C under agitation The samples were then alkylated with 25 mM iodoacetamide (1 µl/20 µl sample) for 30 min in the dark at room temperature. Proteins were precipitated using glacial acetone for 4 hours at -20 °C and centrifuged at 4 °C for 20 min at 18,000 x *g*. The pellet was resuspended with a solution of trypsin (Promega) at a concentration of 2 µg/20 µl in 50 mM NH_4HCO_3 and left at 37 °C overnight. The trypsinolysis was stopped by adding 0.5% formic acid. Samples were centrifuged for 15 min at 11,000 x *g* at 4 °C and the supernatant was mixed with loading buffer (2% Acetonitrile (ACN) + 0.1% formic acid) for injection.

The peptides were analysed by liquid chromatography (LC) connected to a hybrid quadrupole time-of-flight TripleTOF 6600 mass spectrometer (AB SCIEX, Concord, ON). MS/MS data were searched for protein candidates against a database consisting of the six open reading frames (ORFs) from the transcriptome of the anterior part of *Sabellaria alveolata* (NCBI SRA: SRR29446350) using the Protein Pilot software (version 5.0.1). Samples with a false discovery rate (FDR) exceeding 1.0% were excluded from subsequent analyses.

2.6. *In silico* analyses

First, the different cement precursor proteins from *S. alveolata* present on NCBI (Accession number from HE599563 to HE599646) were searched for in the proteomic results.

Then, the sequences of each protein identified in the proteomic analysis were retrieved and translated from the transcriptome of *S. alveolata* anterior part and analyzed using InterProScan (version 5.48-83.0) to determine their protein domain composition (Jones *et al*., 2014). Only proteins identified with a number of peptides superior or equal to two were analyzed. Sequences were then assigned to a protein category based on their domain composition. To do so, a keyword-based research was conducted on gross InterProScan results. 44 keywords (Supplementary Document 1) were used in the data mining process in order to classify most important domains into six functional categories: structural proteins; calcium-binding proteins; glycoproteins; protective proteins; enzymes; putative adhesion. Some domains may be classified into multiple functional categories. Based on domain classification, analyzed proteins were classified into the same six functional categories. Some sequences may therefore be classified into multiple functional categories. This reflects reality as proteins often have two or more functional units that may have a totally different role (Forslund *et al.*, 2019). Additionally, for three categories (structural proteins, protective proteins and putative adhesion), proteins were classified into subcategories, corresponding to keywords that were used to mine the original data. Barplots were then computed using programming language R (version 3.5.2) and packages ggplot2 (version 3.3.5) and dplyr (version 1.0.6).

In parallel to functional domain classification, differential expression data were retrieved from a transcriptomic analysis conducted by Buffet *et al*. (2018). Using sequence correspondence via BLASTn analysis, these data were crossed with functional domain classification of the proteins identified in the extracted granules.

2.7. Selection and characterization of the FAM20C transcript candidates

Local tBLASTn searches were performed in the transcriptome of *S. alveolata* anterior part using different Fam20C sequences retrieved from the NCBI database (NCBI Database numbers: AVI57681.1 (*Pinctada fucata*), XP_033744735.1 (*Pecten maximus*), CAD7192288.1 (*Sepia pharaonic*), XP_035824787.1 (*Aplysia californica*), Q5MJS3.1 (*Mus musculus)* and Q8IXL6.2 (*Homo sapiens*)) (retrieved in January 2021). Five transcripts showing the highest similarity with the reference sequences were selected.

Each sequence was then translated *in silico* and analyzed by looking at their computed parameters including the molecular weight, theoretical pI and amino acid composition using the ProtParam tool (https://web.expasy.org/protparam/). Conserved domains were also searched using NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Lu *et al*., 2020) or InterPro (v75.0) (Paysan-Lafosse *et al*., 2023). The presence of a signal peptide was predicted using the online SignalP 6.0 tool (https://services.healthtech.dtu.dk/service.php?SignalP). Finally, homology against known nucleotide sequences was assessed using NCBI Basic Local Alignment Search Tool (tBLASTn) to confirm these sequences as FAM20C enzymes.

2.8. Total RNA extraction and cDNA construction

Total RNA was extracted from different parts of three honeycomb worms using TRIzol™ Reagent kit (Thermofisher). The parts selected were the operculum, the parathoracic part, the abdominal part and the caudal part. The concentration and purity of the extracted RNA was measured with a UV-Vis spectrophotometer (DENOVIX DS-11). A cDNA library was synthesized from the RNA extracted by reverse transcription-PCR using the Reverse transcription kit (Roche).

2.9. Amplification by PCR

Double-stranded DNA templates were amplified by PCR using the Q5 High-Fidelity DNA Polymerase kit method (New England BioLabs), with primer designed by Open Primer 3 (bioinfo.ut.ee/primer3/) with an optimal amplicon length between 700 and 900 bp. For the adhesive proteins, the primers were designed using the first clone of each cement precursor protein available on NCBI (Becker *et al*., 2012). For *in situ* hybridization probe synthesis, a second PCR was done with a T7 promoter binding site (5'-GGATCCTAATACGACTCACTATAGG-3') added to reverse strand PCR primers. After quality and size check by gel electrophoresis, PCR products were purified using the Wizard

SV Gel and PCR clean-up system kit (Promega). The purified products were used for RNA probe synthesis after sequencing to check if the amplified sequence corresponds to the desired transcript.

2.10. Localization of the candidates using *in situ* hybridization

Anti-sens digoxigenin (DIG)-labelled RNA probes were synthesized with DIG RNA Labelling Kit (Roche) with T7 RNA polymerase and DIG–dUTP. *In situ* hybridization was performed according to the protocol of Lengerer *et al*., 2019. The RNA probes were used at a concentration of 0.2 ng/µl on 14 µm thick frontal sections of *S. alveolata*, and detected with antidigoxigenin-AP Fab fragments (Roche) at a dilution of 1: 2000. The signal was developed using the NBT/BCIP substrate (Roche) at 37 °C. The sections were observed with a 100 X objective to distinguish both types of cement glands based on their secretory granule morphology, and images were taken with a Zeiss Axio Scope A1 microscope.

3. Results

The cement produced by *S. alveolata* is often compared to that of *P. californica* (Fewkes, 1889), a closely related species that inhabits the west coast of North America (Abbott & Reish, 1980). However, differences between these two species have been noticed in some articles. For instance, there are limited percentages of identity in the alignment of their adhesive proteins, and the sulphated polysaccharides described in *P. californica* seem to be absent in *S. alveolata* (Hennebert *et al*., 2015b; Hennebert *et al*., 2018). These observations raise questions about whether the adhesive secretion of *S. alveolata* differs from that of *P. californica*, and if the elemental composition of their biocement is the same.

3.1. Morphology and ultrastructure of the adhesive glands and cement

We first analyzed the ultrastructure of both types of adhesive gland granules using transmission electron microscopy. Both types of granules have a size between 2.5 and 4 µm in diameter. Homogeneous granules have a uniform electron density with no internal structure (Fig. 1A). In contrast, heterogeneous granules contain inclusions of various shapes within a matrix that is less electron-dense and resembles the contents of homogeneous granules (Fig. 1B). These inclusions, which vary in size from 100-1500 nm, appear as spherical to elliptical inclusions and are made up of electron-dense concentric lamellate layers (Fig. 1D). Some of them, particularly the larger ones, show an apparently empty cavity in their centre.

A decalcified adhesive cement dot that held two glass beads together was also observed in TEM (Fig. 1E). The cement dot matrix is homogeneous and electron-dense. It encloses hollow spheroids of various sizes, as well as small electron-dense granules and small lacunae (Fig. 1F). The exception is the periphery of the adhesive dot, which is made entirely of the matrix, giving it a smooth appearance. The hollow spheroids, measuring about 0.3-6.8 μm in diameter, often appear empty at their centres. However, some of them are partially or totally filled with a very electron-dense material. The cortex surrounding these spheroids is electron-dense and possesses a concentric lamellar structure. The thickness of this cortex also seems to increase with the spheroid size and can measure up to 400 nm in thickness. The sizes of the electron-dense granules and lacunae are 50-700 nm and 50-1400 nm in diameter, respectively. They are homogeneously distributed in the matrix between the hollow spheroids (Fig. 1F).

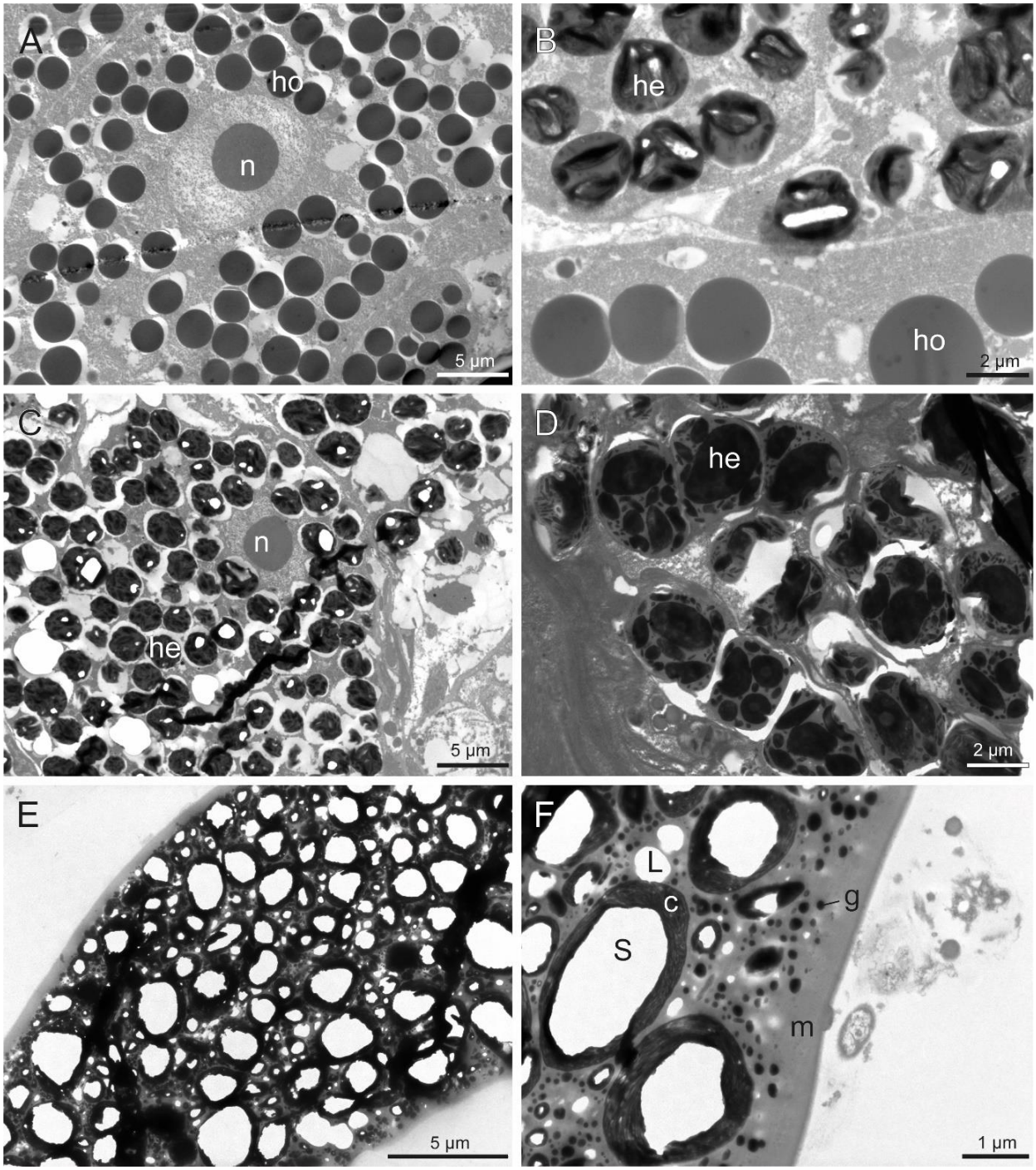


Figure 1. TEM images of the two types of adhesive gland: the cement glands with homogeneous granules (A, B) and the cement glands with heterogeneous granules (C, D). TEM image of a dot of cement holding two glass beads together (E, F). Legend: c – cortex; g - electron-dense granule; he - heterogeneous granules; ho – homogeneous granules; L – lacunae; m – matrix; n – nucleus.

3.2. Elemental composition of the adhesive secretion

To investigate the elemental composition of cement gland granules, we used energy-dispersive X-ray spectroscopy (EDX) coupled with scanning electron microscopy (SEM). In *P. californica*, the inclusions of the heterogeneous granules are mainly composed of hyperphosphorylated adhesive proteins Pc-3A and -3B with approximately 0.5 Mg^{2+} ion per phosphate group to balance the charges (Stewart *et al*., 2017). In *S. alveolata*, the elemental composition was measured on 4 secretory granules of both types of glands in the parathoracic part of worms embedded in Spurr resin. Using the backscattered electron detector, the heterogeneous granules could be distinguished by their appearance, as they appear lighter than the homogeneous granules due to their composition. The heterogeneous granules exhibited high concentrations of phosphorus (7.25% ± 1.23), magnesium (2.35% ± 0.95), and sodium (2.55% ± 0.52) (Fig. 2A) (Supplementary Table 1). On the other hand, the homogeneous granules (Fig. 2B) presented much smaller quantities of these elements: 1.7% ± 0.36 for phosphorus, 1.125% ± 0.09 for sodium, 0.575 % ± 0.15 for magnesium, and no detectable amount of calcium (Fig. 2B) (Supplementary Table 2).

We also conducted an elemental analysis on six cement dots embedded in epoxy resin (Fig. 2C). Our measurements revealed that, in addition to carbon and oxygen, the cement primarily consisted of calcium (15.54% ± 4.5) and phosphorus (3.74% ± 1.08). Additionally, we detected smaller amounts of sulphur (0.64% ± 0.15), magnesium (0.56% ± 0.18), and sodium (0.18% ± 0.04) (Supplementary Table 4). This finding aligns with previous studies that have emphasized the presence of calcium and magnesium in the structure of cement dots (Gruet *et al*., 1987; Sanfillipo *et al*., 2019). The elemental composition of the cement secreted by *P. californica* shows more magnesium and less calcium than what we measured in *S. alveolata* (Stewart *et al*., 2004).

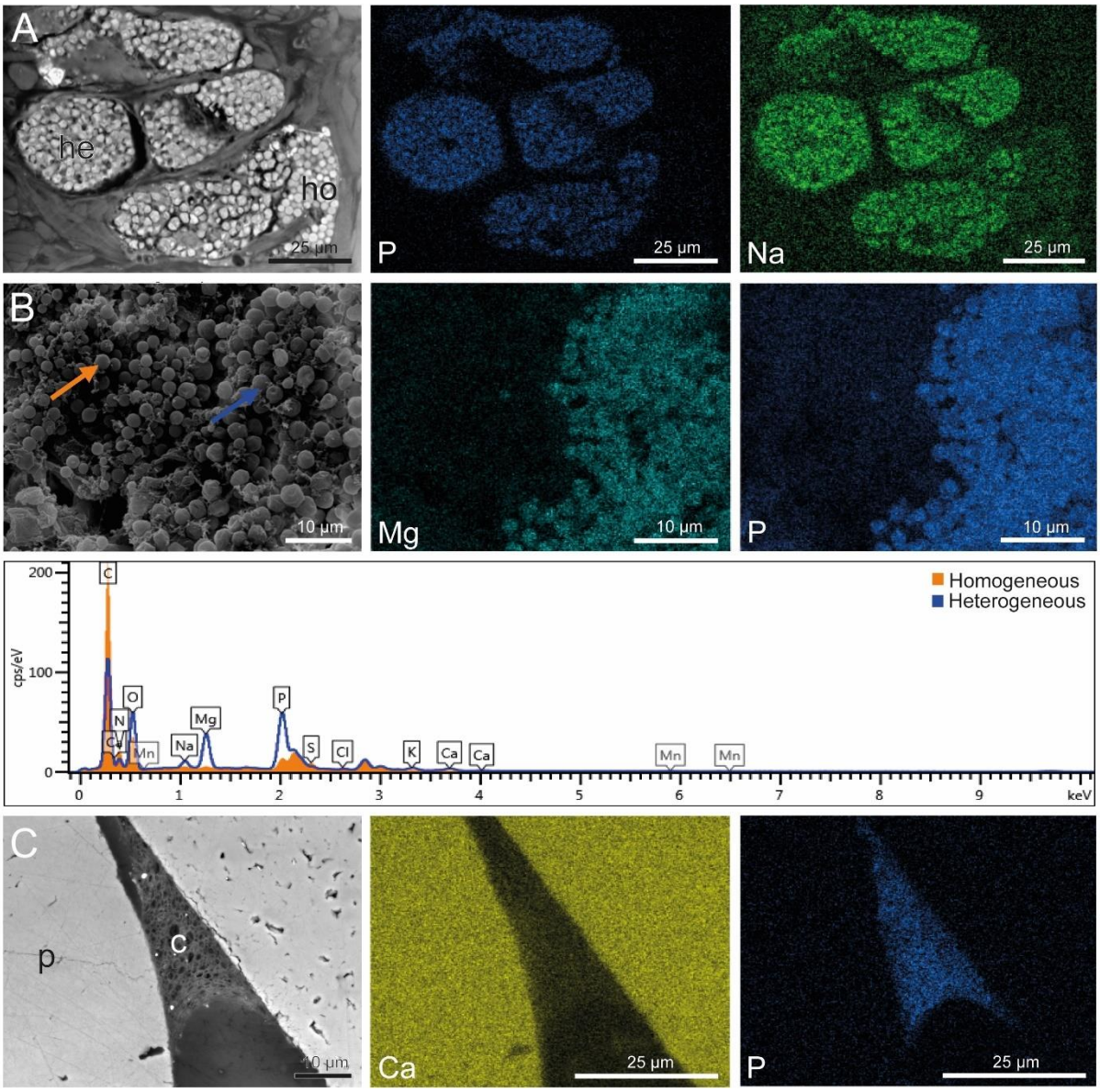


Figure 2. SEM image with composition mapping of the two types of granules in the cement glands of *S. alveolata* (A), with their EDX spectra (B), and of the two main elements (apart from C and O) of a cement dot sticking two particles together (C). Legend: c - cement; he – heterogeneous granules; ho – homogeneous granules; p – particles of the worm tube.

3.3. Microscopic analysis of secretory granules extracted from the honeycomb worm

Inspired by the work of Wang & Stewart (2013), we attempted to extract cement gland secretory granules from *S. alveolata*. The extracted granules were then observed using light and electron microscopy. Several vesicles with a diameter of 2 µm with two kinds of aspect were observed: vesicles with a smooth surface that could correspond to homogeneous granules and vesicles with a more granular surface that could correspond to heterogeneous granules. By observing their elemental composition using EDX, the elements phosphorus and magnesium were detected in some granules. We also stained the granules using different staining methods.

Heidenhain's azan, a general stain, turned the granules red, showing the presence of proteins. NBT glycinate and Arnow staining confirmed the presence of DOPA in the extracted granules by turning them in purple and red, respectively. All these results confirmed that we successfully extracted the granules from the two kinds of cement gland of *S. alveolata* (Fig. 3). However, the images also showed the presence of various other cell debris (Fig 3D, E).

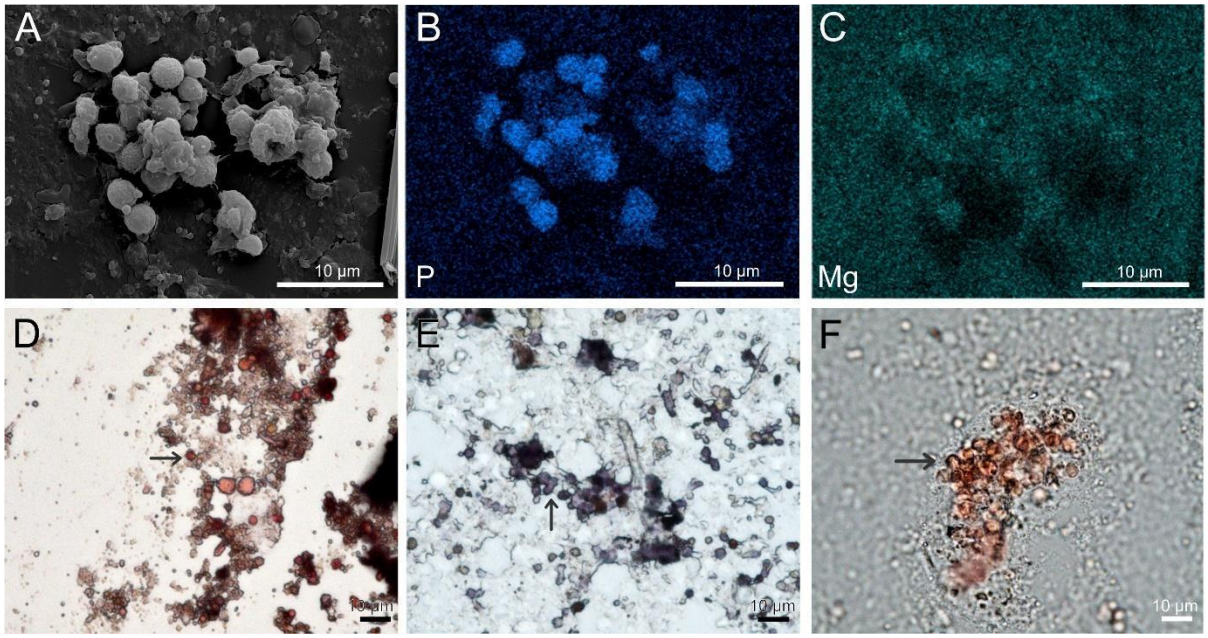


Figure 3. SEM image of a cluster of heterogeneous granules extracted from *Sabellaria alveolata* (A), with corresponding EDX images showing the presence of phosphorus (B) and magnesium (C). Heidenhain's Azan (D), NBT-glycinate (E), and Arnow (F) staining of the extracted granules confirmed the presence of proteins and DOPA (arrows indicate an extracted granule).

3.4. Proteomic analyses of the extracted granules

The proteins from the extracted granules were extracted using a buffer composed of 8M urea with 5% acetic acid, the same buffer used for extracting the DOPA-based adhesive byssus of mussels (Rzepecki *et al*., 1992). The proteins were digested using trypsin and analyzed with a TripleTOF mass spectrometer. By crossing the results with the transcriptome of the anterior part of *S. alveolata*, 2,685 proteins were identified. Among these proteins, two known adhesive proteins from *S. alveolata* were retrieved: Sa-1 (comp284608_c3) identified with 4 peptides and Sa-3B (comp268814_c4_seq3) identified with 2 peptides. The adhesive proteins of *S. alveolata* are known to be short and intrinsically disordered, with repeated peptide motifs often rich in lysine residues. This proteomic analysis was based on a trypsin digestion, which involves cleaving proteins at the C-terminal of arginine and lysine residues, may therefore have generated peptides too short to be analyzed.

For each sequence identified with more than 2 peptides, we examined the conserved domains present using the software InterProScan (Supplementary Table 4). The composition of the protein domains obtained from the contents of the granules allowed us to assume the functions of the proteins extracted from the granules, which may be multiple for each protein.

To analyze granule protein functional composition more systematically and more integratively, proteins were classified into six functional categories using a keyword-based research strategy on gross InterProScan results. 44 keywords (Supplementary Document 1) were used to classify 14,127 domains amongst the 24,783 identified by InterProScan. 808 domains were classified in two different categories, yielding an artificial 14,935 "classified domains". These functional domains led us to the classification of 835 proteins amongst the 1,125 granule proteins identified with a number of peptides superior or equal to two, 20 of which did not have any domain identified by InterProScan. For 270 proteins, domains have been identified but have not been classified with our sets of keywords. Out of the 835 classified proteins, 230 were assigned two or more categories, yielding an artificial 1,121 "classified proteins" (Fig. 4). Subcategory protein classification has also been conducted to add precision to structural proteins, protective proteins and putative adhesion categories (Fig. 4; Supplementary Document 2). Quantitative analysis shows a large part of enzymes and structural proteins in granule proteins, followed by a significant amount of calcium-binding-related peptides and proteins that could be linked to a potential role in adhesion.

Among the proteins identified, we found many structural proteins such as collagen, tropomyosin, myosin, actin, and troponin that link myosin and actin. Spectrin domains and several protein domains that participate in the anchoring or integrity of actin, such as ankyrin, filamin, and profilin, were also found in these proteins. Moreover, other protein components like tubulin and intermediate filaments were also identified, as well as proteins that participate in the anchoring of the cytoskeleton to the membrane, like FERM or PDZ-like domain-containing proteins. These types of proteins are often considered contaminants; however, they can also be associated with tension-bearing functions, making them candidate proteins as cohesive elements in the adhesive matrix (Galli *et al*., 2005; Santos *et al*., 2009b).

Several proteins involved in calcium binding processes were also identified in this proteomic analysis. These proteins include those with EGF-like or calcium-binding site domains like C2 domain-containing proteins. Additionally, proteins containing EF-hand domains, which is also a calcium-binding motif that can coordinate Ca^{2+} ions, were found

(Lewit-Bentley & Réty, 2000). Both calcium and magnesium are known to play a role in the complex coacervation process during the formation of granules (Stewart *et al*., 2004). Moreover, proteins with cation-transporting ATPase domains were identified.

Proteins containing glycoprotein domains, such as laminin, which are important building blocks for cellular networks, and chitin-binding domains (Aumailley, 2013), as well as fibronectin, which has been shown to be a glycoprotein domain involved in cell adhesion, migration, and proliferation (Romberger, 1997), were found in the proteomic granule results. Additionally, enzymes containing glycosyl or glycoside hydrolase domains and galactose mutarotase-like domains were identified, suggesting a potential involvement glycoprotein metabolism.

Unexpected proteins have also been identified. For example, proteins containing Histones H3, H2A, or H4 domains were observed. A study on sea urchin adhesives also found histones, indicating that they are not confined to the nucleus alone but can be present in the cytoplasm or on the cell surface (Santos *et al*., 2009b). Proteins with cysteine-rich domains were identified in the MS-MS results: comp264519_c1_seq1 (identified with 16 peptides), comp151502_c0_seq1 (with 5 peptides), and comp255750_c0_seq1 (with 5 peptides). Adhesive proteins often contain specific amino acids like cysteine, which can form disulfide bonds. Additionally, some proteins were found to have leucine-rich repeat domains. Leucine, with its hydrophobic side chain, is an abundant amino acid in the adhesive proteins of *Ciona robusta* larvae (Cheng *et al*., 2022). Proteins containing immunoglobulin domains were also found. It is known that immunoglobulin-like domains may be involved in protein–protein and protein–ligand interactions (Williams & Barclay, 1988). Lastly, several proteins with low-density lipoprotein (LDL) receptor class A repeat domains were identified, which are known to play a role in activities against bacteria and fungi (Liang *et al*., 2019b). An adhesive protein containing this domain has been identified in the pedal mucus of the limpet *Patella vulgata* (Kang *et al*., 2020). Finally, lots of different enzymes such as Peptidases, kinases (serine/threonine kinase, pyruvate kinase, histidine kinase, adenylate kinase, …) or oxidoreductases were found.

By examining the differential expression of each transcript using the Log2FoldChange from the study by Buffet *et al*., 2018, many transcripts were removed (Supplementary Table 5). In the category of putative adhesion proteins, only two transcripts with immunoglobulin domains showed differential expression, along with three transcripts characterized as calcium-

binding proteins. No transcripts were retained in the glycoproteins and protective proteins categories. However, several transcripts characterized as enzymes (24 vs 433 initially) and structural proteins (34 vs 299 initially) were retained after this analysis (Fig. 5).

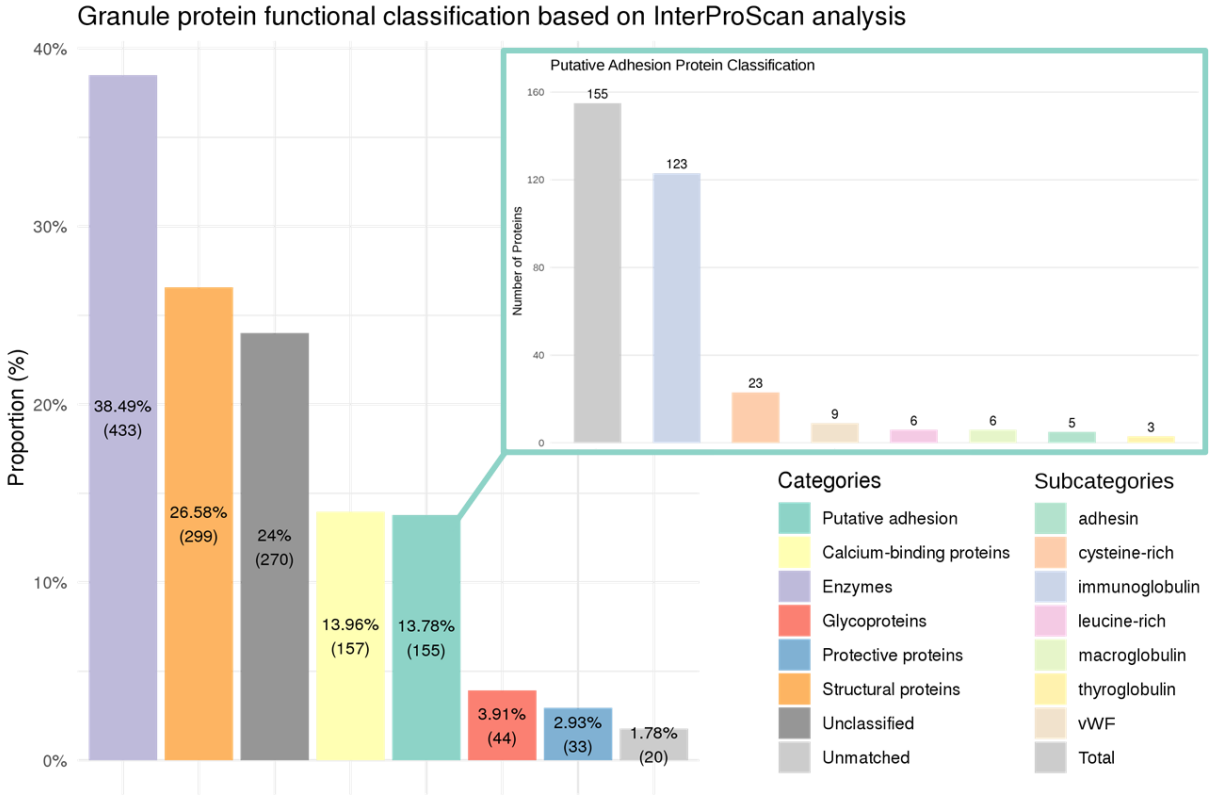


Figure 4. Assumed protein functional composition in the adhesive secretions of *Sabellaria alveolata*, based on an InterProScan analysis of the protein content of cement gland secretory granules. Protein functions were derived from functional domain composition. Emphasis has been put on proteins with a potential role in adhesion.

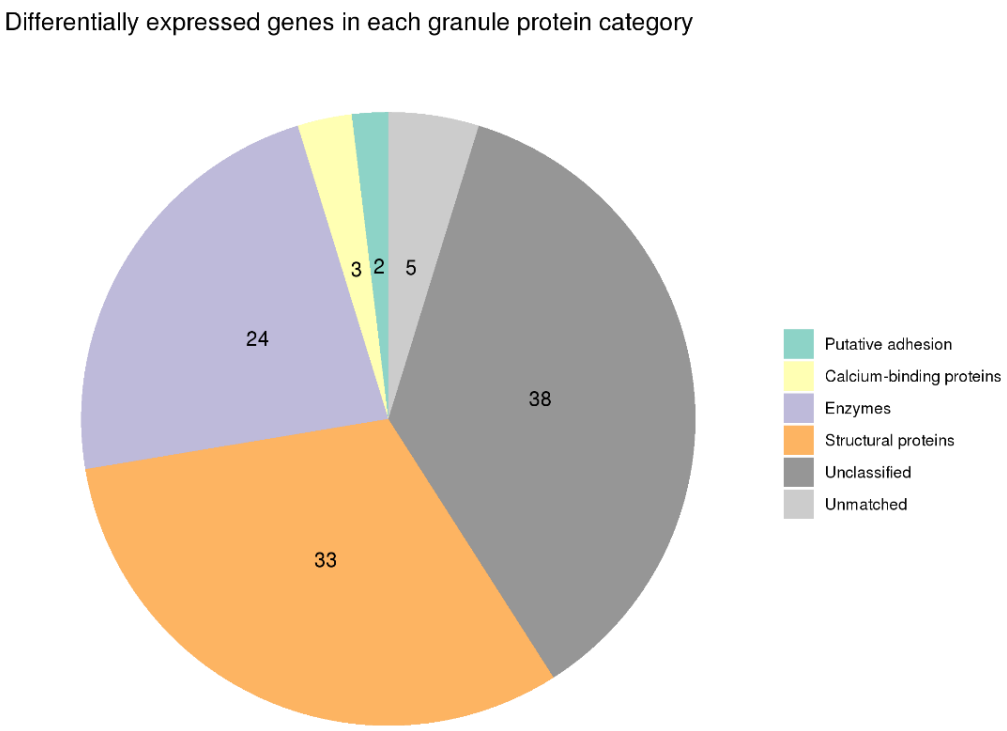


Figure 5. Pie chart showing the transcripts identified from the MS-MS analysis of extracted granules that are differentially expressed between the parathoracic part of *Sabellaria alveolata* and the rest of the body. Overexpressed transcripts are classified in each of the predefined categories; underexpressed transcripts are all grouped together.

3.5. Identification of FAM20C sequences putatively involved in adhesive protein maturation

In this study, we hypothesized that Fam20C enzymes may be the enzymes responsible for the phosphorylation of the serine residue in the adhesive proteins of the honeycomb worms. An earlier study showed that there is an elevated protein sequence homology across different species in the FAM20 family (Tagliabracci *et al.*, 2013a). A comparison of the transcriptome of *S. alveolata* with different Fam20C sequences from other species retrieved from the NCBI database was therefore done. From this BLAST analysis, five transcripts were obtained, named SaFAM20C-1 to -5 (Table 1), and their translated protein sequences were analyzed. According to Tagliabracci *et al.*, 2013b, there are several signature amino acids that define proteins as FAM20 Kinase. First, the sequences must contain a glycine-rich loop, a characteristic feature of this enzyme active site (Bossemeyer, 1994). Fam20 family of kinases have a highly conserved DRHHYE motif in which a conserved aspartate acts catalytically by abstracting a proton from the hydroxyl group of a residue within the phosphor:acceptor substrate. They also require a divalent cation for catalysis, which is bound to a highly conserved motif (a variant motif of DFG) (Fig. 6) (Tagliabracci *et al.*, 2012). Of the five candidates we selected, two do not meet these criteria. SaFAM20C-1 (encoded by transcript comp253537) contains a signal

peptide and has the FAM20C domain (IPR009581 InterProScan reference) in its sequence. However, it lacks some features of the kinase active site, such as the glycine-rich loop and the DRHHYE motif. The reciprocal hit blast revealed similarities between this sequence and sequences that are not described as FAM20C sequences. SaFAM20C-5 (encoded by transcript comp278295) is complete and has the FAM20C domain in its sequence but does not contain a signal peptide or any features of the FAM20C enzymes. Moreover, its reciprocal hit blast results don't correspond to FAM20C sequences, and therefore was not selected for further experiment.

SaFAM20C-2 (encoded by transcript comp288995), SaFAM20C-3 (encoded by transcript comp284991) and SaFAM20C-4 (encoded by transcript comp280217) are complete, contain a signal peptide, a FAM20C domain, and the active site features of the enzyme which are a glycine-rich loop, and a variant of the DFG motif (Fig. 6) (Supplementary Fig. 2). Their reciprocal hit blast corresponds to a sequence described as belonging to the FAM20C family, making them good candidates.

Table 1. List of FAM20C kinase candidate in *S. alveolata* after *in silico* analyses (see methods). Indicated are the transcript ID from the transcriptome of the anterior part of the worm, the amino acid length, completeness of the ORF, proportion of transcripts in the transcriptome, presence of a signal peptide, the domain, and the top reciprocal hit BLAST.

							The reciprocal hit tBLASTn (Database: nucleotide collection (nr/nt))		
Name	**ID Transcripts**	**Length (aa)**	**Complete sequence**	**FPKM**	**Signal peptide**	**CDD**	**Name**	**e-value**	**Id (%)**	**Accession number**
SaFAM20C-1	comp253537_c0_seq2	545	Y	1.3	Y	Fam20C-like Superfamily	*Helobdella robusta* hypothetical protein partial mRNA	1e-103	46.54	XM_009033198.1
SaFAM20C-2	comp288995_c0_seq4	635	Y	5.5	Y	Fam20C	*Octopus sinensis* extracellular serine/threonine protein kinase FAM20C-like	3e-159	51.03	XM_029780782.2
SaFAM20C-3	comp284991_c0_seq1	577	Y	4.14	Y	Fam20C	*Octopus sinensis* extracellular serine/threonine protein kinase FAM20C-like	8e-142	50.98	XM_029780782.2
SaFAM20C-4	comp280217_c0_seq2	433	Y	2.22	Y	Fam20C-like Superfamily	*Crassostrea gigas* glycosaminoglycan xylosylkinase (LOC105346290)	3e-141	47.58	XM_011454801.3
SaFAM20C-5	comp278295_c0_seq4	465	Y	3.65	N	Fam20C-like Superfamily & Signal peptide peptidase Sppa	*Helobdella robusta* hypothetical protein partial mRNA	7e-126	43.8	XM_009033198.1

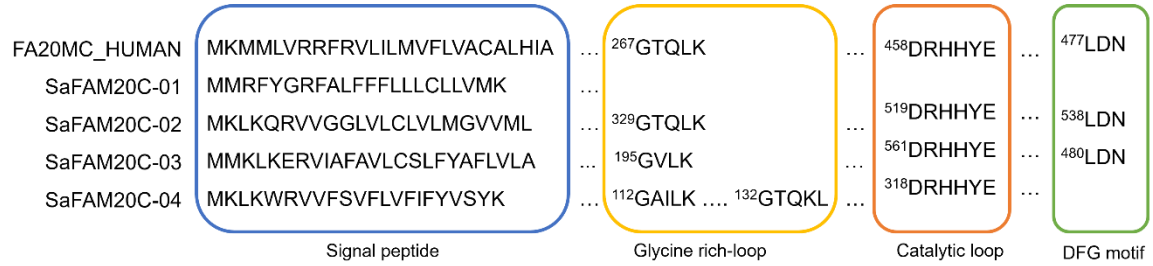


Figure 6. The four candidate SaFAM20C protein sequences aligned with the human FAM20C reference sequence (NCBI accession number Q8IXL6.2). The FAM20C signatures include the N-terminal signal peptide (blue), the glycine-rich loop (yellow), the highly conserved aspartate residue in the DRHHYE catalytic loop (orange), and the DFG-binding motif (green).

To support the role of previously identified candidates in the maturation of adhesive proteins, the localization of these sequences in the adhesive glands of *S. alveolata* was done using the *in situ* hybridization technique with DIG-labelled RNA probes (Fig. 7). Control was made with sense RNA probes, without probes and without antibody (Supplementary Fig. 3). The results show that all the sequences are expressed in both types of adhesive glands. However, those candidates were also found present in other parts of the honeycomb worm, suggesting that these FAM20C kinase candidates may have other roles than the maturation of adhesive proteins (Supplementary Fig. 4).

In parallel to this enzyme localization, the localization of the main adhesive proteins (Sa-1 to -3), whose sequences are available on NCBI, was carried out for comparison and to see if their localization corresponds to that of their homologues in *P. californica*. The proteins Sa-1 and Sa-2 have a positive charge and a mass of 22 kDa. They are rich in glycine, tyrosine, and basic residues. Sa-3 exists in two variants: Sa-3A (22 kDa) and Sa-3B (21 kDa). These proteins are mainly composed of serine residues (75%), which potentially undergo post-translational phosphorylation to generate polyphosphoserine (pSer). As a result, Sa-3A and Sa-3B are highly acidic and negatively charged proteins. Using DIG-labelled RNA probes, we labelled the mRNAs encoding the adhesive proteins of *S. alveolata* on a section of the worm's parathoracic part that displayed the cement glands. Sa-1 and Sa-2 proteins are expressed in cement glands with heterogeneous granules and in glands with homogeneous granules, respectively. These expression sites are similar to those observed for Pc-1 and Pc-2. Like Sa-1, both variants of Sa-3 were expressed in cement glands with heterogeneous granules, consistent with the distribution pattern observed for *P. californica* (Flammang *et al*., 2009; Becker *et al*., 2012; Wang & Stewart, 2012).

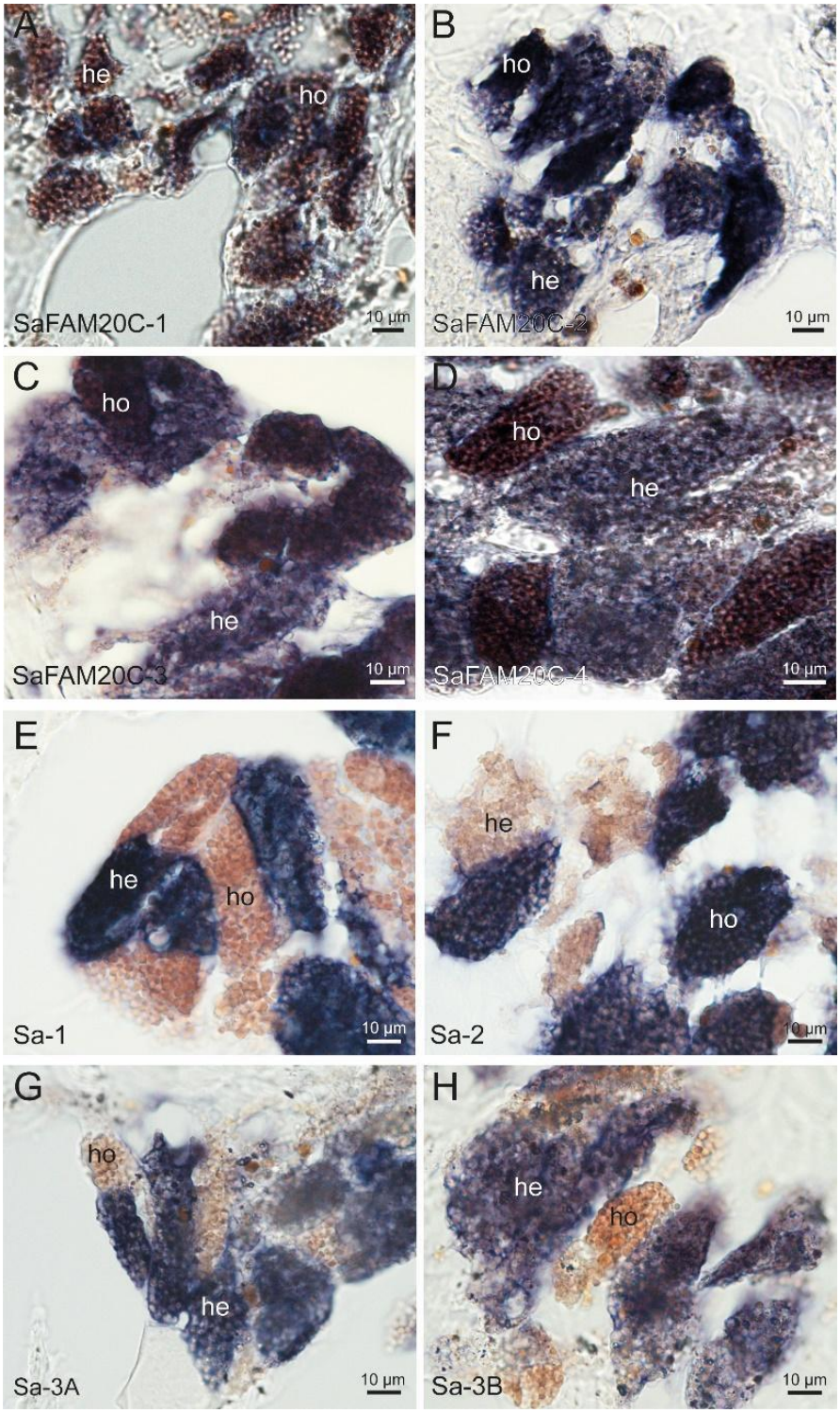


Figure 7. *In situ* hybridization of the kinase FAM20C candidates (A-D) and of the adhesive proteins of the honeycomb worm *S. alveolata* (E-H). Legend: he – cement gland with heterogeneous granules; ho – cement gland with homogeneous granules.

4. Discussion

Sabellaria alveolata is a species of tube-building polychaete worm that is commonly found in intertidal zones along the coasts of Europe. These worms bind together sand grains and shell fragments with a strong proteinaceous cement which hardens in less than 30 seconds under cold and salty water (Stevens *et al*., 2007). This allows the worm to build a hard and resistant tube that serves as a protective home against the hydrodynamic forces of the waves.

4.1. Production of a solid porous material forming highly resistant cement dots

The ultrastructural study of the adhesive glands definitively confirms the presence of two types of cement glands in *S. alveolata*: cement glands containing homogeneous granules and those containing heterogeneous granules. These glands are located in the three parathoracic segments, around the digestive tract and in the parapodia of the honeycomb worm. Their ultrastructure shows that both types of granules have the same spherical shape and size. Homogeneous granules have a uniform content with no internal substructure, while heterogeneous granules contain inclusions of various shapes embedded in a homogeneous matrix. The inclusions are formed by concentric lamellate that are very dense to electrons. In some TEM images, these inclusions show an apparently empty cavity in their centre, which could indicate a form of maturation preceding their secretion. Further studies are needed to confirm this.

TEM observation of the cement dot revealed a homogeneous matrix containing three types of structures: hollow spheroids of variable size, small dense granules, and small lacunae. Both types of granules from the cement glands are excreted simultaneously through individual pores on the epidermal surface of the building organ. There, the homogeneous material from the granules of the two types of cement cells would coalesce to form the matrix of the cement dot. During this process, pockets of seawater can be trapped in the adhesive secretion, giving rise to the lacunae visible in TEM. The inclusions of the heterogeneous granules disperse in this matrix to form dense granules and empty hollow spheroid structures. Among these, the largest spheroids (with diameters exceeding the size of a cement gland secretory granule) appear to originate from remarkable swelling of the inclusions of the heterogeneous granules. This swelling, which is suggested by the empty appearance of some inclusions even before secretion, would be occurring through a still unknown process. The link between inclusions and spheroids is corroborated by simple volume calculations (i.e. the volume of the largest inclusions equals the volume of the cortex in the largest spheroids), as well as by their similar

bright appearance under scanning electron microscopy observations using backscattered electrons (see also Stewart *et al.* 2017). After secretion, intermolecular quinone bonds form between adhesive proteins, involving their DOPA residues (Endrizzi & Stewart, 2009). These bonds allow the adhesives to solidify within a few hours, potentially explaining the porosity gradient observed in the cement dots, as curing would prevent the spheroids from expanding further.

4.2. The inorganic content of the cement is modified during secretion

Both the European and Californian species have heterogeneous granules with inclusions that contain phosphorus and magnesium. In *P. californica*, the concentration of magnesium is sufficient to balance the negative charges of the phosphates (Stewart *et al*., 2017). This high Mg concentration is indicative of the presence of an ATP-dependent H^+/Mg^+ antiporter in the granule membrane (Stewart *et al*., 2017). In *S. alveolata*, the heterogeneous granules also contain a significant amount of phosphorus and magnesium, but also sodium and calcium. Like in *P. californica*, the divalent ions Mg^{2+} and Ca^{2+}, as well as Na^+, can also contribute to the neutralization of negative charges in the granules. It is worth noting that a previous study conducted on *S. alveolata* found small amounts of iron and manganese in the glands' periphery (Gruet *et al*., 1987). However, these metals were not detected in this study. The composition of the granules can also vary according to the fixative buffer used, as shown by Gruet *et al*. (1987).

The elemental analysis was also conducted on cement dots. Deias *et al*. (2023) analyzed the elemental composition of the cement dots from different *S. alveolata* reef sites and found that while the concentrations of most trace elements were similar to those in seawater, those of Ca^{2+} and Mg^{2+} were significantly higher than the mean seawater composition. In our samples, the phosphorus content is twice lower in the cement dots than in the heterogeneous granules, presumably because of the mixing of heterogeneous granules with homogeneous granules in similar quantities. We observed that the amount of calcium was 15 times higher in the cement than in the heterogeneous granules. On the other hand, the magnesium and sodium content strongly decreased. This suggests that, upon secretion, Mg^{2+} and Na^+ might be replaced by Ca^{2+} for complexation with the negative phosphoserine residues of the adhesive proteins. In *P. californica*, it was suggested that secretion is accompanied by a jump in pH from 5 in the secretory granule to 8.2 in seawater that could trigger a change in bonding between Ca^{2+} and phosphate from electrostatic to ionic, the effect of which would be to harden spontaneously and solidify the adhesive (Sagert *et al*., 2006). This could explain why it is important to use an

EDTA treatment for decalcification of the cement dots prior to sectioning, while the cement glands did not require decalcification despite the presence of bivalent cations. SEM analyses conducted on *P. californica* revealed no significant differences in the structure of cement dots treated with EDTA compared to untreated dots, aside from a slight distortion of the pores, which appeared more ovoid (Sun *et al*., 2007). However, EDTA treatment had a strong effect on the mechanical properties of the cement (Sun *et al*., 2007).

4.3. Proteins potentially involved in the adhesive system

The adhesive system of *S. alveolata* is often compared to that of the closely related species, *P. californica*. The localization of their adhesive proteins (Sa-1 to Sa-3 and Pc-1 to Pc-3) in the two types of cement cells is the same, as shown in this study. However, the most expressed adhesive protein in *S. alveolata* is Sa-2, whereas in *P. californica*, it is Pc-1 (Becker *et al*., 2012). In the Californian species, two additional adhesive proteins (Pc-4 and Pc-5) located in the heterogeneous and homogeneous granules, respectively, have been identified, and other putative adhesive proteins (Pc-6 to Pc-18) reported (Endrizzi & Stewart, 2009; Wang & Stewart, 2012). To date, no homologues of these proteins have been characterized in the European species. Furthermore, it has been shown that sulfated polysaccharides are present in the homogeneous secretory granules of *P. californica*, but not in those of *S. alveolata* (Hennebert *et al*., 2018).

To try to identify new components from the cement, we analyzed the protein content of cement granules extracted from the parathoracic part of the honeycomb worm. In addition to a few previously identified adhesive proteins, the proteomic analysis of the adhesive granules of *S. alveolata* identified 2,685 peptides potentially involved in its adhesive system. This high number of proteins may be explained by the cellular debris contaminating the granules samples. Among the identified proteins, actin, histone, tubulin, and myosin, were found to be abundant. These proteins are often considered contaminants, as they may come from the movement of secretory granules, given their known role in granule translocation and attachment to the plasma membrane during exocytosis (Abbineni *et al*., 2013). However, they were also found to be abundant in the adhesive secretion of the sea urchin, comprising up to 65% of its adhesive content (Santos *et al.*, 2013; Lebesgue *et al*., 2016), as well as in that of the sea star (Hennebert *et al*., 2015c). It has been proposed that they may be associated with tension-bearing functions, making them candidate proteins for cohesive elements in the adhesive matrix (Galli *et al*., 2005; Santos *et al*., 2009b). Moreover, a study showed that non-muscle myosin can control cell adhesion through cross-linking and contractile properties, with actin-binding and ATPase

activities present in its head domains (Vicente-Manzanares *et al*., 2009). Some proteins containing glycoprotein domains were also identified in the cement gland's granules. It has already been shown that one protein involved in the permanent adhesion of barnacles is glycosylated (Liang *et al*., 2019a), and in general, glycosylation is highly prevalent in the adhesive system of many organisms. Glycoconjugate proteins may contribute to adhesion and cohesion through electrostatic interactions and hydrogen-bonding of functional groups in their glycan chains (Hennebert *et al*., 2011; Ventura *et al*., 2023). Even if there is no information available regarding the involvement of glycoproteins in *S. alveolata* cement, it should be pointed out that the presence of those structural proteins, the adhesive glycoproteins such as fibronectin and laminin, as well as possibly proteoglycans, suggests an ECM-like composition (Frantz *et al*., 2010). This composition was previously suggested for the adhesive system of sea anemones, with metalloproteases and proteases, such as serine proteases, potentially involved in the remodelling of their ECM-like proteinaceous matrix (Davey *et al*., 2019).

In addition, multiple sequences containing cysteine-rich domains were identified in the MS-MS results, some of which with a high number of peptides. The barnacle protein Cp20K is rich in cysteine, participating in disulfide bonds, and adheres to substrates through non-covalent interactions (Liang *et al*., 2019a). In mussels, the thiol groups of cysteine residues in mfp-6 help prevent the oxidation of DOPA to dopaquinone (Nicklisch *et al*., 2016). Both the formation of disulfide bonds in the adhesive proteins, which would also explain the presence of a disulphide isomerase enzyme, or the prevention of the oxidation of DOPA in adhesive proteins might take place in the cement of *S. alveolata*.

Many proteins involved in calcium binding were found, as well as proteins with cation-transporting ATPase domains. Calcium, which is involved in actin-myosin contraction, was detected in such abundance in the cement dots (15.33%) that we needed to add EDTA to prepare it for TEM observation, highlighting its important role in the adhesive system of *S. alveolata*. Both calcium and magnesium are known to play a role in the complex coacervation process during the maturation of granules.

Finally, multiple identified proteins may have a potential role in the protection of the adhesive secretions from external attack. For instance, a study showed that histones can be present on the cell surface, where they function as receptors for bacterial and viral proteins (Parseghian & Luhrs, 2006; Santos *et al*., 2009b). Proteins such as proteinase inhibitors may protect the adhesive secretions from the degradation due to proteinases produced by bacteria

and other organisms (Zhang *et al*., 2019; Inoue *et al*., 2021; Kang *et al*., 2023), or prevent degradation by an oxidative environment, as suggested for oysters (Whaite *et al*., 2022). Several Low-density lipoprotein (LDL) receptor class A repeat domains were found and are known to have a role in the endocytosis process but were also found to have antibacterial and antifungal activities (Liang *et al*., 2019b).

Among the identified proteins, only two known adhesive proteins (Sa-1 and Sa-3B) were identified, and no tyrosinase enzymes were found, even though they are known to be secreted with the adhesive materials (Wang & Stewart, 2013). This raises questions about the success of this extraction. Moreover, attempts to extract the cement dots present on the tubes have already been tried and failed, indicating their high degree of polymerization (Chapter 1). However, several extracted proteins could be new putative adhesive/cohesive protein candidates based on their functional domain composition. These proteins need further investigation and could lead to the discovery of new adhesive proteins in *S. alveolata*.

4.4. Enzymes responsible for the phosphorylation of serine residue in adhesive proteins

The occurrence of protein phosphorylation in biological adhesion has been reported in various organisms such as sandcastle worms, sea cucumbers, and mussels, and proposed to be an important component for their adhesion (Sagert *et al*., 2006; Flammang *et al*., 2009). For example, mfp-5, an adhesive protein found in the mussel foot, has been shown to contain phosphoserine residues that can bind to calcareous mineral surfaces (Waite & Qin, 2001). This residue can also be involved in protein–protein cross-linking (Taylor & Wang, 2007). But it is in sabellariid tubeworms that this post-translationally modified amino acid seems to be the most important. In sandcastle worms, more than 25% of the cement was found to be composed of phosphoserine (Zhao *et al*., 2005). In this organism, it may play a role in the condensation of the adhesive proteins in the secretory granules through complex coacervation (Stewart *et al*., 2004). As mentioned above, it also participates in the gelation of the secreted adhesive through ionic bonding with calcium ions (Sagert *et al*., 2006). Despite the important roles of phosphorylated amino acid in adhesion, the enzymes involved in phosphorylation are not fully understood.

Currently, little is known about the enzymes involved in the post-translational modifications of the adhesive proteins in sandcastle worms, and only a few studies have been conducted on the tyrosinases that catalyze the formation of DOPA in *P. californica* (Endrizzi & Stewart, 2009; Wang & Stewart, 2012) and *S. alveolata* (Buffet *et al*., 2018; Chapter 1). In

a previous study, researchers attempted to identify and locate a serine kinase responsible for phosphorylating the serine residues of the Pc3 adhesive protein in *P. californica*, but the sequence could not be found (Wang & Stewart, 2012). In this study, we hypothesized that FAM20C could be the serine kinase involved in this modification. It is a secreted protein kinase that phosphorylates the S-x-E/pS motifs within proteins (Tagliabracci *et al.*, 2012). Several FAM20C sequences were retrieved from the NCBI database and a BLAST search was performed in the transcriptome of *S. alveolata*. Five sequences were obtained, all containing a FAM20C domain. The sequences were analyzed to determine if they possessed the amino acid signatures of FAM20 kinases. Four of the sequences had a signal peptide, with three of them containing the glycine-rich loop, the conserved DRHHYE motif and the DFG motif. To confirm their involvement in adhesive protein maturation, we localized their mRNAs using *in situ* hybridization on paraffin sections of the parathoracic part of the worm. All four candidates were expressed in both types of cement glands, supporting the hypothesis that these enzymes might indeed be involved in the maturation of adhesive proteins. However, these FAM20C kinases were also found to be expressed in other parts of *S. alveolata*, as shown by PCR results from other body parts of the worm (Supplementary Fig. 4). It has already been shown that FAM20C kinases are involved in a wide range of biological processes and that they generate the majority of the secreted phosphoproteome in humans, suggesting several roles for these enzymes in honeycomb worms (Tagliabracci *et al*., 2015).

The kinase candidates are expressed in both types of cement glands in *S. alveolata*, but we showed that the polyphosphoserine adhesive proteins are localized exclusively in the glands with heterogeneous granules. This raises questions about why the kinases are also present in the glands with homogeneous granules of *S. alveolata*. Our elemental composition analysis of the granules of the adhesive glands revealed the presence of phosphorus in the heterogeneous granules, as expected, but a significant amount was also found in the homogeneous granules (1.7%). Previous research also showed that a methyl green staining and anti-pSer antibodies stained and labelled the inclusions present in the heterogeneous granules, but also the homogeneous granules with a lower signal (Flammang *et al*., 2009; Becker *et al*., 2012). Moreover, a transcriptomic analysis conducted by Buffet *et al*. (2018) identified a large diversity of cement-related proteins, with over 68% of the overexpressed transcripts assigned to the Poly(S) category. These findings suggest that other unidentified polyphosphoproteins could be present in the homogeneous granules.

5. Conclusion

The findings of this study highlight the complexity of the adhesion system in *S. alveolata* and accentuate the need for further research into the composition and formation of this cement. Moreover, additional studies on the enzymes involved in post-translational modifications of adhesive proteins are also necessary. A better understanding of the cement composition and these enzymes would provide valuable insights into the physical and chemical processes that underlie the assembly of biological materials, which could inspire the design and fabrication of innovative hierarchical materials with diverse applications in various fields.

Chapter III

Supplementary materials for

Characterization of the adhesive system in the honeycomb worm, a marine tube building polychaete

Emilie Duthoo[1], Carla Pugliese[1], Leandro Smacchia[1], Aurélie Lambert[1], Jean-Marc Baele[2], Ruddy Wattiez[3], Matthew J. Harrington[4] and Patrick Flammang[1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Department of Geology, Faculty of Engineering, University of Mons, 7000 Mons, Belgium

[3] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[4] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

This part file includes:

Supplementary Document 1 and 2

Supplementary Table 1, 2 and 3

Supplementary Fig. 1, 2, 3 and 4

Other Supplementary Materials for this manuscript include the following:

Supplementary Table 4: List of the identified sequences in the proteomic analyses with their conserved domains identified by InterProScan. The different columns of the table correspond to the following criteria in order: the accession number of the transcript from the transcriptome of the anterior part of *S. alveolata* (NCBI SRA: SRR29446350), the sequence MD5 digest, the sequence length, its domain analysis, its signature accession, the signature description, the start location of the domain on the sequence, its stop location, the e-value score of the match reported by the member database method, the status of the match (True or False), the date of the analysis, the InterPro annotations – accession, and the InterPro annotations – description.

Supplementary Table 5: List of transcripts identified as differentially expressed in each category, based on the Log2FoldChange values from Buffet *et al.*, 2018. The table includes the transcript ID from the transcriptome, the sequence category, the differential expression value, the main domain, and the top reciprocal BLAST result.

These two supplementary tables are available on:

https://hdl.handle.net/20.500.12907/50241

Chapter III

Supplementary Document 1: Words used for the keyword-based research strategy on gross InterProScan results.

Structural proteins

collagen, tropomyosin, myosin, actin, troponin, spectrin, ankyrin, filamin, profilin, tubulin, intermediate filament, FERM, PDZ, microtubule, ZU5

Calcium binding proteins

EGF, calcium, C2 domain, EF-hand, cation transport, cation-transport, calmodulin

Glycoproteins

laminin, chitin, fibronectin

Protective proteins

Histone, LDL, proteinase inhibitor, inhibitor of metalloproteinase, trypsin inhibitor

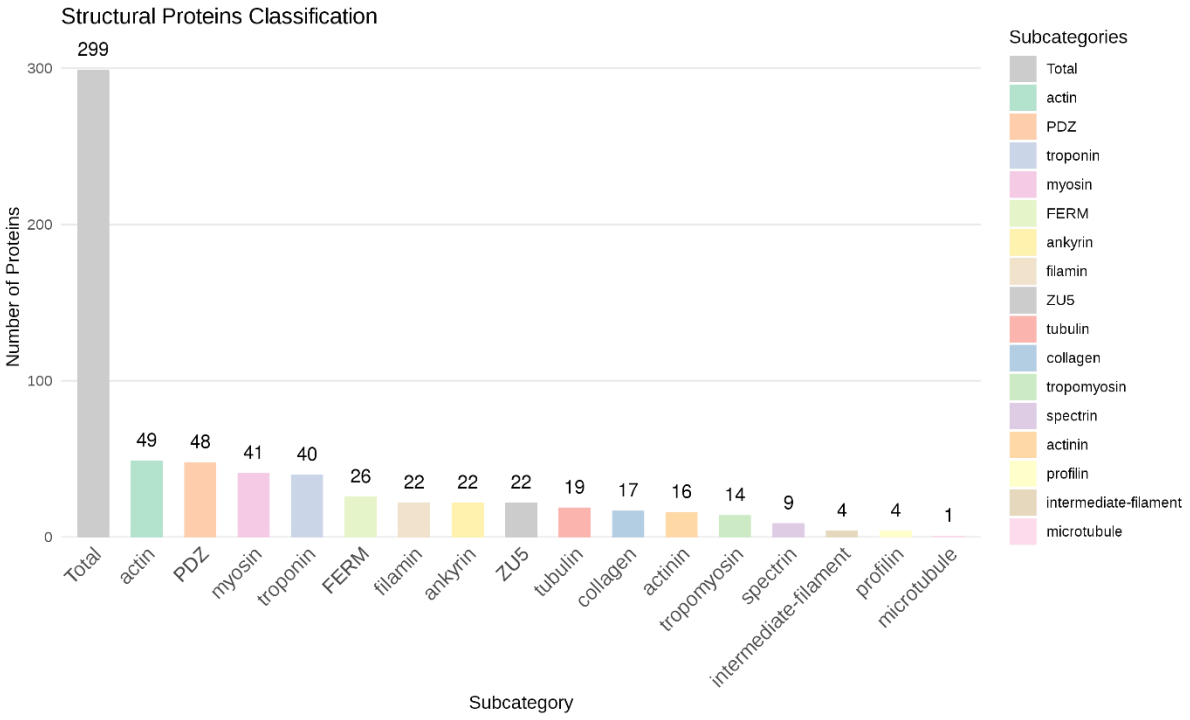
Enzymes

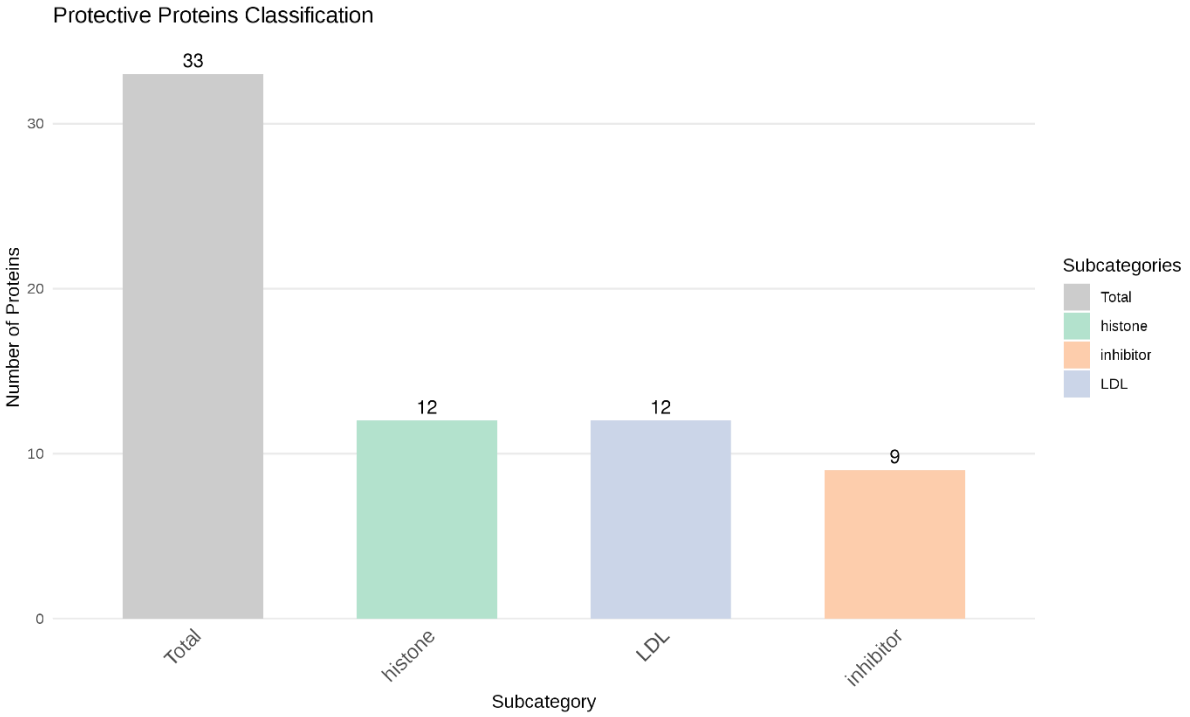
peptidase, kinase, phosphorylase, phosphatase, catalase, protease, proteinase, oxidase, dehydrogenase, synthase, isomerase, transferase, reductase, lyase, ligase, hydrolase, ATPase, mutarotase, transhydrogenase, glucanase, aldolase, GTPase, hydratase, carboxylase, synthetase, anhydrase, glycosidase, glucosidase, hydantoinase, dihydropyrimidinase, aconitase, nuclease, cyclase, dehalogenase, crotonase, fucosidase, enolase, glucuronidase, galactosidase, dismutase, transaminase, glyoxalase, oxygenase, epimerase, nucleotidase, IPPase, trypsin, enzyme

Proteins with potential role in adhesion

cysteine-rich, cysteine rich, leucine rich, leucine-rich, globulin, immunoglobulin, cell wall adhesion, macroglobulin, vWF, adhesin, thyroglobulin

Supplementary Document 2: Bar chart showing the number of proteins in the structural and protective protein subcategories.





Supplementary Table 1. Elemental composition (in wt%) of the heterogeneous granules present in the parathoracic part of *S. alveolata* embedded in Spurr resin**.**

Elements	Hetero 1	Hetero 2	Hetero 3	Hetero 4	Mean	Standard deviation
C	53.7	43.9	45.9	58	50.375	6.61179502
O	30.7	40.2	36.8	30.5	34.55	4.76829809
P	7.4	7.2	8.7	5.7	7.25	1.22882057
Mg	2.6	1.1	3.4	2.3	2.35	0.9539392
Na	2.4	3.1	2.8	1.9	2.55	0.51961524
Ca	0.8	1.2	1.1	0.6	0.925	0.27537853
Cl	1.1	0.7	0.7	0.4	0.725	0.28722813
S	0.8	0.6	0.6	0.4	0.6	0.16329932

Supplementary Table 2. Elemental composition (in wt%) of the homogeneous granules present in the parathoracic part of *S. alveolata* embedded in Spurr resin.

Elements	Homo 1	Homo 2	Homo 3	Homo 4	Mean	Standard deviation
C	66.3	66.1	64.7	64	65.275	1.10867789
O	28.4	28.7	29.1	29.5	28.925	0.47871355
P	1.5	1.3	2	2	1.7	0.35590261
S	1.2	1.3	1.3	1.4	1.3	0.08164966
Na	1.1	1.2	1	1.2	1.125	0.09574271
Cl	0.8	0.8	1.2	1.1	0.975	0.20615528
Mg	0.5	0.4	0.7	0.7	0.575	0.15

Supplementary Table 3. Elemental composition (in wt%) of the interior of six different cement dots sticking two particles together and embedded in epoxy resin.

Elements	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Mean	Standard deviation
C	35.18	49.35	30.23	30.28	55.20	34.54	39.13	10.5536382
O	39.56	32.99	43.16	44.17	33.13	43.62	39.438	5.19839751
Na	0.32	0.16	0.21	0.22	0.20	0	0.185	0.10502381
Mg	0.70	0.36	0.64	0.63	0.32	0.75	0.566	0.18129166
Al	0.05	0.03	0.03	0.00	0.00	0	0.018	0.02136976
Si	0.13	0.03	0.09	0.14	0.00	0	0.065	0.06348228
P	5.59	2.82	3.81	3.80	2.49	3.93	3.74	1.08406642

S	0.86	0.48	0.69	0.70	0.47	0.63	0.638	0.14770466
Cl	0.21	0.13	0.15	0.12	0.12	0.13	0.143	0.03444803
Ca	17.11	13.21	20.20	19.41	7.99	15.33	15.541	4.50850049
Mn	0.29	0.14	0.20	0.28	0.00	0	0.151	0.12967909

Supplementary Figure 1. Picture of a glass bead tube constructed by the honeycomb worm. A dozen sand tubes containing honeycomb worms were placed in a petri dish filled with glass beads, the upper part of the tube having been broken off beforehand. The honeycomb worms were able to reconstruct the upper broken part of their tube in less than 24 hours.

>SaFAM20C-1

MMRFYGRFALFFFLLLCLLVMKKQVITLFGTAREDTALRIVEGDTDLINKPTFIKNKENRNIESPKEGNGILQRENQV
INKFMSTEFLKPQYGSSALPDKHMEVIDELQKHIQSLNISLLEIKELLSVNSTMTASSKHRMPTRVKHINTTDNDIITR
APNDYVSSERSTVHTQATDEYFASDIVQYQKYRNIWNWANNSISHESLFPGKSPEVDTVLNALATARILSIEWFTM
GKKFYENGTSFKWIAMLEGGQKAVIKLAWEEFGGQKKGGRCNDGHELPAAEIAAFHLHRILGFYNTPYVSGRWID
LIHEIYPVACEFVRKQITLLPNGDACVSSGILMYNDNQTLCVSGGKIKASIGYWIPRPLKLYTWYPKYAPFSMARKE
WLDIGFNNKSYCDKIPKTIKPYDQVNYYNDLFDFGLLDMLMYHFDTKHYVIDDGSSANGLTIRLDHGRAFCCYER
DSMNRVFAPIKQCCSLRKTTYERLKAFRHGNVLSLQLKSALKKDPLYPVLFEGWYSALERRLGVLFDTLEKCIQAN
GRDNVLLSS

> SaFAM20C-2

MKLKQRVVGGLVLCLVLMGVVMLKGSLSFPSHLTDTDDTVLKSPHRVKPRHMDPFVRNSVRLSNDANNAGANH
MPNAYANNMPNVDAPESQMKNGLNELMGHLNVMRQELDASNYTSNDRSPEGQRVQSRQSPPSPDQQLQKLILGD
KYIMNRADDIIRLAHVIEETRAGGDVSKNDKPKEFIPNADVEFDKRSQFETGLNDDSDIPPGHEYIRIRKAKERELMA
DQVEKYEREQKQIQRMEVYILPLVMHELEDIATKMTKTRRPSHRRRGATFKIASGSIKNLTGWEKFHNHITQFDMY
DPKDPAIQQVLDDLARQPLVELYQHDGGTQLKIVYWLENDGRALFKPMRFTRDHETLPNHFYFSDYERHTAEIASF
HLDKVLDFHRVPPTAGRAVNMTQDIKMLADSKLLKTFFISPAGNVCFHGQCDYYCDSGHAICGTPDMLEGSLAAY
LPSSRYVSRKTWRHPWRRSYHKRKKAEWETDNTYCQRVRETPPYDTGRRLLDLMDMSVFDFLQGNLDRHHYETF
KDFGNYTFTMHLDNGRAFGKPRYDDLSILAPVYQCCLIRYSTFMKLVSFEQGPHLSEVIDKSMSHDMLYPILIDKHL
RALDRRVRIILKTIHRCVKKNGYQAVIQDDGF

> SaFAM20C-3

MMKLKERVIAFAVLCSLFYAFLVLAKYTNILNNTPLRSHQFSADEIFKDSKSQDAHLSMLRRELSLQKEDSGSDSLK
FDPKEALKVFILNDKVILQEADSIIQIAEILKESRNPKKQKYPEYPNNTHYHFQTYGDYHIIRKKPNFNESSNDHVAY
RKQVESEIKEVRETQWQREQRHQKRISVNLPEHVIEELEGVLKKMKENDLTVPKKTAPNLDIWHLKPGRTELRTWE
VFHNHITPFDLYKPEDPIIDKILHDLTNLPVVELYQHAGGTQIKFVYWFEHNGRALFKPMRFKRDMETLPDHYYFSD
YERHNAEIASWHLDKVLGFHRVPPTAGRIMNITRDIQELADKDLHGTFFISPAGNLCFYGQCDYYCDSNHAVCGNP
TMIEGSMAQYLPSTNHFDRETWRHPWRRSYSRKRKAEWEVTPDYCHRVRKTPPYGSGRRLLDIMDLAVFDFLQD
NLDRHHYETFQKFGNYTALLHLDNGRAFGHPRVDDMSILAPISHCCEIRYSTLQKLIGFERGPKLSEIFDKSMSKDPL
YPLLLDKHLVAIDRRVKILLKTIYKCIAANGYKSVVKDDGF

> SaFAM20C-4

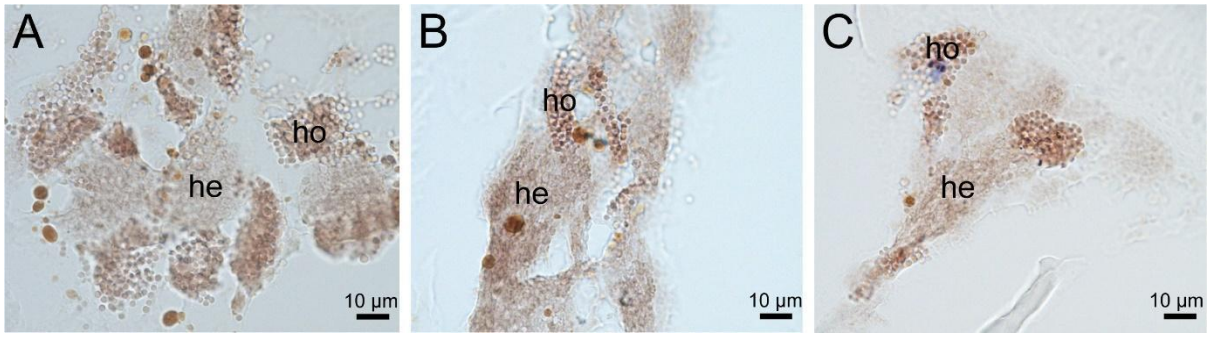
MKLKWRVVFSVFLVFIFYVSYKILEPELHQGGGRRDDGGIYDKADTGNSREFDSDIQEMPLPYNASRGVREKYLLQ
MYDLQWDRPLKEDPWKVAESWVTHRQVHPDYIPELGAILKWMSVATIQLADVGYKGTQLKMLLQLQGNIPVAF
KPKWFGRDEIIPGKAWNGADRHFGEIAGFHLNRILGLNKVPLVVGRTLDLKEEIMPVAKKELLKTFYTKGANTCFY
GKCLYCKNESTGVCGEGEFMEGAIVLWLPTKYKYQKWRHPYQRTYKEGKSARWEEDDSYCSLVVKQQELFQNG
PRLGDILDAAVYDYLILNADRHHYETMSEPWDSMLVMFDSGKSFASPHYDEESILAPLTQCCVIRKNLYERLLMLK
DGVLSKVLKDILNNDPISPVLSNLHLAAMDRRLVRILEAVADCLEKWGPEIVFINDEYR

> SaFAM20C-5

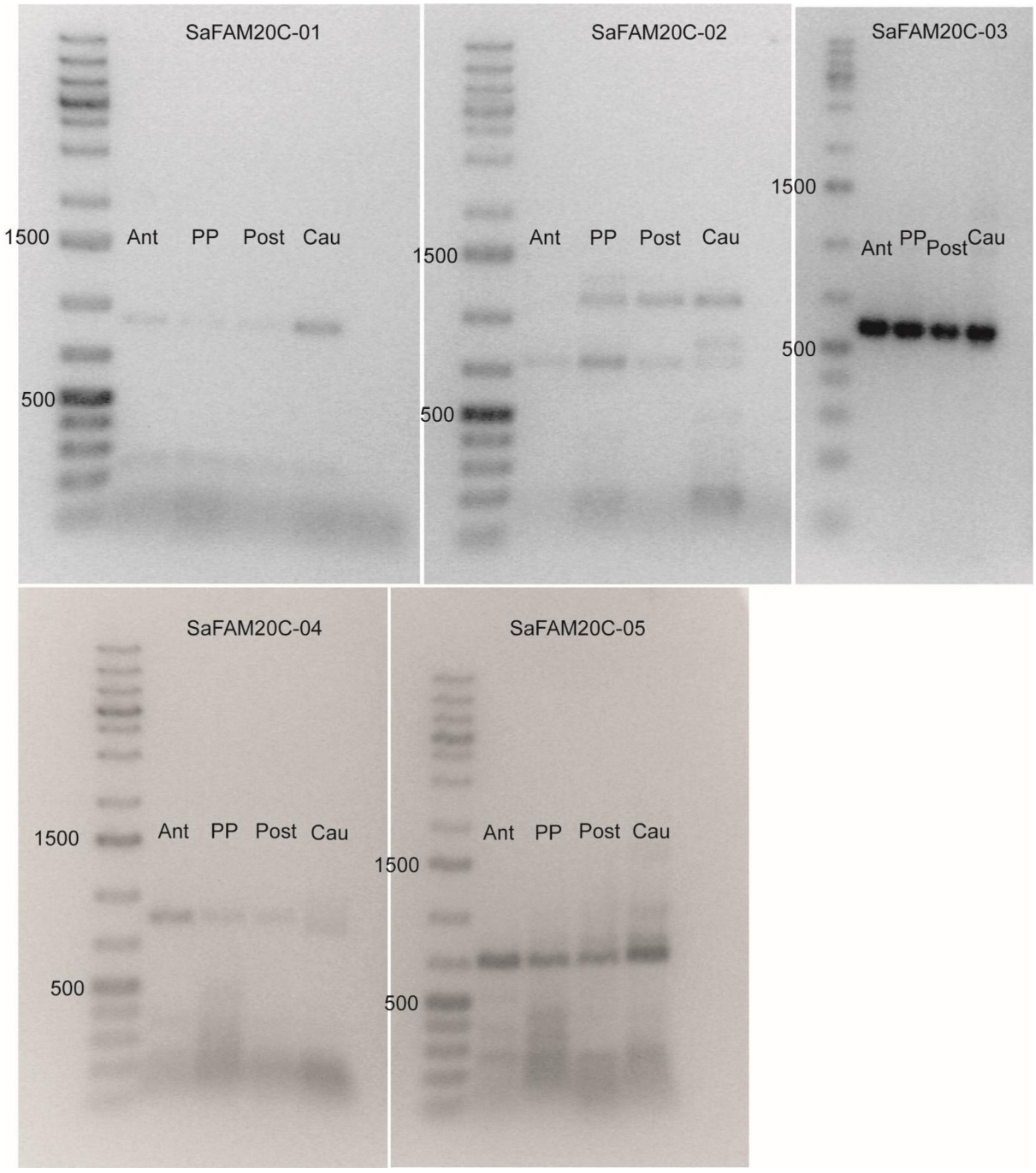
MRFRIRGIRCQVAFFVVLMGLVSVYFILFESSTLPDPSSVYFSKDELEKLILNKNIFHNVTYKSPFEAVYKNIMERQA
KNKSMWQMLHPTESPVVREEKYRKVWQWANDSISQEGLFPGDAPQADLVLKALATARIVSVGGLDMSEYESGTS
RVKWIAQLEGGQKAVIKLVWEKDEWSFSKAKNVIDGQLQSPCNAGHEMCFSEIAGWHLHRVLGFYNTPYVSGRQ
LSLIEEIYPVANEAVRKQISILPGGDACITAKCYLCKYPQRLCVARGMIDASIAYWIPRPLKLYTYPPGYMPYSTPRM
DRWEHIGFNNKTYCKELRQTIEPYEKQRYYNDLFDFAVLDTLMYHYDSKHYVVDDNSKARGLTIRLDHGRAFCFF
DEDNAEIFLAPITQCCSLRKTTYSRLQELRYGDRLTSRLRSALEKDPLYPVLSEVWYPTLERRLKLIFAMLEKCIDAN
GLNNVMLKVFKP

Signal peptide / Glycine-rich loop / Beta-3 / Delta-C / the catalytic loop / DFG motif

Supplementary Figure 2. The five amino acid SaFAM20C sequences with the signature that defines those proteins as FAM20C Kinase are highlighted as follows: the signal peptide in turquoise, the glycine-rich loop in green, the Beta-3 in mallow, the catalytic loop in grey, and the DFG motif in dark teal.



Supplementary Figure 3. *In situ* hybridization control without probes, without antibody, and with sense RNA probes of the parathoracic cement glands of *S. alveolata*. Legend: he – cement gland with heterogeneous granules; ho – cement gland with homogeneous granules.



Supplementary Figure 4. PCR results for each SaFAM20C candidate in different parts of *S. alveolata*. A band is visible in each column, indicating their presence in every part of the honeycomb worm. Legend: Ant - anterior part; PP - parathoracic part; Post - posterior part; Cau - caudal part of the honeycomb worm.

Chapter III

Chapter IV

Morphological and molecular characterization of the organic tube lining in the honeycomb worm *Sabellaria alveolata*

Emilie Duthoo, Jérôme Delroisse, Aurélie Lambert, Ruddy Wattiez, Matthew J. Harrington, and Patrick Flammang

Publication: in preparation, soon to be submitted to Acta Biomaterialia

Contribution: I carried out histological staining, proteo-transcriptomic analyses and *in situ* hybridization experiments. Ultrastructure analyses were carried out by Aurélie Lambert and I. The results were interpreted by Jérôme Delroisse, Patrick Flammang and myself.

Chapter IV

Morphological and molecular characterization of the organic tube lining in the honeycomb worm *Sabellaria alveolata*

Emilie Duthoo[1], Jérôme Delroisse[1,2], Aurélie Lambert[1], Ruddy Wattiez[3], Matthew J. Harrington[4], and Patrick Flammang [1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cellular and Molecular Immunology, GIGA, University of Liège, Belgium

[3] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[4] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

Abstract

The honeycomb worm *Sabellaria alveolata* is a gregarious tubeworm that inhabits a self-constructed tube. It builds this tube by collecting particles from its surrounding environment and by binding them together with a strong cement secreted by a specialized building organ. In nature, the tubes of different individuals are joined side by side, resulting in the formation of impressive reef-like structures. *S. alveolata* also secretes a thin organic sheath that covers the inner surface of its tube, referred to as the organic tube lining in this study. While the tube construction, including the secretion of specific adhesive material, has been extensively studied, the tube lining remains poorly investigated. In *S. alveolata*, the only information on this structure comes from histological and histochemical analyses on the parathoracic ventral shield that secretes it. To date, little is known about the structure and composition of the tube lining. The aim of this study is, therefore, to better characterize the ultrastructure and molecular composition of this inner sheath, as well as to clarify its biosynthesis. Our results indicate that the tube lining is a fibrous structure, consisting of multiple layers of parallelly organized fibres embedded in a matrix. Three main types of secretory cells secreting the inner sheath have been identified: collagen-secreting, acidic mucopolysaccharide-secreting, and catechol-secreting cells. Through a combination of proteomic and transcriptomic analyses, we identified eight main proteins within the organic sheath. Interestingly, most of these proteins do not contain any known conserved functional domains nor share homology with any proteins from public databases.

Chapter IV

1. Introduction

The phylum Annelida exhibits high species diversity, especially among Polychaeta, these species inhabiting a wide range of ecological niches (Rouse & Pleijel, 2001; Eklöf, 2010). Tube-dwelling marine worms, also called tubeworms, are found in different families of Polychaeta. They live in various environments and are found in phylogenetically distant families. However, all of these families appear to share a common feature: an organic tube lining that coats the inner surface of their tubes (Merz, 2015; Shcherbakova, & Tzetlin, 2016; Shcherbakova *et al*., 2017). Studies carried out on 13 different families of tubeworms showed the presence of this tube lining, characterized by a range of micro-textures mainly composed of secreted fibres (Merz, 2015; Shcherbakova & Tzetlin, 2016; Shcherbakova *et al*., 2017). This thin sheath can serve various purposes: (*i*) protecting the worm against abrasion from its mineral tube, (*ii*) facilitating the worm movement within its tube, or (*iii*) acting as a scaffold to maintain the integrity of the tube and prevent fractures. Merz (2015) explained that by regulating the traction between their body and the tube wall, tube-dwelling polychaetes can efficiently move from one end of the tube to the other, support their body during normal functions (e.g., ventilation and feeding), and anchor themselves to prevent dislodgement by predators (Merz, 2015). The fibres making up this tube lining may also contribute to providing the tube wall with high mechanical stability and resistance against deformation or tearing forces, thereby ensuring a durable habitat (Brown & McGee-Russell, 1971). Despite its wide distribution within the families of tubeworms, this structure has been poorly studied, particularly at the molecular level. In Serpulidae, however, the tube lining appears as a thin layer consisting of aligned collagen fibres interspersed with carboxylated and sulfated polysaccharides (Vinn, 2011; Tanur *et al*., 2010).

This tube lining is also present in the European sabellariid tubeworm, *Sabellaria alveolata* (Linnaeus, 1767), commonly referred to as the honeycomb worm, a well-studied species in terms of tube construction (Vovelle 1965; Gruet *et al*., 1987; Le Cam *et al*., 2011; Becker *et al*. 2012; Sanfilippo *et al*. 2019). This species is gregarious and famous for its impressive reef-like constructions, which can be found along most European coasts, from the North Sea to the Mediterranean (Read & Fauchald, 2021). In a few places, these reefs can reach a height of 1.8 metres and extend over several hectares (Noernberg *et al*., 2010; Curd *et al*., 2019; Lisco *et al*., 2020). They are composed of interconnected tubes constructed by different individuals attached to one another. These tubes are made up of sand grains and shell fragments bound together by several dots of a proteinaceous cement secreted by the building organ. The

building organ comprises two types of cement glands located in the parathoracic part of the worm, around the gut and in the parapodia: the cement glands containing so-called homogeneous granules and the ones containing heterogeneous granules with inclusions (Vovelle, 1965). The cement is primarily composed of proteins containing modified amino acids such as DOPA (3,4-dihydroxyphenylalanine) and phosphoserine (Becker *et al*., 2012; Hennebert *et al*., 2015b).

The tube of the honeycomb worm has a specific architecture and is made of three concentric layers of mineral particles cemented together covered internally by a thin organic tube lining (Sanfilippo *et al*., 2019; Fig.1). Watson (cited in McIntosh & Carmichael, 1922) was the first to mention the tube lining and explain how the honeycomb worm applies this layer within its tube. Other authors have described it as a brownish, translucent, and smooth layer (Gruet *et al*., 1987; Sanfilippo *et al*., 2019). According to Vovelle (1965), this inner organic tube lining is secreted by the parathoracic shield, which is located on the ventral surface of the three parathoracic segments of the worm. This parathoracic shield contains different clusters of secretory cells embedded in the subepidermal connective tissue. The first two cell types are located in the first two parathoracic metameres. The first secretory cell type produces a mucoprotein-like secretion with sulfhydryl groups, while the second type secretes acidic mucopolysaccharides. The third type of glandular cell is located in the third metamere and produces an acidophilic secretion comprising orthodiphenol compounds (Vovelle, 1965).

The tube lining, found in most families of tubeworms, holds significant biological importance for these species (Mertz, 2015) and, from a biomimetic perspective, this fibrous structure merits further investigation. However, knowledge about this material remains fragmentary. In this study, we investigated the ultrastructure and composition of the tube lining of *S. alveolata* using electron microscopy as well as proteomic and molecular analyses. We also characterized the morphology of the parathoracic shield. Our findings indicate that the tube lining is a highly complex fibrous material consisting predominantly of proteins that do not resemble any known proteins. One of them, however, contains a collagen domain which appears to show some similarities with bacterial collagen triple helix repeat proteins.

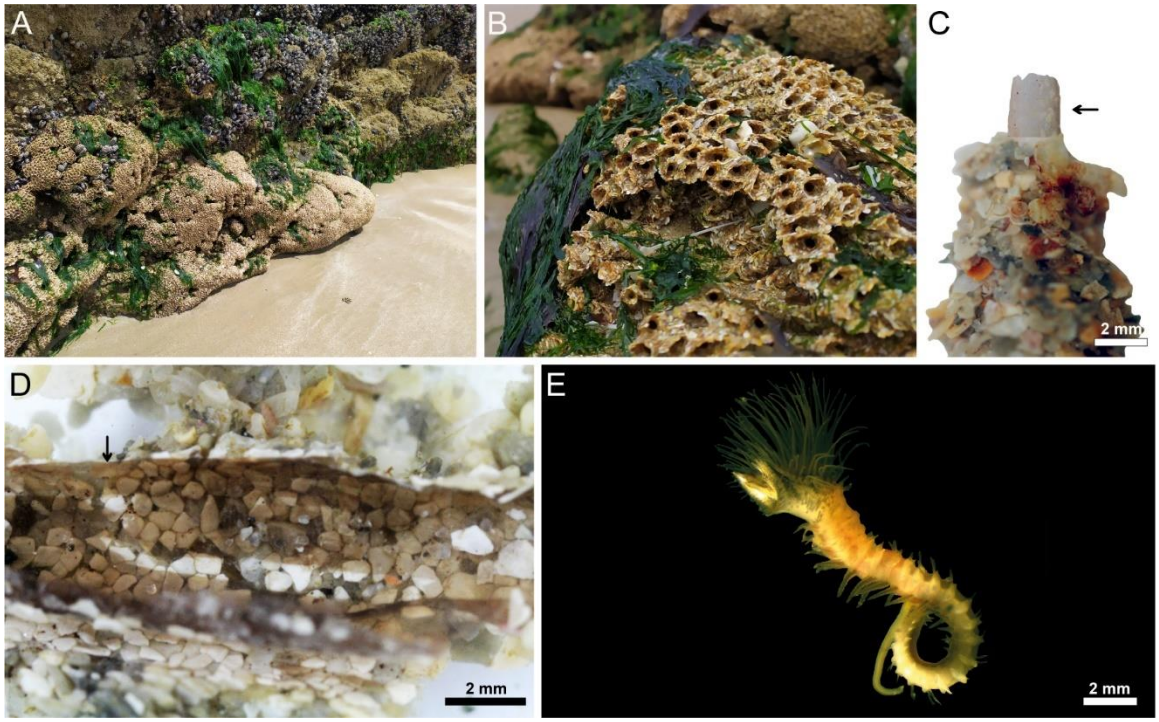


Figure 1. Massive tubes built by the honeycomb worm *Sabellaria alveolata* (Douarnenez, France) (A), with a close-up view of the sand tubes attached to each other (B) (Photos courtesy of Alexia Lourtie). View of the organic tube lining (arrow) inside a longitudinally cut sand tube (C & D) and lateral view of an individual removed from its tube (E).

2. Methods

2.1. Honeycomb worm collection and maintenance

Fragments from colonies of *Sabellaria alveolata* (Linnaeus, 1767) were obtained from the Biological Sample Collection Service of the Station Biologique de Roscoff in Brittany, France. The animals were transported to the Biology of Marine Organisms and Biomimetics Laboratory at the University of Mons, Belgium, where they were kept in a recirculating aquarium maintained at 13 °C and 33 psu salinity.

2.2. Histology

The parathoracic parts of the worms were dissected and fixed in Bouin's fluid for a minimum of 24 hours. Subsequently, the samples were dehydrated in a graded ethanol series, followed by embedding in paraffin wax and sectioning into 7 µm thick sagittal sections using a Microm HM 340 E microtome. Sections were dewaxed using toluene, rehydrated, and stained with Heidenhain's Azan trichrome (Gabe, 1968), Alcian blue at pH 1 and 2.5 (Gabe, 1968), Sirius red (Junqueira *et al*., 1979), Arnow staining (Arnow, 1937), or Periodic Acid Schiff (PAS) (McManus, 1948; Lüllmann-Rauch, 2008). All sections were mounted using Roti®-Histokitt

(Roth) and examined with a Zeiss Axioscope A1 microscope equipped with a Zeiss AxioCam 305 colour camera.

Additionally, the Arnow staining procedure was performed on whole individuals after they were anesthetized in seawater containing 7.5% magnesium chloride for 30 min at room temperature (Williams GC & Van Syoc, 2007). This staining was also performed on sagittal sections through the parathoracic part of the honeycomb worms which were embedded in OCT medium and sectioned using a cryotome.

2.3. Scanning Electron Microscopy

The distal parts of several isolated tubes still containing their honeycomb worms were broken, and individuals were given glass beads (Glass beads, unwashed, 425–600 µm Sigma) to reconstruct their tubes. The glass bead tubes were collected and fixed in Bouin's fluid for 24 hours. The samples were then dehydrated using a sequence of graded ethanol and dried using the critical-point method (Agar Polaron Critical Point Dryer). The different samples were mounted on aluminium stubs and coated with a gold/palladium mixture using a sputter coater (JEOL JFC-1100E). The tubes were observed under a Jeol JSM-7200F scanning electron microscope.

2.4. Transmission Electron Microscopy

Tube linings and parathoracic shields of the honeycomb worms were fixed in 3% glutaraldehyde in cacodylate buffer (0.1 M, pH 7.8, adjusted to 1030 mOsm l-1 with NaCl). The samples were subsequently rinsed three times in cacodylate buffer (0.2 M, pH 7.8, adjusted to 1030 mOsm) for 10 min each. They were then post-fixed in 1% osmium tetroxide in cacodylate buffer (0.1 M, pH 7.8, adjusted to 1030 mOsm) for 1 hour in the dark. Following a final rinse in cacodylate buffer, the samples underwent dehydration through an ethanol series and were embedded in Spurr resin. These embedded samples were subsequently sectioned using a Leica Ultracut UCT ultramicrotome equipped with a diamond knife. Sections were contrasted with uranyl acetate and lead citrate and observed with a Zeiss LEO 906E transmission electron microscope.

2.5. Protein extraction and mass spectrometry analyses

Proteins were extracted from four samples containing approximately 10 tube linings, which were removed freshly from the honeycomb worm tubes. This extraction was performed using a solution of 1.5 M Tris–HCl buffer (pH 8.5) containing 7 M guanidine hydrochloride (GuHCl), 20 mM ethylenediaminetetraacetate (EDTA), and 0.5 M dithiothreitol (DTT) (Tris-GuHCl

buffer). The samples were incubated for 1 hour at 60 °C under agitation. Then, they were treated with iodoacetamide in a 2.5-fold excess (w/w) to DTT in the dark at room temperature for 20 min, and the reaction was then stopped with mercaptoethanol (βMSH), added in the same quantity as the iodoacetamide. The suspension was centrifuged at 13,000 rpm for 15 min at 4 °C. The protein concentration of the supernatant was determined using the Non-Interfering Protein Assay Kit (Calbiochem, Darmstadt, Germany) with bovine serum albumin as a protein standard. For each sample, 50 µg of proteins were precipitated in 80% acetone overnight at −20 °C. After a 15-min centrifugation at 13,000 rpm and acetone evaporation, the resulting pellet was subjected to overnight enzymatic digestion using trypsin (sequencing grade, Promega) at an enzyme/substrate ratio of 1/50 at 37 °C in 25 mM NH_4HCO_3. The reaction was terminated by adding formic acid to a final concentration of 0.1% (v/v) (Hennebert *et al*., 2015c).

The peptides were analysed by LC connected to a hybrid quadrupole time-of-flight TripleTOF 6600 mass spectrometer (AB SCIEX, Concord, ON). MS/MS data were searched for protein candidates against a database consisting of the six open reading frames (ORFs) from the transcriptome of *S. alveolata*'s anterior part (NCBI SRA: SRR29446350) using ProteinPilot software (version 5.0.1). Samples with a false discovery rate (FDR) exceeding 1.0% were excluded from subsequent analyses.

2.6. *In silico* analyses

The sequences of the identified transcripts were translated *in silico* (Expasy, Translate tool) and characterized by analyzing its conserved domains using InterProScan (https://www.ebi.ac.uk/interpro/search/sequence/) (Paysan-Lafosse *et al*., 2023), the presence of a signal peptide using the SignalP-6.0 software (https://services.healthtech.dtu.dk/service.php?SignalP) (Teufel *et al*., 2022) , and the amino acid composition using SAPS (https://www.ebi.ac.uk/jdispatcher/seqstats/saps) (Madeira *et al*., 2024). The translated protein sequences corresponding to these different transcripts were subsequently employed in a reciprocal tBLASTn search against the NCBI non-redundant protein database. An additional BLAST analysis was conducted using the genomes of tubeworm species available on NCBI (see Supplementary Table 2).

2.7. *In situ* hybridization

Parathoracic parts of the worms were fixed in 4% paraformaldehyde in phosphate-buffered saline. They were then dehydrated through a graded ethanol series and embedded in paraffin wax. Sagittal sections of 14 µm in thickness were cut from the samples using a Microm HM 340 E microtome. These sections were mounted on Superfrost Ultra Plus slides (Thermo Scientific) with a drop of Milli-Q water and used for *in situ* Hybridization experiments.

RNA was extracted from the parathoracic part of three honeycomb worms using the TRIzol Reagent kit (Thermofisher). The cDNA synthesis from the extracted RNA was performed using a Reverse Transcription kit (Roche). Subsequently, double-stranded DNA templates were amplified via PCR using the Q5 High-Fidelity DNA Polymerase kit method, with primers designed using Open Primer 3 (bioinfo.ut.ee/primer3/) ensuring an optimal probe length ranging between 700 and 900 bp. A second PCR was conducted with a T7 promoter binding site (5'-GGATCCTAATACGACTCACTATAGG-3') added to the reverse strand PCR primers. PCR products were purified using the Wizard SV Gel and PCR Clean-Up System kit (Promega) and used for RNA probe synthesis.

Digoxigenin (DIG)-labelled RNA probes were synthesized using the DIG RNA Labeling Kit (Roche) with T7 RNA polymerase and DIG–dUTP. *In situ* hybridisation was then performed following the procedures of Lengerer *et al*., 2019, and Rodrigues *et al*., 2014. The probes were applied on sections at a concentration of 0.2 ng/µl and detected using antidigoxigenin-AP Fab fragments (Roche) at a dilution of 1:2000. The signal was developed using the NBT/BCIP system (Roche) at 37 °C until a visible precipitate formed. Sections were observed with a Zeiss Axioscope A1 microscope connected to a Zeiss AxioCam 305 colour camera.

3. Results

3.1. The honeycomb worm secretes a fibrous tube lining

Several studies have noted the presence of a fibrous layer covering the inner surface of the tubes in different families of tubeworms (see e.g., Mertz, 2015). Scanning electron microscopy (SEM) of the inner surface of the tubes of *S. alveolata* revealed the presence of a continuous fibrous sheath approximately 10 µm in thickness. This lining appears as an overlay of layers of fibres embedded in a matrix (Fig. 2). These fibres measure about 40 nm in thickness (Fig. 2G-I). Within each layer, most fibres are oriented parallel to each other, with only a few fibres running in different directions (Fig. 2).

Transmission electron microscopy (TEM) analysis of the tube lining confirmed this complex structure characterized by multiple layers, measuring between 200 and 400 nm in thickness, with fibres mostly oriented in one direction. The fibres, corresponding in size to those visible in SEM, are very electron-dense and embedded in a finely fibrillar matrix (Fig. 2G-I). The areas between fibrous layers contain numerous pores or vacuoles, some of which enclose bacteria. One side of the organic tube lining exhibited a higher fibre density and fewer pores within the matrix (Fig. 2G). This structural arrangement suggests that the tube lining is secreted in successive layers by the worm.

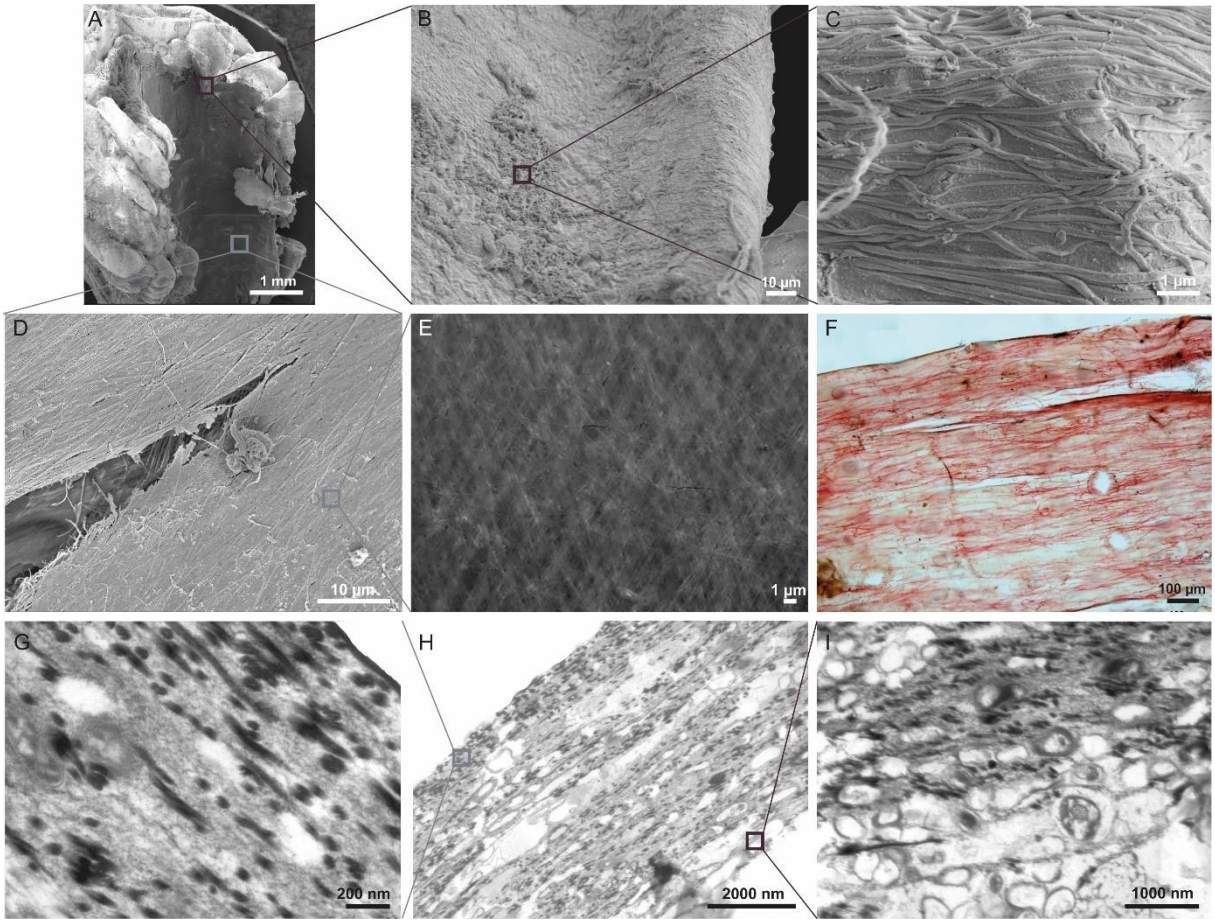


Figure 2. The organic tube lining covering the inside surface of the tube of *Sabellaria alveolata* was observed using scanning (SEM) and transmission (TEM) electron microscopies. Honeycomb worm tube with its organic tube lining: an overview (A) and close-up views of the outer face (B, C) and the inner face (D, E) of it. The tube lining has also been stained with Sirius Red (F). Transverse section of the organic tube lining: an overview (H) and close-up views of the inner (G) and outer (I) faces of it.

3.2. The parathoracic ventral shield of the honeycomb worm contains different unicellular glands

Vovelle (1965) showed that the tube lining is secreted by the ventral part of the parathoracic region of the worm which he named the "ventral shield". This region, comprising the ventral side of the first three metameres located directly posterior to the mouth of the worm, forms a discrete plate (Fig. 3A). Its surface is homogeneously covered with secretory pores about 1 µm in diameter (Fig. 3B).

Several histological stainings were conducted on sagittal sections of the honeycomb worm to investigate the cells responsible for the secretion of the tube lining. Heidenhain's Azan stain provided an overall understanding of the ventral shield histological organization and enabled the identification of various glands within it. All unicellular gland cells are very elongated and subepithelial. Their cell bodies were found to be distributed both deep in the connective tissue, near the cement glands, or more superficially, beneath the epidermis. Long gland cells located in the first two metameres were stained in blue with Heidenhain's Azan stain (Fig. 3C & Supplementary Fig. 1). These same gland cells also exhibited a positive reaction to Sirius red, a dye specific for reticular fibres and collagen (Fig. 2F & 3D), and to Periodic Acid Schiff (PAS) staining, which gave a red coloration to neutral polysaccharides within the samples (Fig. 3E). The glandular cells in the third metamere also displayed positivity for P.A.S., but with fewer instances of positive reactions for Sirius red compared to the first two metameres (see Supplementary Fig. 1).

Only the epidermis of the second and third metameres, along with the shorter underlying glands, were stained with Alcian blue at pH 1 and 2.5 (Fig. 3G-I). Alcian blue is a cationic dye known for its ability to stain sulfated molecules at low pH (pH 1), where sulfated groups are the only ionized groups present. At a higher pH (pH 2.5), it also stains carboxylated molecules, as both types of groups are ionized (Gabe, 1968).

Arnow staining for catechols was conducted on whole individuals of *S. alveolata*. By transversely sectioning the worm at various levels within the parathoracic part, we observed that the three metameres displayed red coloration on their ventral sides. However, a distinct and vivid red coloration was evident in the second and third metamere (see Supplementary Fig. 1). To further explore the possibility of catechol secretion, we extended Arnow staining to sagittal frozen sections of the honeycomb worm. With this method, clusters of cement glands were brightly stained (Fig. 3F). Within the parathoracic shield, we observed more weakly

stained glands, indicating the presence of catechol in all metameres but mostly in the second and third ones.

According to these observations, there would be at least three types of gland cells in the parathoracic shield of the worm: collagen-secreting glands in all metameres, acidic mucopolysaccharide-secreting glands under the epidermis of the second and third metameres and catechol-secreting glands mostly in the second and third metameres.

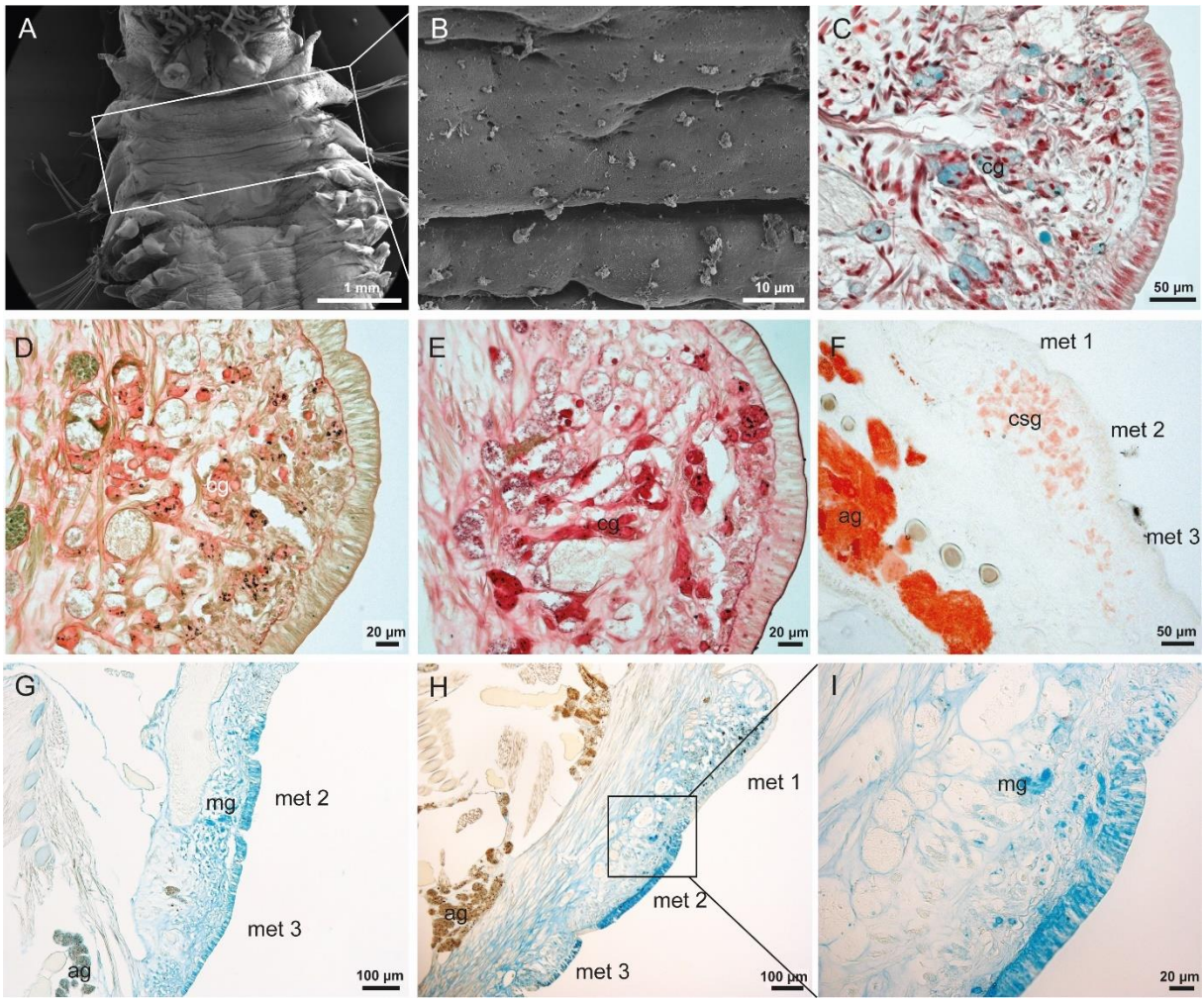


Figure 3. SEM image of the anterior part of *S. alveolata* with the ventral shield framed in white (A) and a close-up view of the secretory pores covering its surface (B). Picture of sagittal sections of the first metamer stained with Heidenhain Azan (C), Sirius Red (D), and P.A.S. (E). Picture of a sagittal section of the ventral shield subjected to Arnow staining (F), Alcian Blue at pH 2.5 (G) and Alcian Blue at pH1 (H) with a close-up view (I). Legend: ag; adhesive glands – csg; catechol-secreting glands - cg; collagen-secreting glands – met 1, 2 and 3; the different metameres of the parathoracic shield - mg; acidic mucopolysaccharide-secreting glands.

In parallel to the histological study, transmission electron microscopy (TEM) was employed to examine the ultrastructure of the various gland cells found within the parathoracic shield, revealing the presence of five distinct types of secretory cells. All the secretory cells are

completely filled with secretory granules and release the contents of these granules on the ventral surface of the parathoracic shield via pores opening through the worm's cuticle. These pores correspond to those observed in SEM (Fig. 3B). The most prevalent cell type (type 1) is characterized by granules containing aligned fibrils (Fig. 4B-C). These secretory granules are ovoid and measure up to 1 µm in width by 3 µm in length. They contain fibrillar material with clear striations perpendicular to the long axis of the granule being separated by 600 nm. Some TEM sections showed different maturation stages of those granules. They appear to derive from more condensed spherical granules, about 1 to 1.2 µm in diameter, containing a two-phase heterogeneous material. In some cells, some granules also appear enlarged and their fibrillar contents decondensed. Type 1 secretory cells are found in all three metameres but are more abundant in the first one. The first metamere also contains another secretory cell type (type 2) characterized by small, spherical electron-dense granules with a size ranging between 0.5 and 1.2 µm of diameter (Fig. 4B). These cells are less conspicuous and abundant than type 1 cells. In the second and third metameres, an abundance of two types of mucocyte-like cells was observed under (type 3) or within (type 4) the epidermis. The former contain secretory granules (around 1 to 1.5 µm in diameter) filled with swirled fibrils and the latter granules (around 1 µm in diameter) of a more granular appearance with various electron densities (Fig. 4D-E). Finally, the second and third metameres house type 5 secretory cells characterized by granules containing a material arranged in concentric circles (Fig. 4F). These granules are spherical with a diameter of about 1 to 1.5 µm. They appear to derive from smaller (0.5 µm), more homogeneous granules.

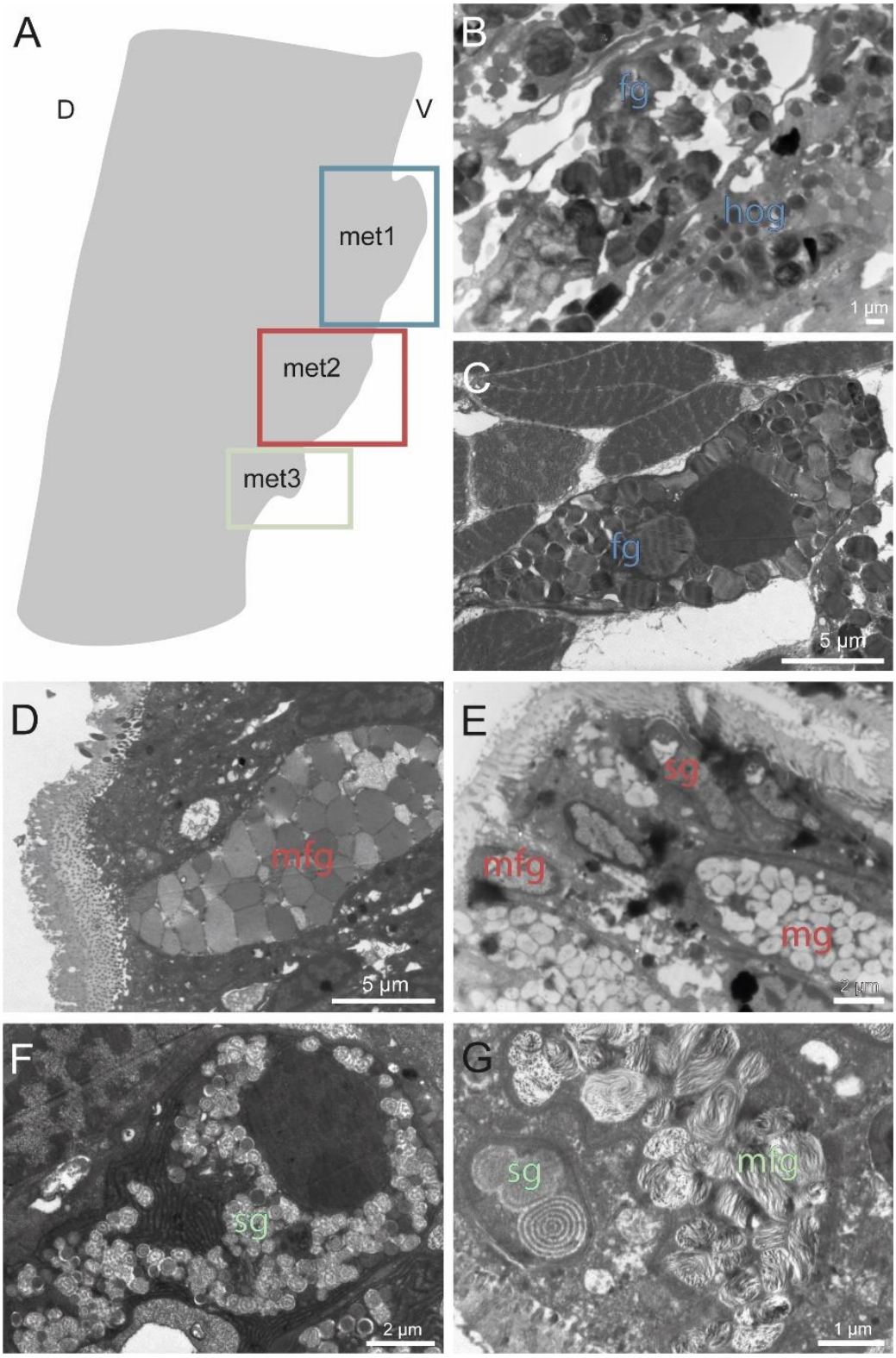


Figure 4. Illustration of the different metameres of the ventral shield (A) and TEM micrographs showing the five types of secretory cells in the parathoracic shield (B-G). Legend: D – dorsal side; fg - secretory cell containing heterogeneous granules filled with fibres in a condensed and decondensed form (type 1); hog- small homogeneous electron-dense granules secretory cell (type 2); mg - mucocyte cells with inclusions containing no fibres (type 4); mfg- mucocyte cells with inclusions containing fibres (type 3); sg - secretory cells containing spiral-shaped granules (type 5); V – ventral side.

3.3. The tube lining contains proteins with no homology to known proteins

Four samples, each consisting of multiple tube linings, were analyzed using mass spectrometry. By comparing these results with the transcriptome of the anterior part of *S. alveolata*, nine transcripts with a peptide95 than 2 were obtained, presumably corresponding to five proteins named SaTLP for *Sabellaria alveolata tube lining protein 1 to 5* (Table 1) (Supplementary Table 1). After *in silico* analyses, the major proteins were found not to contain any known conserved domains and did not have homologous or similar proteins in public databases. Two of them (SaTLP-3 and SaTLP-5), however, show homology with a translated mRNA sequence from another Sabellariidae species, *Phragmatopoma caudata* Krøyer in Mörch, 1863 (formerly named *Phragmatopoma lapidosa*), and are also very similar with one another. Additional BLAST analyses were also done in the genomes of multiple tubeworm species available on NCBI (see Supplementary Table 2) using the newly discovered tube lining protein sequences as queries. From these BLAST searches, no significant matches were found for any SaTLP except for SaTLP-4. However, the coverage and e-value were too weak to validate these results.

Only one of the candidates, SaTLP-4 (comp264488), shows a collagen domain of 61 amino acids in its sequence. Its best reciprocal blast hit corresponds to a bacterial sequence. Bacterial collagen proteins typically exhibit (Gly-X-Y)n regions, lack hydroxyproline, but possess a triple helical arrangement like metazoan collagens. Bacterial collagens maintain stability through electrostatic interactions and/or a high content of glycosylated threonine or a very high content of proline residues (Yu *et al*., 2014). SaTLP-4 is incomplete, features the structure (Gly-X-Y)n and its proline content is approximately 15% (Supplementary Table 1). Although this sequence shares similarities with bacterial collagen proteins, the incomplete nature of the corresponding transcript complicates the analysis of its evolutionary origin (see Supplementary Document 3).

The proteomic analysis also identified an incomplete sequence containing a tyrosinase domain. Tyrosinases are found in almost all living organisms (Aguilera *et al*., 2013). This enzyme belongs to the type-3 copper protein family and catalyzes the o-hydroxylation of monophenols into o-diphenols (i.e., catechol) and the further oxidation of *o*-diphenols to *o*-quinones (Ullrich & Hofrichter, 2007). It could potentially catalyze the production of catechol, which has been demonstrated by Arnow staining, in some of the glands of the worm's parathoracic shield.

Table 1. List of transcript sequences identified in the mass spectrometry analysis of the tube lining of *S. alveolata*. The table includes the following information: transcript ID from the transcriptome of the worm's anterior part, number of peptides matching the transcripts from the four MS analyses (P95), protein length in amino acids (aa), completeness of ORF, normalized expression level of the transcripts in the transcriptome (FPKM), presence of a signal peptide, presence of a conserved domain (CDD), and top reciprocal BLAST hit. Legend: NA – No data available.

		Peptides identified by proteomic analyses									The reciprocal hit tBLASTn (Database: nucleotide collection (nr/nt))			
ID	**Transcript ID**	**P95 1**	**P95 2**	**P95 3**	**P95 4**	**Length (aa)**	**Sequence complete**	**FPKM**	**Signal peptide**	**CDD**	**Name**	**e-value**	**Id (%)**	**Accession number**
SaTLP-1	Comp264093_c0_seq2	13	32	19	51	676	N	76.45	Y	NA		NA		
	Comp264093_c0_seq1	13	32	20	52	642	Y	76,45	Y	NA		NA		
SaTLP-2	Comp258733_c0_seq1	10	27	9	24	153	N	529.64	N	NA		NA		
	Comp258733_c1_seq2	10	22	8	23	116	N	2.87	N	NA		NA		
SaTLP -3	Comp249094_c0_seq1	2	14	7	37	230	Y	62.67	Y	NA	*Phragmatopoma lapidosa*	6e-14	33.72	KM267659
SaTLP -4	Comp264488_c0_seq3	4	4	2	7	415	N	45.52	N	Collagen	*Hungatella xylanolytica*	5e-07	60	CP070896
SaTLP -5	Comp267302_c0_seq1	2	8	4	8	177	Y	43	Y	NA	*Phragmatopoma lapidosa*	4e-14	33.52	KM267659
SaTLP-6	Comp258812_c0_seq1	1	2	5	2	191	Y	68.77	Y	NA		NA		
SaTLP-7	Comp268765_c2_seq7	1	2	4	2	288	N	160.15	N	NA		NA		
SaTLP-8	comp253654_c0_seq1	1	1	1	10	370	N	36.84	N	Tyrosinase	*Lumbricus terrestris*	4e-49	41.57	OX457036

To investigate the expression of the most abundant candidates (estimated by the number of peptides identified by MS analyses) within the tissues responsible for secreting the tube lining, *in situ* hybridization (ISH) experiments were conducted on sections through the parathoracic shield (Fig. 5). Controls were performed using sense RNA probes, as well as without probes and without antibody (Supplementary Fig. 2). Using ISH, the mRNAs coding for SaTLP-1 (corresponding to transcript comp264093) were localized in the third metamere, whereas those coding for SaTLP-2 (comp258733) were identified in the second metamere. Additionally, SaTLP-3 mRNAs (comp249094) appear to be present in gland cells distributed across all metameres, but with a weaker signal (Fig. 5). Unfortunately, *in situ* hybridization experiments could not be achieved on the mRNAs coding for the proteins SaTLP-4 and -5.

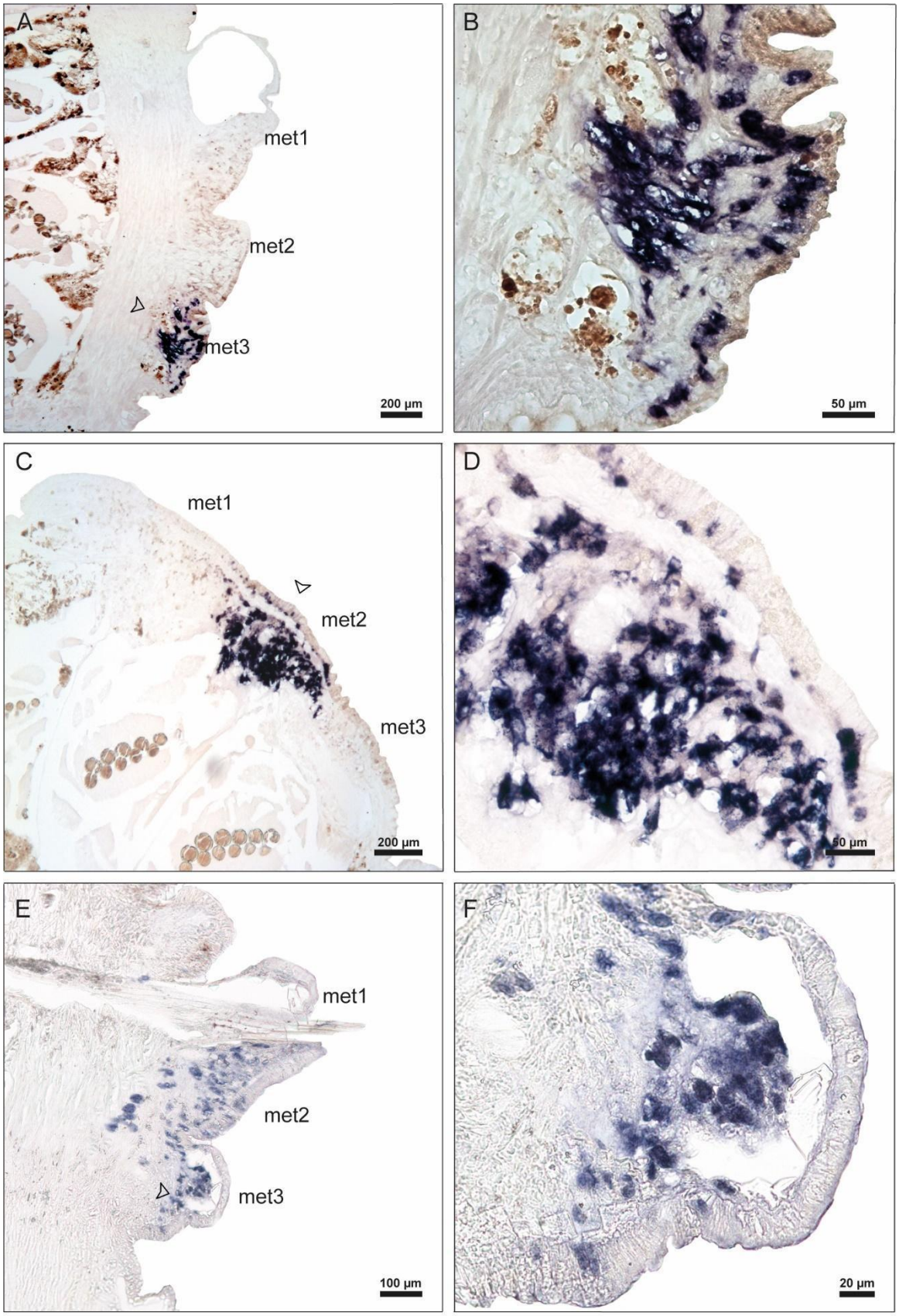


Figure 5. Localization of SaTLP-1 (A-B), SaTLP-2 (C-D) and SaTLP-3 (E-F) transcripts using *in situ* hybridization on sections through the parathoracic shield of the worm (with closer view on the left). Legend: met1 – first metamere; met2 – second metamere; met3 – third metamere.

4. Discussion

The honeycomb worm *Sabellaria alveolata* secretes a thin, fibrous organic tube lining that covers the interior surface of its tube. The tube lining, secreted by the ventral parathoracic shield (Vovelle, 1965), has received little attention leaving its composition, structure and evolutionary origin largely unknown. In *S. alveolata*, observations of the tube lining once isolated from the tube indicate a thickness of approximately 10 µm. It comprises multiple layers of fibres, each with slightly varying orientations, along with a matrix that binds these fibres together. Furthermore, the distinct arrangement of fibre layers between the region in contact with the tube surface and the non-contact area, suggests a progressive, overlapping layer-by-layer secretion process.

Histological and histochemical results indicate the presence of three most prevalent glandular types in the parathoracic shield: cells stained with Sirius red and PAS, cells stained with Alcian blue, and cells stained with the Arnow's method. On the other hand, in TEM, five secretory cell types were distinguished. In most cases, however, the distribution of the cells in the three parathoracic metameres and the ultrastructure of their secretory granules allows to reconcile light and electron microscopy. Type 1 secretory cells are filled with heterogeneous granules containing arrays of fibrils. They occur in all three metameres but are more abundant in the first one. Their contents therefore likely consist of collagen (as evidenced by Sirius red staining), with the PAS staining corresponding to the glycosylated part of collagen. It could also indicate the presence of other proteins in the vesicles, like muco-rich proteins. Type 2 secretory cells, showing a similar distribution, contain small homogeneous electron-dense granules. These cells have not been observed in histological sections. Vovelle (1965) described mucous glandular formations within and beneath the epithelium of the second metamere of the parathoracic shield. These glands would secrete acidic mucopolysaccharides. Type 3 and 4 secretory cells correspond to this description. They are mucocyte-like glands located within or below the epithelium in the second and third metameres, produce a polysaccharidic secretion bearing carboxylated and sulfated groups, as demonstrated by Alcian blue staining at different pHs. Type 3 secretory cells are subepidermal and are packed with granules containing fibrils. Type 4 secretory cells are intraepidermal and their granules lack fibrils. Acidic polysaccharides could form the matrix that holds the fibres together within the tube lining (Fig. 6). This hypothesis is supported by the fact that one of the two mucous glands contains smaller fibrils resembling those observed in the matrix.

The last main gland contains catechol, as shown by Arnow staining. The staining reveals that these glands are essentially present in the second and third metameres. They could therefore correspond to type 5 secretory cells which are located more deeply in the second and third metameres and which contain granules with a content arranged in concentric layers. Catechols, such as DOPA (3,4-dihydroxyphenylalanine), are known to be an important component for the adhesion of tubeworms (Zhao *et al*., 2005), as they can form links with the substrate. The oxidized form of DOPA residues is also known to play important roles in cross-linking reactions (Yu *et al*., 1999). We can therefore assume that catechols may be involved in the cross-linking of the tube lining. DOPA is produced by the hydroxylation of a tyrosine residue by a tyrosinase enzyme (Sánchez-Ferrer *et al*., 1995) and a protein with a tyrosinase domain (SaTLP-8) has been identified by the proteomic analysis of the tube lining, supporting this hypothesis. However, according to Vovelle (1965), the catechol secreted by the parathoracic shield is not bound to proteins (i.e., not in the form of DOPA). It could still be involved in tube lining cross-linking but in a way more similar to the sclerotization of arthropod cuticle (Fig. 6) (Hopkins *et al*., 1999).

4.1. The tube lining secreted by *Sabellaria alveolata* consists of novel proteins

Eight major proteins were identified in the tube lining of the honeycomb worm. Based on the differential expression data of Buffet *et al*. (2018), all the transcripts coding for these proteins appear to be overexpressed in the parathoracic area, except those encoding SaTLP-3 and -5. The expression profiles of the transcripts corresponding to the three proteins having the highest number of identified peptides (SaTLP-01 to -03) were verified by *in situ* hybridization. All three proteins were expressed in the parathoracic shield in different metameres, confirming their secretion in the tube lining.

Proteins SaTLP-1 and -2 do not resemble any known proteins, precluding any inference about their evolutionary origin, structures, or specific functions. These sequences exhibit an acidic isoelectric point (between 4.11 and 4.45). SaTLP-1 is especially rich in aspartate residues (Supplementary Table 1) and is expressed exclusively in the third metamere of the parathoracic shield. SaTLP-2, on the other hand, is expressed in the second metamere. Among the different proteins identified, only proteins SaTLP-3 and SaTLP-5 exhibit similarity with a secreted protein from *Phragmatopoma lapidosa*, albeit with no known conserved domain. *P. lapidosa* is a closely-related species from the family Sabellariidae which presumably also possesses a tube lining. However, the other SaTLPs do not show similarity with any *P. lapidosa*

protein in the BLAST search. This may indicate that the organic tube lining in *P. lapidosa* has a different composition compared to the honeycomb worm.

Only one transcript, SaTLP-4, was found to contain a functional domain in the proteomic results. This transcript shows similarity with a bacterial collagen sequence. Different *in silico* analyses were conducted to see if this gene could have been transferred from bacteria to a honeycomb worm ancestor (see Supplementary Document 3). The results of those analyses show that there is some similarity between SaTLP-4 and bacterial sequences from the public databases. However, SaTLP-4 is incomplete and the repetitive motif (here (GPT)n) within its sequence makes similarity analyses difficult to carry out.

It is worthy of note that the proteomic analysis was based on trypsin digestion, which involves cleaving proteins at the C-terminal end of arginine and lysine residues. It is possible that not all proteins were extracted, especially if their sequences do not contain or contain too many of these amino acids. Additionally, the efficiency of this digestion method depends on multiple factors, including the digestion and denaturing conditions, and the presence of post-translational modifications (Proc *et al*., 2010; Walmsley *et al*., 2013). To identify additional tube lining proteins, it may be useful to test other proteolytic enzymes, such as pepsin.

4.2. A potentially conserved organic tube lining structure and composition in tubeworms?

The phylogenetic relationships among Annelida, notably the Polychaeta, are often controversial. In the early 1800s, Lamarck proposed a taxon within the Polychaeta, composed of sessile and mostly tubicolous worms, called Sedentaria (Lamarck, 1818). However, this taxon was shown to be polyphyletic and was discarded by Fauchald in 1977 (Fauchald & Rouse, 1997). In recent phylogenies, tubeworms are found in diverse and phylogenetically distant clades of polychaetes, such as Sabellaridae and Serpulidae in Canalipalpata, and Maldanidae in Scolecida (Eklöf, 2010). Additionally, different types of tubicolous worms have been identified: those living in secreted calcareous tubes like the Serpulidae (Vinn *et al*., 2009), those building their tubes by cementing exogenous materials like Sabellariidae (Vovelle, 1965), or those living in entirely organic tubes like members of the Alvinellidae family (Gail & Hunt, 1986). Despite their diversity and distant phylogenetic positions, an inner organic sheath lining the inner surface of their tubes is a common feature among most tubeworms. This sheath usually comprises successive layers made of sheets or fine threads (Daly, 1973; Nishi, 1993; Shillito *et al*., 1995; Merz, 2015). For example, evidence of a tube lining has been found in the families Terebellidae (Shcherbakova & Tzetlin, 2016), Maldanidae (Kongsrud & Rapp, 2012;

Shcherbakova *et al*., 2017), and Alvinellidae (Tunnicliffe *et al*., 1993). Merz (2015) noted the presence of this tube lining in many other families, such as Oweniidae and Nereididae, all showing deposition of superimposed layers or series of fibres with changing orientations from one layer to another inside the tube (Merz, 2015). Moreover, it has been shown that the tube lining structure in the Maldanidae family varies among species, being sometimes thicker in certain species and containing many fibre layers that provide elasticity and strength to the tube (Shcherbakova *et al*., 2017). The organic tube lining thus seems to be structurally similar in most tubeworms, which suggests it potentially derived from a common ancestor.

In terms of composition, it has been shown that worms from the Serpulidae family secrete a thin layer of collagen fibres mixed with carboxylated and sulfated polysaccharides inside their calcareous tube (Vinn *et al*., 2009; Vinn, 2011). This general composition being similar to what we observed in *S. alveolata*, it would be tempting to postulate that all polychaete tube lining are also similar in terms of composition. However, we could not find any homology between the tube lining proteins we identified and proteins from annelid databases. Tube lining in polychaetes seems therefore to have evolved independently in different lineages of tube-dwelling worms but converged to morphologically similar structures.

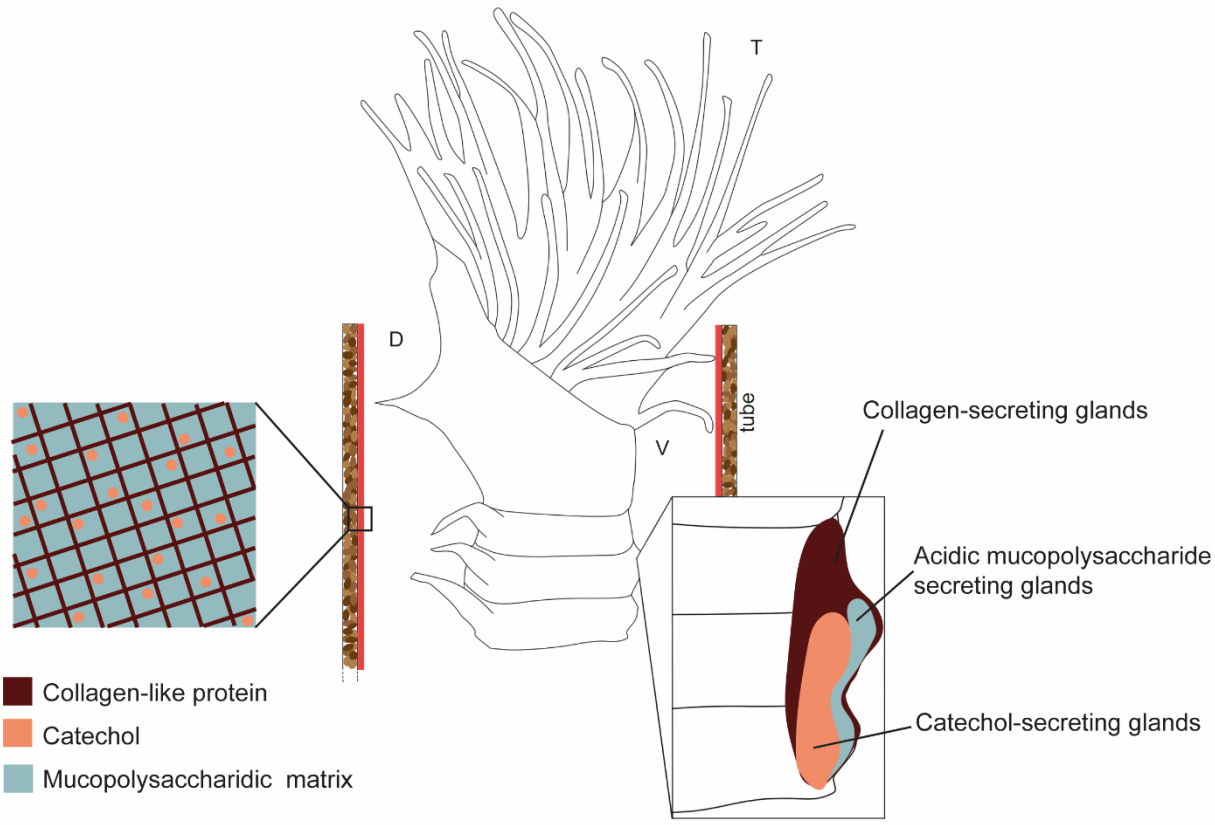


Figure 6. Schematic model of the honeycomb worm with its organic tube lining covering the inner surface of the tube formed from collected exogenous particles: *S. alveolata* applies several layers of a fibrous material to cover the inner surface of its tube by rotating around itself. The tube lining is secreted by three distinct glands of the ventral parathoracic shield of the worm: the collagen-secreting gland, the acidic mucopolysaccharide gland, and the catechol-secreting gland. Together, their secretions form a thin, hard layer of collagen filaments embedded in a matrix of mucopolysaccharides and catechol. Legend: D – dorsal part; T – tentacles; V – ventral part.

5. Conclusion

This first morpho-molecular study of the organic tube lining covering the inner surface of the tube of *S. alveolata* provides new insights into the structure and composition of this complex material. Although analogous structures are present in most tubeworm species, the tube lining proteins we identified are novel and have no homologues in databases. An in-depth study of the tube lining found in other polychaetes is required to understand the evolution of this structure which plays crucial roles for these tubeworms.

Chapter IV

Chapter IV

Supplementary materials for

Morphological and molecular characterization of the organic tube lining in the honeycomb worm *Sabellaria alveolata*

Emilie Duthoo[1], Jérôme Delroisse[1,2], Aurélie Lambert[1], Ruddy Wattiez[3], Matthew J. Harrington[4], and Patrick Flammang [1*]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[2] Laboratory of Cellular and Molecular Immunology, GIGA, University of Liège, Belgium

[3] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium

[4] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

This part file includes:

Supplementary Fig. 1, 2 and Document 3

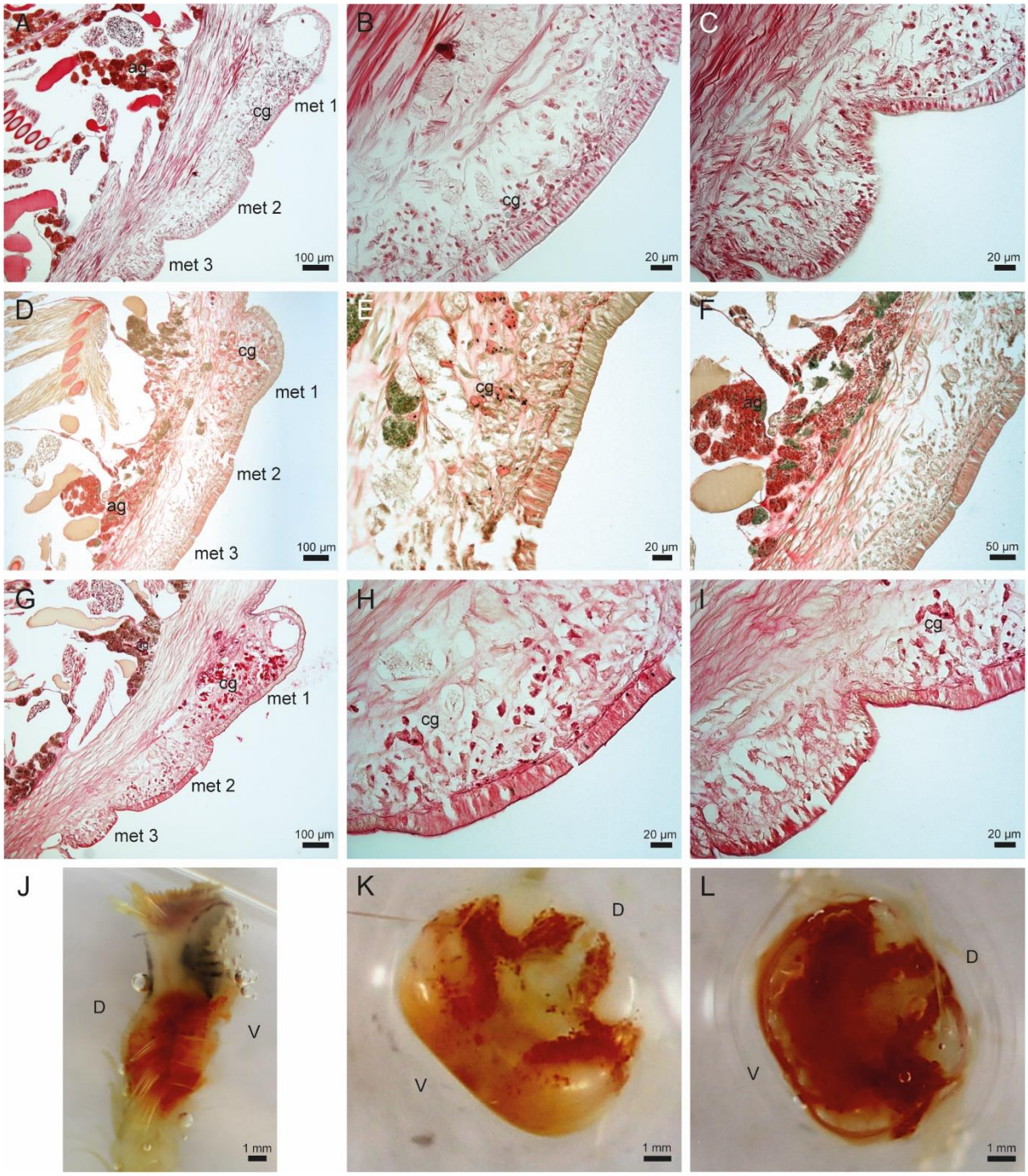
Other Supplementary materials for this manuscript include the following:

Supplementary Table 1. Excel file with the result of the proteomic analyses made on four different samples containing organic tube linings.

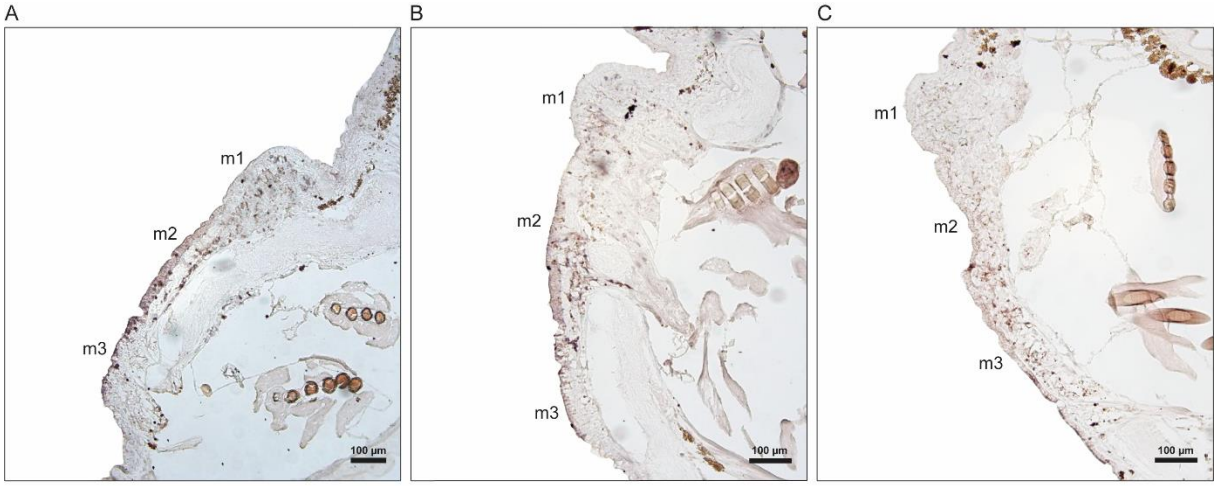
Supplementary Table 2. Excel file with the BLAST results of SaTLP-04 with different tubeworm genomes available on NCBI.

These two supplementary tables are available on:

https://hdl.handle.net/20.500.12907/50241



Supplementary Figure 1. Sagittal sections of the parathoracic shield with close-up views of the second and third metamers of the worm subjected to Heidenhain Azan staining (A-C), Sirius Red staining (D-F), and P.A.S. staining (G-I). Arnow staining of the worm with transverse sections of the first (K) and third metameres (L). Legend: ag; adhesive glands – csg; catechol-secreting glands - cg; collagen-secreting glands; D – dorsal side; met 1, 2 and 3; the different metameres of the parathoracic shield - mg; acidic mucopolysaccharide-secreting glands; V – ventral side.



Supplementary Figure 2. *In situ* hybridization (ISH) controls with sense RNA probes (A), without probes (B) and without antibodies (C) on a section of the parathoracic shield of *S. alveolata*. Legend: met 1, 2 and 3; the different metameres of the parathoracic shield.

Supplementary Document 3.

In silico analyses of the transcript SaTLP-4

During the proteomic analysis of the inner organic sheath, several proteins were identified. Only one of them, SaTLP-4 (encoded by transcript comp264488), shows a functional domain in its sequence. This sequence has indeed a collagen domain, and its best reciprocal BLAST hit corresponds to a bacterial sequence translated from a genome. Collagen sequences have (Gly-X-Y)n regions that form a stable triple helix arrangement, with a high content of proline and the post-translationally modified hydroxyproline in the Y position. This hydroxyproline allows the stabilization of the helix thanks to the stereoelectronic effects (Yu *et al*., 2014). In SaTLP-4, we have a (Gly-Pro-Thr)n motif which is more similar to bacterial collagens, as bacterial collagen structures have threonine as the dominating amino acid in the Y position. It has been suggested that threonine establishes direct hydrogen bonds with neighboring carbonyls in the backbone of the helix, replacing hydroxyproline in the stabilization of the collagen triple helix (Rasmussen *et al*., 2003). Therefore, several *in silico* analyses were conducted to have a better understanding of the evolutionary origin of the sequence SaTLP-4.

Method

To gain a deeper understanding of the evolutionary origin of the sequence SaTLP-4, we conducted a CLANS analysis on a comprehensive dataset of 504 retrieved sequences in FASTA format. This dataset included the two longest contigs coding for SaTLP-4 (Comp264488_c0_seq1 and Comp264488_c0_seq3), 402 reviewed proteins containing the same collagenic domain as SaTLP-4 (PF01391) identified using the SMART software (sequences retrieved on February 1st, 2024) (Letunic *et al*., 2020), as well as the first 50 reciprocal BLASTp hits of this sequence in the NCBI non-redundant protein sequences (Nr) database and the first 50 top reciprocal BLASTp hits that were restricted to Metazoan sequences. The CLANS analysis was based on sequence similarity using an all-against-all BLASTp with the scoring matrix BLOSUM62 and linkage clustering was performed with a maximum E-value of $1E^{-20}$ to identify coherent clusters (https://toolkit.tuebingen.mpg.de/clans/) (Frickey & Lupas, 2004). In addition to this CLANS analysis, we also conducted a multiple alignment with the top reciprocal hit of the two types of BLASTp searches previously done using the MAFFT algorithm (v7.490) with the automated parameters implemented in the Geneious Prime 2023.2.1 software (Katoh *et al*., 2019). Finally, we looked at the 3D structure of SaTLP-4 using AlphaFold3 (Abramson *et al*., 2024).

Results and discussion

Two types of BLASTp analyzes have been conducted: BLASTp with the translated SaTLP-4 using default parameters, showing that SaTLP-4 shares similarities with bacterial sequences, and BLASTp with the parameters set to only include Metazoa sequences. Both types of BLASTp results have approximately the same coverages (91 to 99% on average), around the same percentage of identities (43.56% and 37.41%), and more or less the same e-value (5e-48 and 8e-41), indicating low similarity and no clear homology of SaTLP-4 to any particular sequence (Table 1). Looking at the domains present in the best reciprocal sequences, we can see that both sequences contain two collagenous domains, unlike SaTLP-4 which contains one from position 198 to 258 amino acids as well as two internal repetitive motifs (Fig. 2A).

Table 1. Results of two types of SaTLP-4 BLASTp search: The first line corresponds to the BLASTp hits of SaTLP-4 in the NCBI non-redundant protein sequences (Nr) database, and the second line also shows the BLASTp hit, but with only Metazoa sequences set in the parameters.

Database	**BLASTp reciprocal hit**					
	Description	**species**	**Coverage**	**ID**	**e-value**	**references**
NCBI non-redundant protein sequences (nr)	DNRLRE domain-containing protein	*Syntrophobotulus glycolicus*	91%	43.59%	5e-48	WP_013625194
Only Metazoa (taxid:33208) sequence in the NCBI non-redundant protein sequences (nr)	collagen alpha-2(I) chain-like	*Procambarus clarkii*	99%	37.41%	8e-41	XP_045588259.1

To investigate this further, a CLANS analysis was conducted with 504 sequences retrieved from both BLASTp analyses and sequences with the same collagen domain as SaTLP-4 obtained from SMART (see methods). This all-against-all BLAST analysis shows that most of the sequences are grouped together in a large cluster, although, we can see that the bacterial sequences are grouped together between all the metazoan sequences. The two SaTLP-4 sequences are close to each other in the CLANS analysis and are placed between the metazoan and bacterial sequences (Fig. 1).

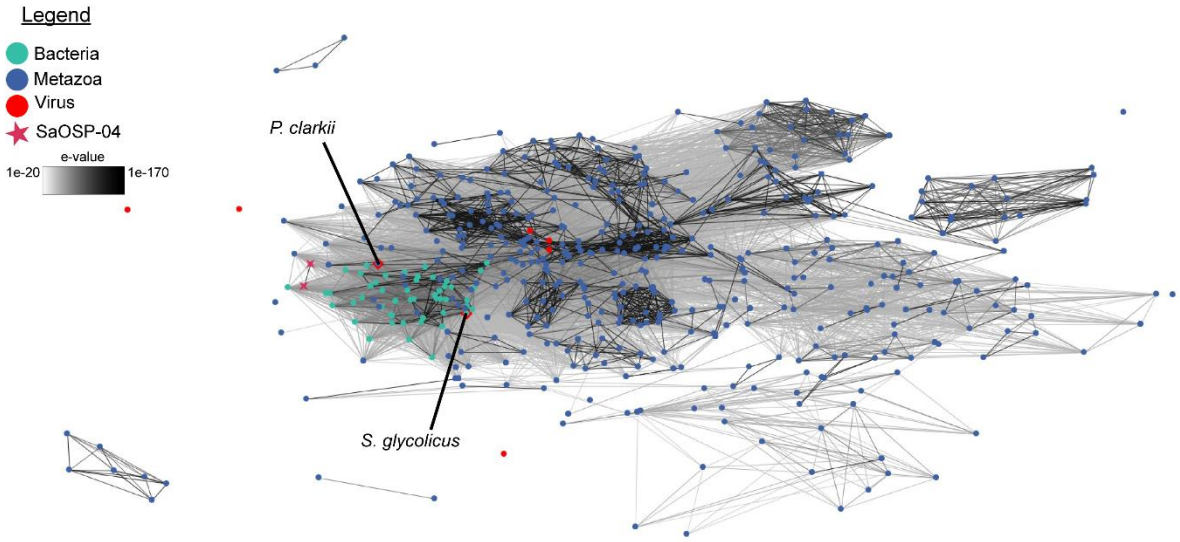


Figure 1. CLANS analysis: A sequence-similarity-based clustering approach was conducted using BLASTp e-values with 504 retrieved sequences from the 50 top hits from both types of BLASTp and sequences having the same collagenic domain as SaTLP-4 in the SMART software.

A MAFFT alignment of SaTLP-4 with the two best-hit reciprocal BLASTp sequences was also conducted, revealing some regions with gaps and only 30% of identity (Fig. 2B).

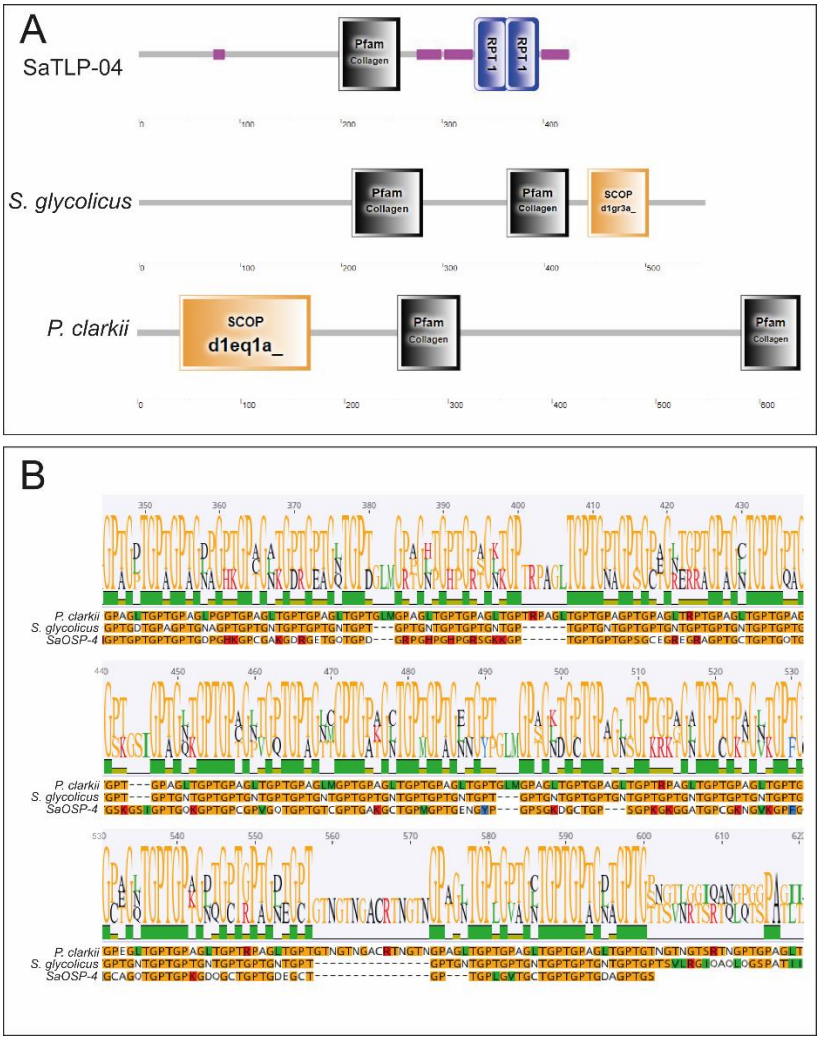


Figure 2. *In silico* analysis of the SaTLP-4 sequence: SMART domains of SaTLP-4 and of both best reciprocal hit BLASTp sequences, showing that they all contain a collagenic domain pfam PF01391 (a SCOP domain is also present in *S. glycolicus* and *P. clarkia* sequence) (A). A part of the MAFFT alignment of SaTLP-4 with the two best BLASTp hit sequences (B).

Finally, to have an idea of the predicted 3D structure to SaTLP-4, we submitted the SaTLP-4 protein sequence in the AlphaFold Database Protein Structure (https://alphafold.ebi.ac.uk) (Jumper *et al*., 2021; Varadi *et al*., 2024). The first hit corresponds to a 539-long uncharacterized amino acid sequence from *Bacillus* sp. X1 (gene number HW35_02320), and the second to another 757-long uncharacterized amino acid sequence from *Amphibacillus xylanus* (gene number AXY_02370). Both sequences have an alpha helix structure at the N-terminus of their sequence. There are (GPT)n motifs from amino acids 103 to 372 in the *Bacillus sp.* sequence, and in the *A. xylanus* sequence, from amino acids 109 to 171 and from 236 to 592. Both have a part corresponding to beta sheet-rich structures in their C-terminal part (Fig. 3A, B).

Along with this, we examined the 3D structure of SaTLP-4 using AlphaFold3. SaTLP-4 contains (GPT)n motifs, although it is less present, with only about ten repetitions scattered throughout the sequence. There is also low confidence in the local structure (Fig. 3C, D). By comparing this structure with the two previous results, we see that all three exhibit this (GPT)n motif, characteristic of bacterial collagen, as well as a beta-sheet structure but in the N-terminal region in SaTLP-4 (Fig. 3).

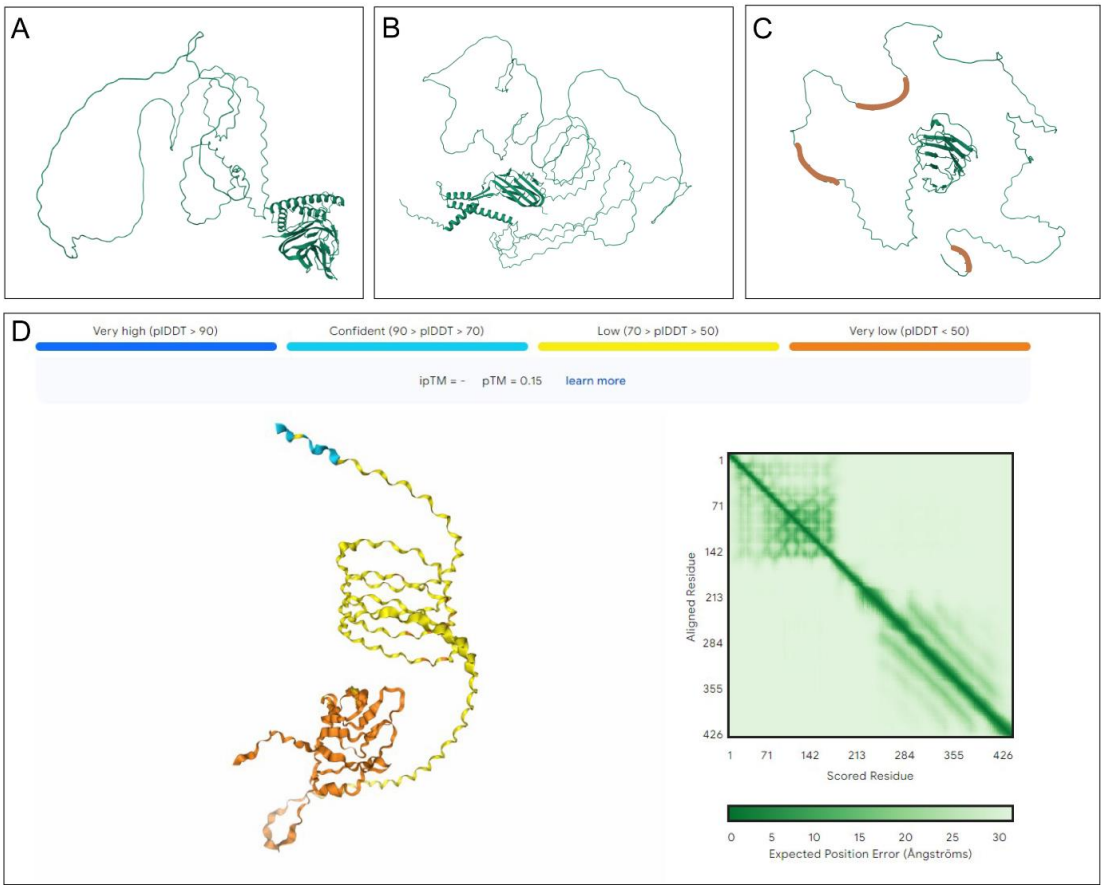


Figure 3. Results of the SaTLP-4 search in the AlphaFold database Protein structure: 3D structure of the first result, the *Bacillus sp*. protein sequence (A), and the second result, the *A. xylanus* protein sequence (B). 3D structure of the SaTLP-4 protein sequence with the (GPT)n motifs highlighted in brown (c), and its AlphaFold prediction showing low confidence in the local structure (D).

All these *in silico* results suggest that SaTLP-4 is a bacterial collagen-like sequence with some similarity (43.59 % of identity) to known bacterial sequences from public databases. This potential homology could then result from a horizontal gene transfer. However, SaTLP-4 is incomplete, and the comparison of its protein sequence with the repetitive motif (here (GPT)n) is difficult to carry out. We hypothesize that several SaTLP-4 proteins can associate together to form the collagen fibres of the tube lining and that the beta sheet-rich structure

found at the N-terminus of each sequence would assemble and act as an anchor domain for other SaTLP-4 proteins within the complex.

General discussion

General discussion

Many marine invertebrate organisms possess the remarkable ability to form strong adhesive bonds, temporary or permanent, that function effectively in wet environments with high ionic strength, such as seawater. This contrasts with man-made adhesive materials, which often fail in the presence of water, salt or slight pH fluctuations (Hofman *et al*., 2018). Underwater adhesion has been extensively studied over the last few decades and is recognized as a complex, multifunctional process involving four key steps: removal of weak boundary layers from the substrate, wetting, establishment of interfacial adhesion and curing (Waite, 1987). In addition, marine invertebrates must adapt to substrates with widely varying chemical and physical properties (Santos *et al*., 2009a).

In 1704, Sir Isaac Newton (cited in Waite, 1983a) once wrote in his Opticks that "there are agents in Nature able to make the particles of joints stick together by a very strong attraction, and it is the business of experimental philosophy to find them out". This thesis aims at improving our knowledge of the fundamental properties of marine adhesive systems, in particular the enzymes involved in post-translational modifications in two of the best-studied organisms: the mussels and the sabellariid tubeworms. This final discussion aims to integrate all the results obtained in the different chapters, starting from a more general point of view and ending with how this research could contribute to the development of biomimetic materials.

Understanding the importance of post-translational modifications in marine adhesion

Chemical adhesion involves bonding two different materials, known as adherents, using an adhesive substance, while cohesion refers to the joining of similar materials (Waite, 1983a). The surface properties of the adherends, and the chemical and physical properties of the adhesive system will determine the strength of adhesion (Waite, 1983a). In marine systems, this strength is often attributed to adhesive proteins and the functional groups formed by their post-translational modifications (PTMs). Post-translational modifications are covalent transformations that change the properties of a protein by proteolytic cleavage or the addition of modifying groups, such as acetyl, glycosyl, methyl and phosphoryl, to one or more amino acids (Ramazi *et al*., 2020; Ramazi & Zahiri, 2021). These PTMs play a fundamental role in protein structure and function (Ramazi *et al*., 2020; Ramazi & Zahiri, 2021). The different physico-chemical interactions between the adhesive and the substrate resulting from these modifications can generate variable forces that influence the type of adhesion (instantaneous, temporary, transitory or permanent). In marine invertebrates, the amino acids targeted by PTMs are mainly, but not limited to, lysine, proline, serine, threonine, tyrosine (see Introduction

section). These PTMs often facilitated by specific enzymes, play a crucial role in optimizing the resistance and functionality of the marine adhesive system. The main known modifications include glycosylation, hydroxylation, phosphorylation, and sulfation (Sagert *et al*., 2006; Hennebert *et al*., 2015a; Hennebert *et al*., 2018; Clarke *et al*., 2020).

Hydroxylation involves the addition of a hydroxyl group on the side chain of several amino acids and is found in all types of adhesion, but particularly in permanent adhesive systems. This type of PTM is often involved in the formation of strong bonds such as covalent and metallic bonds (Introduction – Chapter I). Glycosylation is mainly present in permanent, temporary and transitory adhesion, as demonstrated in barnacle, mussel, sea star and limpet adhesion (Introduction) (Ventura *et al*., 2023). In temporary adhesion, their adhesive cell granules are known to include large cohesive proteins and smaller adhesive proteins, some of which are glycoproteins that segregate in the outer part of the adhesive secretory granules (Wunderer *et al*., 2019; Simão *et al*., 2020; Ventura *et al*., 2023). Although current knowledge about the role of sugars in temporary adhesion is limited, it has been suggested that glycoconjugates contribute to both adhesion and cohesion, possibly through electrostatic interactions and/or lectin-mediated cross-linking (Hennebert *et al*., 2011; Pjeta *et al*., 2019). Phosphorylation is also prevalent in biological adhesion and has been studied in various organisms due to its involvement in protein-protein cross-linking (Sagert & Waite, 2006; Taylor & Wang, 2007; Flammang *et al*., 2009). Additionally, sulfates, along with phosphates, are believed to play a key role in non-covalent adhesive or cohesive interactions, possibly facilitated by Ca^{2+} or Mg^{2+} bridging (Hennebert *et al*., 2015a; Lebesgue *et al*., 2016).

Adhesive secretions are produced and secreted by specific organs/structures within the animal where adhesive protein biosynthesis follows a similar process in all organisms (Li *et al*., 2021). The post-translational modifications of newly synthesized adhesive proteins occur during their transport in the cells. N-glycosylation begins in the endoplasmic reticulum (ER), followed by oligosaccharide processing in the Golgi apparatus (Hennebert *et al*., 2015a). Other modifications, such as O-glycosylation and phosphorylation, take place in the Golgi apparatus (Hennebert *et al*., 2015a; Lebesgue *et al*., 2016). Disulfide bonds, also found in some marine adhesive systems (Introduction), are formed in the lumen of the endoplasmic reticulum by the enzyme disulfide isomerase, which is abundant in the ER of secretory cells (Freedman, 1989). Hydroxylase enzymes are also associated with the ER membrane (Freedman, 1989) (Fig. 1).

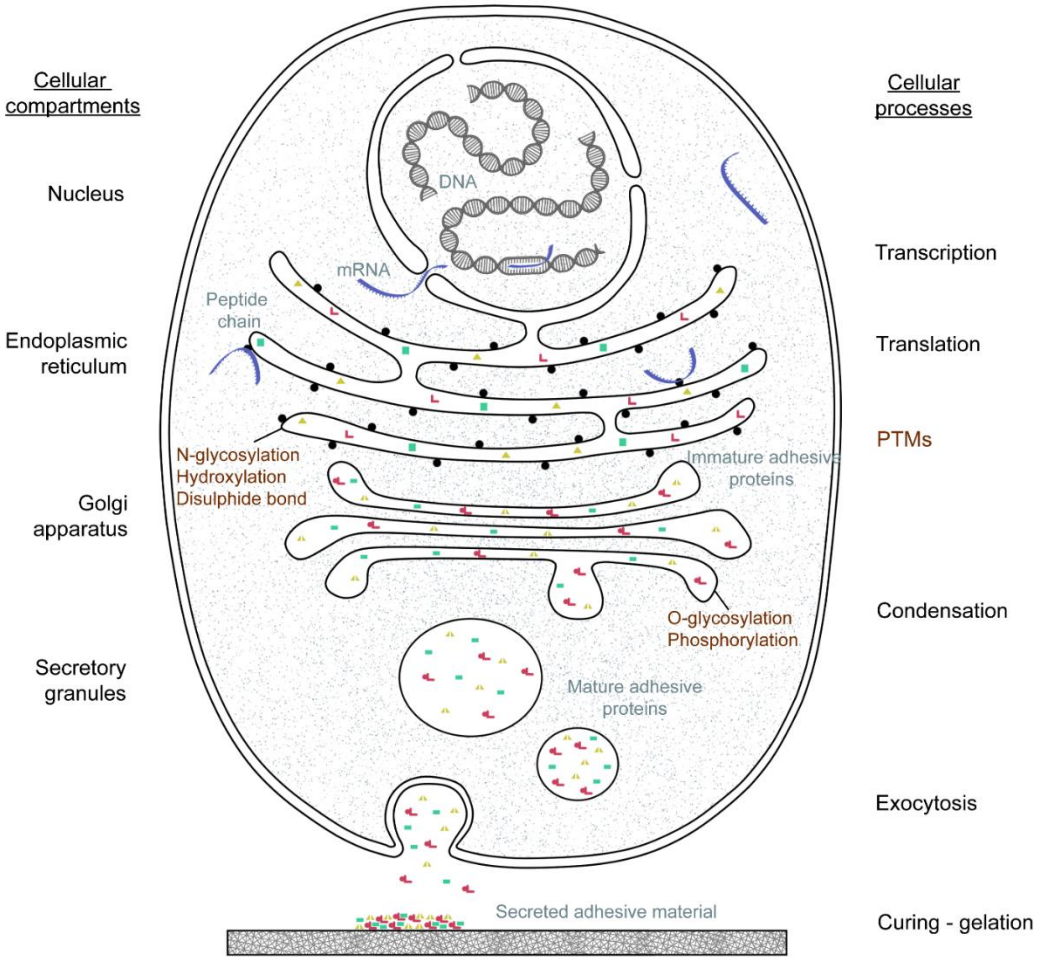


Figure 2. Illustration of the biosynthesis of adhesive proteins in the cell (not to scale) with an indication of the location where the various post-translational modifications take place (inspired by Hennebert *et al*., 2015a; Li *et al*., 2021).

Once secreted, hydrolases, another type of enzyme involved in marine adhesion, can act on the secreted adhesive materials. These enzymes, found in the form of peptidases (also known as proteases), can act by breaking peptide bonds and may play a role in the detachment process of temporary and transitory adhesions (Lebesgue *et al*., 2016; Davey *et al*., 2019; Kang *et al*., 2020; Algrain *et al*., 2022).

Tyrosinase: a key enzyme in marine adhesion and cross-linking systems

The main enzyme studied in this thesis is tyrosinase. Tyrosinases are involved in various biological processes in nature (Chapter I), and they facilitate the production of DOPA, which is well known for its role in adhesion, particularly in the adhesive systems of the blue mussel *Mytilus edulis* and the honeycomb worm *Sabellaria alveolata*, as it can form bonds with the substrate and within the adhesive material (Chapter I) (Becker *et al*., 2012; Petrone, 2013; Davey *et al*., 2021). By using proteo-transcriptomic analyses, several candidate tyrosinases

were identified and characterized in both species (Fig. 2). These candidates were then localized by *in situ* hybridization and shown to be either specific to a single gland type (three in the blue mussel) or expressed in different glands (five in the blue mussel and all candidates in the honeycomb worm). The blue mussel candidates also matched the distribution of their homologs within the byssus of *Mytilus coruscus* (Qin *et al*., 2016) and *Perna viridis* (Guerette *et al*., 2013). Similarly, the honeycomb worm tyrosinases matched their homolog in *Phragmatopoma californica* (Wang & Stewart, 2012). These findings highlight a high diversity of tyrosinases in both species, suggesting either the coexistence of different functions in these enzymes or variations in substrate specificity. Furthermore, studies have long questioned the potential convergent evolution of DOPA-based adhesive mechanisms in these two species (Endrizzi & Stewart, 2009; Flammang *et al*., 2009; Stewart *et al*., 2011a; Hofman *et al*., 2018). Indeed, as the adhesive protein sequences of both studied species are short and intrinsically disordered, establishing their phylogenetic relationships is complicated. On the other hand, the phylogenetic analysis of the tyrosinases involved in the maturation of their adhesive proteins revealed independent expansion and parallel evolution, supporting the convergent evolution of their DOPA-based adhesion and answering this long-standing question (Chapter I).

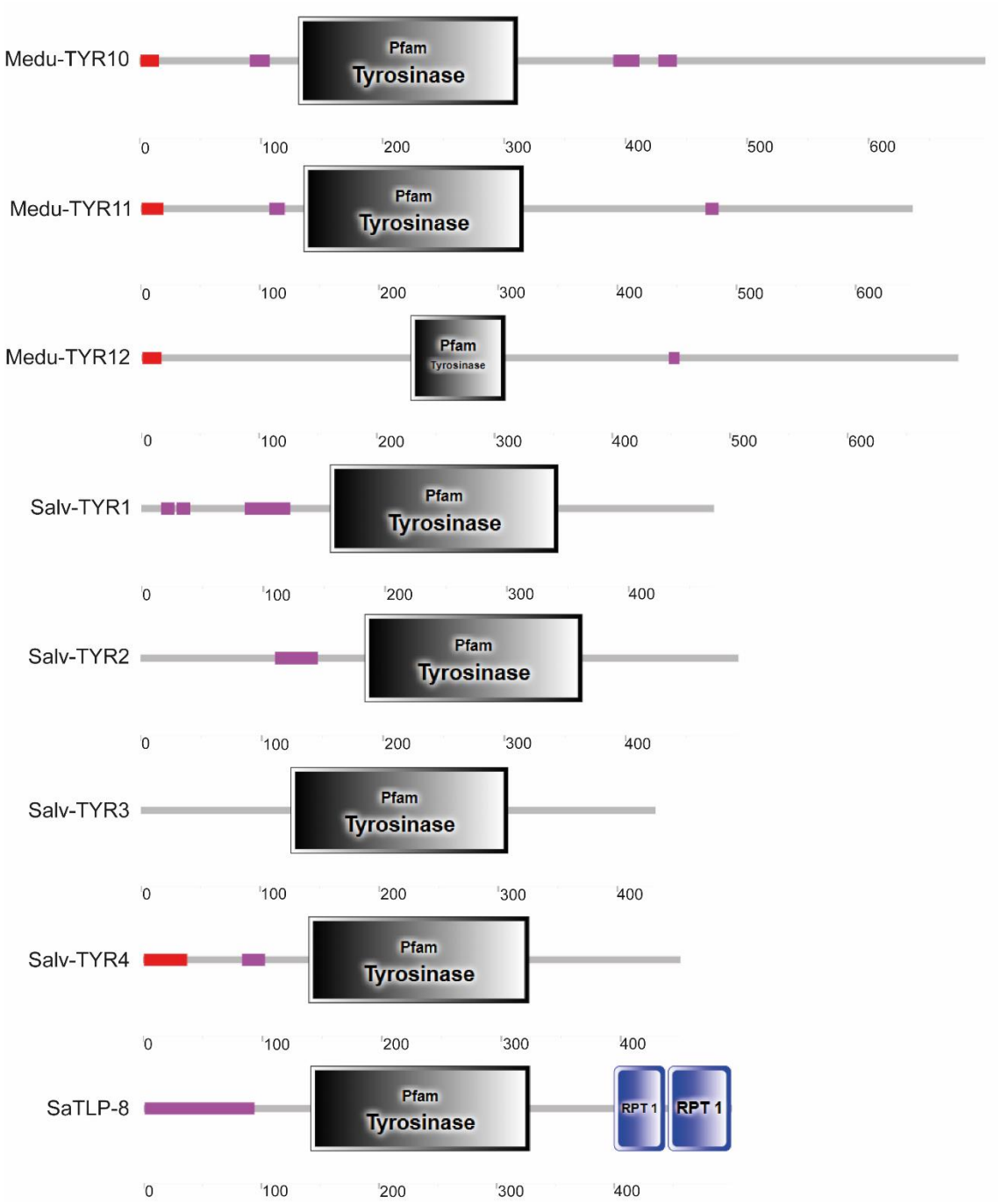


Figure 2. Domain organization of the main tyrosinases identified in both species. The red box indicates the signal peptide, and the purple box highlights the low-complexity region.

To better understand the role of all tyrosinases involved in adhesive protein maturation in mussels and tubeworms, one can search for homology with enzymes of known function or expression patterns. One can also attempt to recombinantly express some candidates, e.g., Medu-TYR1, which was the most highly expressed *M. edulis* tyrosinase candidate in proteomic analyses, or tyrosinases showing gland specificity in blue mussels (i.e., Medu-TYR10 to -12). For the honeycomb worm, this approach could also confirm the potential diphenolase activity

of tyrosinase candidates found in *S. alveolata*, as was observed in *P. californica* (Wang & Stewart, 2013). A proteomic analysis of the plaque gland vesicles identified the candidate Medu-TYR12, reinforcing the proof of its involvement in the maturation of plaque adhesive proteins in the blue mussel (Mathieu Rivard, unpublished data). Medu-TYR12 is the only tyrosinase showing plaque gland specificity. Therefore, this candidate was selected to test its function; we expressed this protein recombinantly and attempted to assess its enzymatic activity using UV-Vis spectrophotometry. However, we were unable to fully produce it, as it tended to precipitate during the refolding process. Following this, several approaches to improve the refolding process have been proposed for future work (Chapter II). Moreover, other approaches, such as enzymatic tests through substrate and inhibitor profiling, could also be conducted to test the enzymatic activity of these tyrosinases (Wang & Stewart, 2013). In addition, we have also demonstrated in this work that a tyrosinase may be involved in the production and/or oxidation of the catechols that cross-link the tube lining in honeycomb worms (Fig. 2) (Chapter IV).

Another important enzyme: Kinase

The second enzyme that has been investigated during this thesis work is the one responsible for phosphoserine production in the honeycomb worm: serine kinase or, as proposed during my work thesis, kinase FAM20C enzyme (Chapter III). This is the protein kinase that phosphorylates the S-x-E/pS motifs within secreted proteins (Tagliabracci *et al*., 2012), a motif that is similar to the poly(S) repeats of the adhesive protein Sa-3 (Becker et al., 2012). FAM20C kinases are part of the FAM20 family which are also involved in diverse biological functions and were shown to generate most of the secreted phosphoproteome in humans (Nalbant *et al*., 2005; Hao *et al*., 2007; Tagliabracci *et al*., 2013a; Tagliabracci *et al*., 2015; Du *et al*., 2018). Several FAM20C kinase candidates were identified thanks to homology searches and were localized in both types of adhesive cement glands as shown by *in situ* hybridization, but also in other body parts of the worm as shown by PCR test (Chapter III).

FAM20C kinase may also be the enzyme responsible for synthesis of the phosphoserine residues found in the mussel foot protein mfp-5 (Waite & Qin, 2001; Zhao & Waite, 2006). In general, kinases appear to be present in all types of adhesion (Introduction). For instance, the adhesive system of the larva of the barnacle *Amphibalanus amphitrite* (Darwin, 1854) is composed of a biphasic system composed of phosphoproteins and lipids (Gohad *et al*., 2014). In temporary adhesion, it has been proposed that proteins with serine/threonine kinase and

phosphorylase activity regulate the exocytosis of adhesive and de-adhesive granules, while also modulating this process via protein phosphorylation and dephosphorylation (Lebesgue *et al*., 2016). Phosphorylation sites were also identified in the adhesive system of sea cucumbers, suggesting that they may also contain kinases (Flammang *et al*., 2009). An in-depth study is needed to determine the type of kinases present in all these organisms and if there is any potential homology (between mussels and tubeworms, for example).

Other important components in marine adhesion

Alongside post-translational modifications (PTMs) and enzymes that contribute to the maturation of adhesive proteins, adhesives also comprise a complex mixture of biomacromolecules, such as lipids and carbohydrates, which play important functional roles (Ga *et al*., 2017; Li *et al*., 2021). Lipids are particularly important in some species, as they create favorable conditions for the spreading of adhesive proteins on substrates and their protection from biodegradation (Fears *et al*., 2018; Gohad *et al*., 2014). Moreover, the presence of proteins and carbohydrates appears to be a common feature of non-permanent adhesives in marine invertebrates, as these complexes tend to form highly hydrated, viscoelastic adhesives (Flammang *et al*., 1998; Santos *et al*., 2009b). In fact, the adhesive proteins of most marine invertebrate species, found in the same type of adhesion, often share similar features such as domains, particularly those associated with temporary adhesion, like the von Willebrand factor domain, C-type or galactose-binding lectin domain, trypsin-inhibitor domains, and multiple EGF-like domains (Davey *et al*., 2021; Bertemes *et al*., 2022; Ventura *et al*., 2023). In transitory adhesion, protein domains are also observed in limpets and sea anemones, notably proteins with protease inhibitor and metallopeptidase domains (Kang *et al*., 2020). In general, both transitory and temporary adhesive proteins tend to be large, multi-domain sequences, often featuring duplicated domains believed to enhance protein–ligand interactions and include numerous glycosylated proteins (Kang *et al*., 2020). Additionally, specific amino acids such as glycine, tyrosine, and cysteine are frequently over-represented in marine adhesive proteins. These amino acids are characteristic of adhesive proteins like barnacle cement and mussel foot proteins (Li *et al*., 2021).

Although some processes linked to adhesion, including protein biosynthesis and secretion, are similar across organisms, others such as the composition, structure and molecular functions of adhesive proteins may differ the type of adhesion considered (Introduction) (Li *et al*., 2021). It has been shown that adhesive proteins can be highly specific to taxonomic groups,

and even species-specific, while exhibiting distinct structural features (Li *et al*., 2021). For example, several differences have already been shown in the adhesive system of the two closely related tubeworm species of the family Sabellariidae: the Californian species *P. californica* and the European tubeworm species *S. alveolata* (Hennebert *et al*., 2015b; 2018). In this thesis an ultrastructural study of the morphology of the cement glands, as well as on the cement dots that attach two tube particles together was carried out. It showed similarities between the two species in terms of the morphology of the cement glands and the structure of the cement dots. However, using elemental composition analyses (EDX) of these structures in *S. alveolata*, differences were noted with the Californian species (chapter III).

During my work, we also aimed to better characterize and understand the adhesion mechanisms of *S. alveolata*. We first performed protein extraction on several tubes reconstructed by the honeycomb worm using glass beads; however, only a few proteins were detected, and no cement proteins were found, suggesting that the proteins could not be extracted due to their high degree of polymerization (Chapter I). Therefore, a proteomic analysis was carried out on the granules extracted from the cement glands, allowing the identification of 2,685 proteins potentially involved in its adhesive system (Chapter III). Many new proteins were identified and categorized according to the domain(s) found in their sequences (Chapter III). Among these proteins, histones, glycoproteins, sequences containing cysteine-rich domains, proteins involved in calcium binding and proteins with a potential cement-protecting role may be involved in the *S. alveolata* adhesive system. However, only two of the known adhesive proteins from *S. alveolata* were retrieved. The significant number of cellular debris observed during granule extraction suggests that many of the 2,685 identified proteins might be contaminants; what which could be confirmed by filtering the data using the differential expression analysis of the proteins between the parathoracic part of the worm and the rest of its body.

These protein extractions show that the adhesion of the honeycomb worm is complex and that extracting its proteins is challenging. However, we have demonstrated that the worm indeed uses a complex polyprotein-based cement, primarily relying on DOPA- and phosphoserine-based adhesion, similar to the adhesive system of mussels (Chapters I and III). Other extraction buffers could be used to extract the cement protein, such as the one used for barnacles, which contains a high concentration of denaturing and reducing agents, as well as requiring heat treatment (their reduction step was performed with 0.5 M of DTT for one hour at 60 °C, instead of using 12.5 mM DTE for 25 min at 56 °C in our study) (Kamino *et al*.,

2000). The adhesive system of barnacles has been shown to differ from that of all other permanently gluing marine animals studied to date (Whittington & Cribb, 2001; Flammang *et al*., 2005; Jonker *et al*., 2012). However, cysteine-rich proteins have been identified in barnacle cement (Kamino, 2001), and three cysteine-rich proteins were also found in the proteome of *S. alveolata* granules (Chapter III), as well as in mussel adhesion (Jehle *et al*., 2020). Moreover, one of the barnacle cement proteins was found to be glycosylated (Kamino *et al*., 2012), and we have shown the potential involvement of glycosylation in the *S. alveolata* adhesive system (Chapter III), suggesting that the adhesion mechanisms of barnacles and tubeworms may involve similar compounds. Further studies are needed to confirm this.

Looking beyond marine adhesives *sensu stricto*, we have also studied the tube lining that covers the inside surface of the tube of the honeycomb worms (Chapter IV). The tube lining is a highly complex fibrous material covering the inner surface of the tube in these organisms (Vovelle, 1965). This organic lining could serve as protecting the worm against abrasion from its tube, facilitating the worm movement within its tube, or acting as a scaffold to maintain the integrity of the tube and prevent fractures. It was previously shown that this organic sheath is secreted by specific glandular cells located on the ventral surface of the three metamers of the parathoracic part of the worm, this part forming a kind of ventral shield (Vovelle, 1965). Vovelle's histochemical study described an acidophilic secretion containing ortho-diphenolic compounds (Vovelle, 1965). He also described two other types of glands: one secreting mucoprotein-like substances with sulfhydryl groups, and the other producing acidic mucopolysaccharides (Vovelle, 1965).

In our work, a histological study of the parathoracic shield revealed collagen-secreting glands in all metameres, acidic mucopolysaccharide-secreting glands under the epidermis of the second and third metameres, and catechol-secreting glands predominantly in the second and third metameres. An ultrastructural study of this region was also carried out, identifying five distinct types of cellular glands. Proteo-transcriptomic analyses identified eight major proteins, including one containing a tyrosinase domain. Interestingly, most of these identified proteins contain no known conserved functional domains and share no homology with proteins in public databases. The most abundant candidates have been localized, confirming their involvement in tube lining production. This tube lining is a common feature of most tubeworm species and its similar structure in different species suggests that it can potentially derive from a common ancestor. However, we could not find any homologues to the tube lining protein we identified, including in tubeworm genomes (Chapter IV). This indicates that, despite their

structural similarity, tube linings from different polychaete species might not be homologous from a molecular point of view.

Biomimetics inspiration and perspectives

Adhesive proteins have become crucial candidates for the development of high-performance materials in biomedicine and biomimetics (Wunderer *et al*., 2019). Over the past 15 years, various types of artificial adhesive proteins have been successfully developed by mimicking marine bioadhesion, improving the performance of biomimetic materials (Cui *et al*., 2017). The main components of the permanent adhesion systems of mussels, sabellariid worms and barnacles have been the most studied to inspire the development of biomimetic approaches (Liu *et al*., 2024). For example, mussel adhesion has inspired the development of DOPA-based adhesives, through the incorporation of DOPA as an adhesive component in polymeric materials, which have been shown to be both practical and effective (Jenkins *et al*., 2017; Lee *et al*., 2020). Another example of biomimetic inspiration is the use of the coacervation process involved in the underwater adhesion of worms of the family Sabellariidae (Shao & Stewart, 2010; Zhao *et al*., 2016; Dompé *et al*., 2019). Research has also been conducted to recombinantly express the nano-fibrous protein present in barnacle cement (Liang *et al*., 2018; Estrella *et al*., 2021). However, a better understanding of the fundamental basis of their adhesive systems, including the enzymes involved in the maturation of the adhesive protein, could only lead to improvements in these materials and advances in the design of new types of biomimetic adhesive materials.

Permanently attaching organisms such as mussels, ascidians, and barnacles are also known to be biofouling species. Ship hull fouling has a significant economic impact and contributes to a loss of cruising efficiency for ships (Schultz, 2007; Schultz *et al*., 2011). Here also, a better understanding of the fundamental mechanisms of biological adhesion, including the enzymes involved, may inspire the design of new biomimetic materials as well as new coating technologies to combat marine biofouling (Callow & Callow, 2011). In addition, marine enzymes are becoming an increasingly attractive source in the industrial world, with marine enzymes such as oxidoreductases, hydrolases, transferases, isomerases, ligases and lyases being useful for food and pharmaceutical applications (Lima & Porto, 2016; Trincone, 2017). Knowledge of these biocatalysts and their habitat-related properties, such as salt tolerance, hyperthermostability, barophilicity and cold adaptability, is necessary for exploitation in some of their industrial bioprocesses (Trincone, 2010; Trincone, 2017).

Finally, collagen has become an appealing biomaterial due to its biodegradability, resistance to proteolysis, low antigenicity, minimal inflammatory response, and its ability to promote cell attachment, growth, and, ultimately, tissue healing and regeneration (Paul & Bailey, 2003; Friess, 1998; Lynn *et al*., 2004; Zeugolis *et al*., 2009). We have shown that the collagen that constitutes the tube lining produced by the honeycomb worm appears to have a composition similar to bacterial collagen (Chapter IV). Bacterial collagen is useful for large-scale collagen production and biomedical applications, serving as a prototype for the engineering of new collagen-based biomaterials (An *et al*., 2014). Moreover, bacterial collagen from *Streptococcus pyogenes* (Scl2 protein) has been shown to be non-immunogenic, non-cytotoxic and non-thrombogenic (Peng *et al*., 2010). This tubeworm particular collagen could therefore become a new source of inspiration for potential biomimetic applications.

General discussion

References

References

Abbineni, P. S., Hibbert, J. E., & Coorssen, J. R. (2013). Critical Role of Cortical Vesicles in Dissecting Regulated Exocytosis: Overview of Insights Into Fundamental Molecular Mechanisms. *The Biological Bulletin*, 224(3), 200-217.

Abbott, D. P., & Reish, D. J. (1980). Polychaeta: the marine annelid worms. *Intertidal invertebrates of California. Stanford University Press, Stanford, California*, 448-489.

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493-500.

Abueg, L. A. L., Afgan, E., Allart, O., Awan, A. H., Bacon, W. A., Baker, D., Bassetti, M., Batut, B., Bernt, M., Blankenberg, D., Bombarely, A., Bretaudeau, A., Bromhead, C. J., Burke, M. L., Capon, P. K., Čech, M., Chavero-Díez, M., Chilton, J. M., … Zoabi, R. (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Research*, 52(W1), W83-W94.

Adams, P. J. M., & Tyler, S. (1980). Hopping locomotion in a nematode: Functional anatomy of the caudal gland apparatus of *Theristus caudasaliens* sp. n. *Journal of Morphology*, 164(3), 265-285.

Agarwal, P., Singh, M., Singh, J., & Singh, R. P. (2019). Microbial Tyrosinases: A Novel Enzyme, Structural Features, and Applications. *In Applied Microbiology and Bioengineering* (p. 3-19). Elsevier.

Aguilera, F., McDougall, C., & Degnan, B. M. (2013). Origin, evolution and classification of type-3 copper proteins: Lineage-specific gene expansions and losses across the Metazoa. *BMC Evolutionary Biology*, 13(1), 96.

Aguilera, F., McDougall, C., & Degnan, B. M. (2014). Evolution of the tyrosinase gene family in bivalve molluscs: Independent expansion of the mantle gene repertoire. *Acta Biomaterialia*, 10(9), 3855 3865.

Akbari, N., Khajeh, K., Ghaemi, N., & Salemi, Z. (2010). Efficient refolding of recombinant lipase from *Escherichia coli* inclusion bodies by response surface methodology. *Protein Expression and Purification*, 70(2), 254-259.

Aldred, N., & Clare, A. S. (2008). The adhesive strategies of cyprids and development of barnacle-resistant marine coatings. *Biofouling*, 24(5), 351-363.

Aldred, N., Scardino, A., Cavaco, A., De Nys, R., & Clare, A. S. (2010). Attachment strength is a key factor in the selection of surfaces by barnacle cyprids (*Balanus amphitrite*) during settlement. *Biofouling*, 26(3), 287-299.

Algrain, M., Hennebert, E., Bertemes, P., Wattiez, R., Flammang, P., & Lengerer, B. (2022). In the footsteps of sea stars: Deciphering the catalogue of proteins involved in underwater temporary adhesion. *Open Biology*, *12*(8), 220103.

References

Almeida, M., Reis, R. L., & Silva, T. H. (2020). Marine invertebrates are a source of bioadhesives with biomimetic interest. *Materials Science and Engineering: C*, 108, 110467.

Ameye, L., Hermann, R., DuBois, P., & Flammang, P. (2000). Ultrastructure of the echinoderm cuticle after fast-freezing/freeze substitution and conventional chemical fixations. *Microscopy research and technique*, *48*(6), 385-393.

Amini, S., Kolle, S., Petrone, L., Ahanotu, O., Sunny, S., Sutanto, C. N., ... & Miserez, A. (2017). Preventing mussel adhesion using lubricant-infused materials. *Science*, *357*(6352), 668-673.

An, B., Kaplan, D. L., & Brodsky, B. (2014). Engineered recombinant bacterial collagen as an alternative collagen-based biomaterial for tissue engineering. *Frontiers in Chemistry*, 2.

Andersen, S. O. (2010). Insect cuticular sclerotization: A review. *Insect Biochemistry and Molecular Biology*, 40(3), 166 178.

Arnold, A. A., Byette, F., Séguin-Heine, M.-O., LeBlanc, A., Sleno, L., Tremblay, R., Pellerin, C., & Marcotte, I. (2013). Solid-State NMR Structure Determination of Whole Anchoring Threads from the Blue Mussel *Mytilus edulis. Biomacromolecules*, 14(1), 132-141.

Arnow L.E. (1937). Colorimetric determination of the components of 3,4-dihydroxyphenylalaninetyrosine mixtures. *Journal of Biological Chemistry*. 118: 531-537.

Aumailley, M. (2013). The laminin family. *Cell Adhesion & Migration*, 7(1), 48-55.

Babonis, L. S., DeBiasse, M. B., Francis, W. R., Christianson, L. M., Moss, A. G., Haddock, S. H. D., Martindale, M. Q., & Ryan, J. F. (2018). Integrating embryonic development and evolutionary history to characterize tentacle-specific cell types in a ctenophore. *Molecular Biology and Evolution*. 35(12):2940–2956.

Bairoch, A. (2000). The ENZYME database in 2000. *Nucleic Acids Research*, 28(1), 304-305.

Bargmann. W., Jacob. K., & Rast, A. (1972). Über Tentakel und Colloblasten der Ctenophore Pleurobrachia pileus. Zeitschrift für Zellforschung und mikroskopische Anatomie 123(1): 121–152

Barlow, D. E., Dickinson, G. H., Orihuela, B., Rittschof, D., & Wahl, K. J. (2009). In situ ATR–FTIR characterization of primary cement interfaces of the barnacle Balanus amphitrite. *Biofouling*, *25*(4), 359-366.

Barlow, D. E., Dickinson, G. H., Orihuela, B., Kulp III, J. L., Rittschof, D., & Wahl, K. J. (2010). Characterization of the adhesive plaque of the barnacle Balanus amphitrite: amyloid-like nanofibrils are a major component. *Langmuir*, *26*(9), 6549-6556.

References

Bartolomaeus, T., Purschke, G., & Hausen, H. (2005). Polychaete phylogeny based on morphological data – a comparison of current attempts. *Hydrobiologia*, 535-536(1), 341-356.

Becker, P. T., & Flammang, P. (2010). Unravelling the sticky threads of sea cucumbers—A comparative study on Cuvierian tubule morphology and histochemistry (pp. 87-98). Springer Vienna.

Becker, P. T., Lambert, A., Lejeune, A., Lanterbecq, D., & Flammang, P. (2012). Identification, characterization, and expression levels of putative adhesive proteins from the tube-dwelling polychaete *Sabellaria alveolata*. *The Biological Bulletin*, *223*(2), 217-225.

Benwitz, G. (1978). Elektronenmikroskopische Untersuchungen der Colloblasten-Entwicklung bei der Ctenophora *Pleurobrachia pileus* (Tentaculifera, Cydippea). *Zoomorphologie* 89:257–278.

Bertemes, P., Grosbusch, A. L., Geschwindt, A., Kauffmann, B., Salvenmoser, W., Mertens, B., Pjeta, R., Egger, B., & Ladurner, P. (2022). Sticking Together an Updated Model for Temporary Adhesion. *Marine Drugs*, *20*(6), 359.

Bessette, P. H., Åslund, F., Beckwith, J., & Georgiou, G. (1999). Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proceedings of the National Academy of Sciences*, 96(24), 13703-13708.

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X.-Z., & Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in *Escherichia coli* and Strategies to Address Them for Industrial Applications. *Frontiers in Bioengineering and Biotechnology*, 9, 630551.

Bleidorn, C., Vogt, L., & Bartolomaeus, T. (2003). New insights into polychaete phylogeny (Annelida) inferred from 18S rDNA sequences. *Molecular Phylogenetics and Evolution*, 29(2), 279-288.

Bonneel, M. (2020). *Sea cucumbers as a source of proteins with biomimetic interest: Adhesive and connective tissue-stiffening proteins from Holothuria forskali* [Thèse de doctorat, Université de Mons]. URL: https://orbi.umons.ac.be/handle/20.500.12907/20066

Borovansky, J., & Riley, P. A. (2011). Melanins and melanosomes: biosynthesis, structure, physiological and pathological functions. *John Wiley & Sons*, pp. 343–381.

Bossemeyer, D. (1994). The glycine-rich sequence of protein kinases: A multifunctional element. *Trends in Biochemical Sciences*, 19(5), 201-205.

Brondyk, W. H. (2009). Chapter 11 Selecting an Appropriate Method for Expressing a Recombinant Protein. *In Methods in Enzymology* (Vol. 463, p. 131-147). Elsevier.

Brown, S. C., & McGee-Russell, S. (1971). Chaetopterus tubes: Ultrastructural architecture. *Tissue and Cell*, 3(1), 65 70.

References

Buffet, J. P., Corre, E., Duvernois-Berthet, E., Fournier, J., & Lopez, P. J. (2018). Adhesive gland transcriptomics uncovers a diversity of genes involved in glue formation in marine tube-building polychaetes. *Acta biomaterialia*, *72*, 316-328.

Burmester, T., & Schellen, K. (1996). Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and dipteran arylphorin receptor. *Journal of Molecular Evolution*, 42(6), 713 728.

Burzio, L. A., & Waite, J. H. (2000). Cross-Linking in Adhesive Quinoproteins: Studies with Model Decapeptides. *Biochemistry*, 39(36), 11147-11153.

Burzio, L. A., & Waite, J. H. (2002). The Other Topa: Formation of 3,4,5-Trihydroxyphenylalanine in Peptides. *Analytical Biochemistry*, 306(1), 108-114.

Cabanes, J., García-Cánovas, F., & García-Carmona, F. (1987). Chemical and enzymatic oxidation of 4-methylcatechol in the presence and absence of l-serine. Spectrophotometric determination of intermediates. *Biochimica et Biophysica Acta (BBA) -Protein Structure and Molecular Enzymology*, 914(2), 190-197.

Cabilly, S. (1989). Growth at sub-optimal temperatures allows the production of functional, antigen-binding Fab fragments in *Escherichia coli*. Gene, 85(2), 553-557.

Callow, J. A., & Callow, M. E. (2011). Trends in the development of environmentally friendly fouling-resistant marine coatings. *Nature Communications*, 2(1), 244.

Capella-Gutierrez, S., Silla-Martinez, J. M., & Gabaldon, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972 1973.

Cárdenas-Moreno, Y., González-Bacerio, J., García Arellano, H., & Del Monte-Martínez, A. (2023). Oxidoreductase enzymes: Characteristics, applications, and challenges as a biocatalyst. *Biotechnology and Applied Biochemistry*, 70(6), 2108-2135.

Cerenius, L., & Söderhäll, K. (2004). The prophenoloxidase-activating system in invertebrates. *Immunological Reviews*, 198(1), 116 126.

Chamoy, L., Nicolaï, M., Ravaux, J., Quennedey, B., Gaill, F., & Delachambre, J. (2001). A Novel Chitin-binding Protein from the Vestimentiferan *Riftia pachyptila* Interacts Specifically with β-Chitin. *Journal of Biological Chemistry*, 276(11), 8051-8058.

Chen, T., Ren, C., Wong, N.-K., Yan, A., Sun, C., Fan, D., Luo, P., Jiang, X., Zhang, L., Ruan, Y., Li, J., Wu, X., Huo, D., Huang, J., Li, X., Wu, F., E, Z., Cheng, C., Zhang, X., … Hu, C. (2023). The *Holothuria leucospilota* genome elucidates sacrificial organ expulsion and bioadhesive trap enriched with amyloid-patterned proteins. *Proceedings of the National Academy of Sciences*, 120(16), e2213512120.

Cheng, J., Li, S., Li, X., Fu, R., Huang, X., & Zhan, A. (2022). Molecular functional analyses of larval adhesion in a highly fouling invasive model ascidian. *Marine Biology*, 169(9), 120.

References

Choi, Y. S., Yang, Y. J., Yang, B., & Cha, H. J. (2012). *In vivo* modification of tyrosine residues in recombinant mussel adhesive protein by tyrosinase co-expression in *Escherichia coli*. *Microbial Cell Factories*, 11(1), 139.

Clancy, S. K., Sodano, A., Cunningham, D. J., Huang, S. S., Zalicki, P. J., Shin, S., & Ahn, B. K. (2016). Marine Bioinspired Underwater Contact Adhesion. *Biomacromolecules*, 17(5), 1869-1874.

Clark, D. J., & Maaløe, O. (1967). DNA replication and the division cycle in *Escherichia coli*. *Journal of Molecular Biology*, 23(1), 99-112.

Clarke, J. L., Davey, P. A., & Aldred, N. (2020). Sea anemones (*Exaiptasia pallida*) use a secreted adhesive and complex pedal disc morphology for surface attachment. *BMC Zoology*, 5(1), 5.

Claus, H., & Decker, H. (2006). Bacterial tyrosinases. *Systematic and Applied Microbiology*, 29(1), 3 14.

Cloney, R. (1977). Larval adhesive organs and metamorphosis in ascidians: I. Fine structure of the everting papillae of *Distaplia occidentalis*. *Cell and Tissue Research*, 183(4).

Coyne, K. J., Qin, X.-X., & Waite, J. H. (1997). Extensible Collagen in Mussel Byssus: A Natural Block Copolymer. *Science*, 277(5333), 1830-1832.

Cranfield, H. J. (1973). A study of the morphology, ultrastructure, and histochemistry of the food of the pediveliger of *Ostrea edulis*. *Marine Biology*, 22(3), 187-202.

Cui, M., Ren, S., Wei, S., Sun, C., & Zhong, C. (2017). Natural and bio-inspired underwater adhesives: Current progress and new perspectives. *APL Materials*, 5(11), 116102.

Curd, A., Pernet, F., Corporeau, C., Delisle, L., Firth, L. B., Nunes, F. L. D., & Dubois, S. F. (2019). Connecting organic to mineral: How the physiological state of an ecosystem-engineer is linked to its habitat structure. *Ecological Indicators*, 98, 49-60.

Daly, J. M. (1973). Behavioural and Secretory Activity during Tube construction by *Platynereis dumerilii* Aud & M. Edw. [Polychaeta: Nereidae]. *Journal of the Marine Biological Association of the United Kingdom*, 53(3), 521-529.

Danner, E. W., Kan, Y., Hammer, M. U., Israelachvili, J. N., & Waite, J. H. (2012). Adhesion of Mussel Foot Protein Mefp-5 to Mica: An Underwater Superglue. *Biochemistry*, *51*(33), 6511-6518.

Davey, P. A., Rodrigues, M., Clarke, J. L., & Aldred, N. (2019). Transcriptional characterisation of the *Exaiptasia pallida* pedal disc. *BMC Genomics*, 20(1), 581.

Davey, P. A., Power, A. M., Santos, R., Bertemes, P., Ladurner, P., Palmowski, P., Clarke, J., Flammang, P., Lengerer, B., Hennebert, E., Rothbächer, U., Pjeta, R., Wunderer, J., Zurovec, M., & Aldred, N. (2021). Omics-based molecular analyses of adhesion by aquatic invertebrates. *Biological Reviews*, 96(3), 1051-1075.

References

Davies, M. S., Jones, H. D., & Hawkins, S. J. (1990). Seasonal variation in the composition of pedal mucus from Patella vulgata L. *Journal of Experimental Marine Biology and Ecology*, 144(2-3), 101-112.

Deias, C., Guido, A., Sanfilippo, R., Apollaro, C., Dominici, R., Cipriani, M., Barca, D., & Vespasiano, G. (2023). Elemental Fractionation in Sabellariidae (Polychaeta) Biocement and Comparison with Seawater Pattern: A New Environmental Proxy in a High-Biodiversity Ecosystem? *Water*, 15(8), 1549.

Del Grosso, C. A., McCarthy, T. W., Clark, C. L., Cloud, J. L., & Wilker, J. J. (2016). Managing redox chemistry to deter marine biological adhesion. *Chemistry of Materials*, *28*(18), 6791-6796.

Del Marmol, V., & Beermann, F. (1996). Tyrosinase and related proteins in mammalian pigmentation. *FEBS Letters*, 381(3), 165 168.

Delroisse, J., Kang, V., Gouveneaux, A., Santos, R., & Flammang, P. (2023). Convergent evolution of attachment mechanisms in aquatic animals. In *Convergent Evolution: Animal Form and Function* (pp. 523-557). Cham: Springer International Publishing.

Demarche, P., Junghanns, C., Nair, R. R., & Agathos, S. N. (2012). Harnessing the power of enzymes for environmental stewardship. *Biotechnology Advances*, *30*(5), 933-953.

DeMartini, D. G., Errico, J. M., Sjoestroem, S., Fenster, A., & Waite, J. H. (2017). A cohort of new adhesive proteins identified from transcriptomic analysis of mussel foot glands. *Journal of the Royal Society Interface*, *14*(131), 20170151.

Demeuldre, M., Chinh Ngo, T., Hennebert, E., Wattiez, R., Leclère, P., & Flammang, P. (2014). Instantaneous adhesion of Cuvierian tubules in the sea cucumber *Holothuria forskali*. *Biointerphases*, *9*(2), 029016.

DeMoor, S., Waite, H. J., Jangoux, M. J., & Flammang, P. J. (2003). Characterization of the Adhesive from Cuvierian Tubules of the Sea Cucumber *Holothuria forskali* (Echinodermata, Holothuroidea). *Marine Biotechnology*, 5(1), 45-57.

Dennell, R. (1958). The hardening of insect cuticles. *Biological Reviews*, 33(2), 178 196.

Dickinson, G. H., Vega, I. E., Wahl, K. J., Orihuela, B., Beyley, V., Rodriguez, E. N., Everett, R. K., Bonaventura, J., & Rittschof, D. (2009). Barnacle cement: A polymerization model based on evolutionary concepts. *Journal of Experimental Biology*, 212(21), 3499-3510.

Dickinson, G. H., Yang, X., Wu, F., Orihuela, B., Rittschof, D., & Beniash, E. (2016). Localization of Phosphoproteins within the Barnacle Adhesive Interface. *The Biological Bulletin*, 230(3), 233-242.

Dolcemascolo, G., Pennati, R., De Bernardi, F., Damiani, F., & Gianguzza, M. (2009). Ultrastructural comparative analysis on the adhesive papillae of the swimming larvae of three ascidian species. *Invertebrate Survival Journal*, *6*(1 (Suppl)), S77-S86.

References

Domínguez-Pérez, D., Almeida, D., Wissing, J., Machado, A. M., Jänsch, L., Castro, L. F., Antunes, A., Vasconcelos, V., Campos, A., & Cunha, I. (2020). The Quantitative Proteome of the Cement and Adhesive Gland of the Pedunculate Barnacle, Pollicipes pollicipes. *International Journal of Molecular Sciences*, *21*(7), 2524.

Domínguez-Pérez, D., Almeida, D., Wissing, J., Machado, A. M., Jänsch, L., Antunes, A., Castro, L. F., Vasconcelos, V., Campos, A., & Cunha, I. (2021). Proteogenomic Characterization of the Cement and Adhesive Gland of the Pelagic Gooseneck Barnacle Lepas anatifera. *International Journal of Molecular Sciences*, *22*(7), 3370.

Dompé, M., Cedano-Serrano, F. J., Heckert, O., Van Den Heuvel, N., Van Der Gucht, J., Tran, Y., Hourdet, D., Creton, C., & Kamperman, M. (2019). Thermoresponsive Complex Coacervate-Based Underwater Adhesive. *Advanced Materials*, 31(21), 1808179.

Donovan, R. S., Robinson, C. W., & Glick, B. R. (1996). Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *Journal of Industrial Microbiology*, 16(3), 145-154.

Drexel, R., Siegmund, S., Schneider, H.-J., Linzen, B., Gielens, C., Préaux, G., Lontie, R., Kellermann, J., & Lottspeich, F. (1987). Complete amino-acid sequence of a functional unit from a molluscan hemocyanin (*Helix pomatia*). *Biological Chemistry Hoppe-Seyler*, 368(1), 617 636.

Du, J., Liu, C., Xu, G., Xie, J., Xie, L., & Zhang, R. (2018). Fam20C participates in the shell formation in the pearl oyster, *Pinctada fucata*. *Scientific Reports*, 8(1), 3563.

Duong-Ly, K. C., & Gabelli, S. B. (2014). Explanatory Chapter: Troubleshooting Recombinant Protein Expression. *In Methods in Enzymology* (Vol. 541, p. 209-229). Elsevier.

Dutta, S., Ware, L. A., Barbosa, A., Ockenhouse, C. F., & Lanar, D. E. (2001). Purification, Characterization, and Immunogenicity of a Disulfide Cross-Linked Plasmodium vivax Vaccine Candidate Antigen, Merozoite Surface Protein 1, Expressed in *Escherichia coli*. *Infection and Immunity*, 69(9), 5464-5470.

Edlund, A. F., & Koehl, M. A. R. (1998). Adhesion and Reattachment of Compound Ascidians to Various Substrata: Weak Glue can Prevent Tissue Damage. *Journal of Experimental Biology*, 201(16), 2397-2402.

Eklöf, J. (2010). Taxonomy and phylogeny of polychaetes. *Department of Zoology, University of Gothenburg.*

Ellis, V. L., Ross, D. M., & Sutton, L. (1969). The pedal disc of the swimming sea anemone Stomphia coccinea during detachment, swimming, and resettlement. *Canadian Journal of Zoology*, 47(3), 333-342.

Endean, R. (1957). The Cuvierian Tubules of *Holothuria leucospilota*. *Journal of Cell Science*, S3-98(44), 455-472.

Endrizzi, B. J., & Stewart, R. J. (2009). Glueomics : An Expression Survey of the Adhesive Gland of the Sandcastle Worm. *The Journal of Adhesion*, 85(8), 546-559.

References

Espín, J. C., Van Leeuwen, J., & Wichers, H. J. (1999). Kinetic Study of the Activation Process of a Latent Mushroom (*Agaricus bisporus*) Tyrosinase by Serine Proteases. *Journal of Agricultural and Food Chemistry*, 47(9), 3509 3517.

Esposito, R., D'Aniello, S., Squarzoni, P., Pezzotti, M. R., Ristoratore, F., & Spagnuolo, A. (2012). New insights into the evolution of metazoan tyrosinase gene family. *PLoS ONE*, 7(4), e35731.

Estrella, L. A., Yates, E. A., Fears, K. P., Schultzhaus, J. N., Ryou, H., Leary, D. H., & So, C. R. (2021). Engineered *Escherichia coli* Biofilms Produce Adhesive Nanomaterials Shaped by a Patterned 43 kDa Barnacle Cement Protein. *Biomacromolecules*, 22(2), 365 373.

Fairhead, M., & Thöny-Meyer, L. (2012). Bacterial tyrosinases: Old enzymes with new relevance to biotechnology. *New Biotechnology*, 29(2), 183 191.

Fauchald, K., & Rouse, G. (1997). Polychaete systematics: Past and present. *Zoologica Scripta*, 26(2), 71-138.

Fears, K. P., Orihuela, B., Rittschof, D., & Wahl, K. J. (2018). Acorn Barnacles Secrete Phase-Separating Fluid to Clear Surfaces Ahead of Cement Deposition. *Advanced Science*, 5(6), 1700762.

Flammang, P., & Jangoux, M. (1993). Functional morphology of coronal and peristomeal podia in *Sphaerechinus granularis* (Echinodermata, Echinoida). *Zoomorphology*, *113*(1), 47-60.

Flammang, P., Demeulenaere, S., & Jangoux, M. (1994). The Role of Podial Secretions in Adhesion in Two Species of Sea Stars (Echinodermata*). The Biological Bulletin*, 187(1), 35-47.

Flammang, P. (1996). Adhesion in echinoderms. *Echinoderm Stud*., 5, 1–60.

Flammang, P., Michel, A., Van Cauwenberge, A., Alexandre, H., & Jangoux, M. (1998). A Study of the Temporary Adhesion of the Podia in the Sea Star *Asterias Rubens* (Echinodermata, Asteroidea) through their Footprints. *Journal of Experimental Biology*, 201(16), 2383-2395.

Flammang, P., Santos, R., & Haesaerts, D. (2005). Echinoderm Adhesive Secretions: From Experimental Characterization to Biotechnological Applications. *In V. Matranga (Éd.), Echinodermata* (Vol. 39, p. 201-220). Springer-Verlag.

Flammang, P., Lambert, A., Bailly, P., & Hennebert, E. (2009). Polyphosphoprotein-Containing Marine Adhesives. *The Journal of Adhesion*, 85(8), 447-464.

Forslund, S. K., Kaduk, M., & Sonnhammer, E. L. (2019). Evolution of protein domain architectures. *Evolutionary genomics: statistical and computational methods*, 469-504.

Franc, J. M. (1978). Organization and function of ctenophore colloblasts: an ultrastructural study. *The Biological Bulletin*, *155*(3), 527-541.

References

Frantz, C., Stewart, K. M., & Weaver, V. M. (2010). The extracellular matrix at a glance. *Journal of Cell Science*, 123(24), 4195-4200.

Freedman, R. B. (1989). Post-translational modification and folding of secreted proteins. *Biochemical Society Transactions*, 17(2), 331 335.

Frickey, T., & Lupas, A. (2004). CLANS: A Java application for visualizing protein families based on pairwise similarity. *Bioinformatics*, 20(18), 3702 3704.

Friess, W. (1998). Collagen – biomaterial for drug delivery. *European Journal of Pharmaceutics and Biopharmaceutics*, 45(2), 113 136.

Fyhn, U. E. H., & Costlow, J. D. (1976). A HISTOCHEMICAL STUDY OF CEMENT SECRETION DURING THE INTERMOLT CYCLE IN BARNACLES. *The Biological Bulletin*, 150(1), 47-56.

Gabe, M. (1968). Techniques histologiques. *Masson et Cie (Eds.), Paris*, 1115.

Gaill, F., & Hunt, S. (1986). Tubes of deep sea hydrothermal vent worms *Riftia pachyptila* (Vestimentifera) and *Alvinella pompejana* (Annelida). *Marine Ecology Progress Series*, 34(3), 267–274.

Galli, C., Guizzardi, S., Passeri, G., Macaluso, G. M., & Scandroglio, R. (2005). Life on the wire: on tensegrity and force balance in cells. *Acta Biomed*, 76(1), 5-12.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein Identification and Analysis Tools on the Expasy Server. *(In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press*. pp. 571-607.

Gautam, S., Dubey, P., Mohmad Rather, G., & Gupta, M. N. (2012). Non-Chromatographic Strategies for Protein Refolding. *Recent Patents on Biotechnology*, 6(1), 57-68.

González-Santoyo, I., & Córdoba-Aguilar, A. (2012). Phenoloxidase : A key component of the insect immune system: Biochemical and evolutionary ecology of PO. *Entomologia Experimentalis et Applicata*, 142(1), 1 16.

Gorb, S. N. (2008). Biological attachment devices: Exploring nature's diversity for biomimetics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1870), 1557-1574.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644 652.

Grenon, J.-F., & Walker, G. (1978). The histology and histochemistry of the pedal glandular system of two limpets, *Patella vulgata* and *Acmaea tessulata* (Gastropoda: Prosobranchia). *Journal of the Marine Biological Association of the United Kingdom*, 58(4), 803-816.

References

Sun, Y. (1972). Aspects morphologiques et dynamiques de constructions de l'Annelide polychete *Sabellaria alveolata* (Linne). *Revue des Travaux de l'Institut des Pêches Maritimes*, 36(2), 131-161.

Ga, S., Na, A., & Sm, H. (2017). Comparative Studies on the Biological Glue of Some Opportunistic Adult Marine Macro-Fouling After Dislodgement and Construction of Temporary Faunal Conglomerations. *Journal of Marine Biology & Oceanography*, 06(01).

Gohad, N. V., Aldred, N., Hartshorn, C. M., Jong Lee, Y., Cicerone, M. T., Orihuela, B., Clare, A. S., Rittschof, D., & Mount, A. S. (2014). Synergistic roles for lipids and proteins in the permanent adhesive of barnacle larvae. *Nature Communications*, 5(1), 4414.

Gruet, Y., Vovelle, J., & Grasset, M. (1987). Composante biominérale du ciment du tube chez *Sabellaria alveolata* (L.), Annélide Polychète. *Canadian Journal of Zoology*, 65(4), 837-842.

Guerette, P. A., Hoon, S., Seow, Y., Raida, M., Masic, A., Wong, F. T., Ho, V. H. B., Kong, K. W., Demirel, M. C., Pena-Francesch, A., Amini, S., Tay, G. Z., Ding, D., & Miserez, A. (2013). Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science. *Nature Biotechnology*, 31(10), 908-915.

Hamel, J., & Mercier, A. (2000). Cuvierian tubules in tropical holothurians: Usefulness and efficiency as a defence mechanism. *Marine and Freshwater Behaviour and Physiology*, 33(2), 115-139.

Hao, J., Narayanan, K., Muni, T., Ramachandran, A., & George, A. (2007). Dentin Matrix Protein 4, a Novel Secretory Calcium-binding Protein That Modulates Odontoblast Differentiation. *Journal of Biological Chemistry*, 282(21), 15357-15365.

Harrington, M. J., Masic, A., Holten-Andersen, N., Waite, J. H., & Fratzl, P. (2010). Iron-Clad Fibers: A Metal-Based Biological Strategy for Hard Flexible Coatings. *Science*, 328(5975), 216-220.

Harrington, M. J., Jehle, F., & Priemel, T. (2018). Mussel Byssus Structure-Function and Fabrication as Inspiration for Biotechnological Production of Advanced Materials. *Biotechnology Journal*, 13(12), 1800133.

Harrison, P. L., & Booth, D. J. (2007). Coral reefs: naturally dynamic and increasingly disturbed ecosystems. *Marine ecology*, 316-377.

He, L.-S., Zhang, G., Wang, Y., Yan, G.-Y., & Qian, P.-Y. (2018). Toward understanding barnacle cementing by characterization of one cement protein-100kDa in *Amphibalanus amphitrite*. *Biochemical and Biophysical Research Communications*, 495(1), 969-975.

Hellio, C., Bourgougnon, N., & Gal, Y. L. (2000). Phenoloxidase (E.C. 1.14.18.1) from the byssus gland of *Mytilus edulis*: Purification, partial characterization and application for screening products with potential antifouling activities. *Biofouling*, 16(2 4), 235 244.

References

Hennebert, E., Viville, P., Lazzaroni, R., & Flammang, P. (2008). Micro- and nanostructure of the adhesive material secreted by the tube feet of the sea star *Asterias rubens*. *Journal of Structural Biology*, 164(1), 108-118.

Hennebert, E., Wattiez, R., & Flammang, P. (2011). Characterisation of the Carbohydrate Fraction of the Temporary Adhesive Secreted by the Tube Feet of the Sea Star *Asterias rubens*. *Marine Biotechnology*, 13(3), 484-495.

Hennebert, E., Wattiez, R., Demeuldre, M., Ladurner, P., Hwang, D. S., Waite, J. H., & Flammang, P. (2014). Sea star tenacity mediated by a protein that fragments, then aggregates. *Proceedings of the National Academy of Sciences*, 111(17), 6317-6322.

Hennebert, E., Maldonado, B., Ladurner, P., Flammang, P., & Santos, R. (2015a). Experimental strategies for the identification and characterization of adhesive proteins in animals: A review. *Interface Focus*, *5*(1), 20140064.

Hennebert E., Maldonado B., Van De Weerdt C., Demeuldre M., Richter K., Rischka K., Patrick Flammang P. (2015b). From sand tube to test tube: The adhesive selection from Sabellariid Tubeworms. *Bioadhesion and biomimetics: from nature to applications. Singapore: Pan Stanford Publishing*.

Hennebert, E., Leroy, B., Wattiez, R., & Ladurner, P. (2015c). An integrated transcriptomic and proteomic analysis of sea star epidermal secretions identifies proteins involved in defense and adhesion. *Journal of Proteomics*, 128, 83-91.

Hennebert, E., Gregorowicz, E., & Flammang, P. (2018). Involvement of sulfated biopolymers in adhesive secretions produced by marine invertebrates. *Biology Open*, bio.037358.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2 : Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518 522.

Hofman, A. H., van Hees, I. A., Yang, J., & Kamperman, M. (2018). Bioinspired underwater adhesives by using the supramolecular toolbox. Advanced Materials, 30(19), 1704640.

Holliday, G. L., Rahman, S. A., Furnham, N., & Thornton, J. M. (2014). Exploring the Biological and Chemical Complexity of the Ligases. *Journal of Molecular Biology*, 426(10), 2098-2111.

Holten-Andersen, N., Fantner, G. E., Hohlbauch, S., Waite, J. H., & Zok, F. W. (2007). Protective coatings on extensible biofibres. *Nature Materials*, 6(9), 669-672.

Holten-Andersen, N., Mates, T. E., Toprak, M. S., Stucky, G. D., Zok, F. W., & Waite, J. H. (2009). Metals and the Integrity of a Biological Coating: The Cuticle of Mussel Byssus. *Langmuir*, 25(6), 3323-3326.

Hopkins, T. L., Starkey, S. R., Xu, R., Merritt, M. E., Schaefer, J., & Kramer, K. J. (1999). Catechols involved in sclerotization of cuticle and egg pods of the grasshopper, *Melanoplus sanguinipes*, and their interactions with cuticular proteins. *Archives of Insect Biochemistry and Physiology*, 40(3), 119 128.

References

Hoyle, G. (1960). Neuromuscular Activity in the Swimming Sea Anemone, *Stomphia Coccinea* (Muller)*. *Journal of Experimental Biology*, 37(4), 671-688.

Huan, P., Liu, G., Wang, H., & Liu, B. (2013). Identification of a tyrosinase gene potentially involved in early larval shell biogenesis of the Pacific oyster *Crassostrea gigas*. *Development Genes and Evolution*, 223(6), 389 394.

Hwang, D. S., Yoo, H. J., Jun, J. H., Moon, W. K., & Cha, H. J. (2004). Expression of functional recombinant mussel adhesive protein Mgfp-5 in *Escherichia coli*. *Applied and Environmental Microbiology*, 70(6), 3352 3359.

Hwang, D. S., Y. Gim and H. J. Cha (2005). Expression of functional recombinant mussel adhesive protein type 3A in *Escherichia coli*. *Biotechnol Prog*, 21(3): 965-970.

Inoue, K., Takeuchi, Y., Miki, D., Odo, S., Harayama, S., & Waite, J. H. (1996). Cloning, sequencing and sites of expression of genes for the hydroxyarginine-containing adhesive-plaque protein of the mussel *Mytilus galloprovincialis*. *European journal of biochemistry*, *239*(1), 172-176.

Inoue, K., Yoshioka, Y., Tanaka, H., Kinjo, A., Sassa, M., Ueda, I., Shinzato, C., Toyoda, A., & Itoh, T. (2021). Genomics and transcriptomics of the green mussel explain the durability of its byssus. *Scientific Reports*, 11(1), 5992.

Ito, L., Shiraki, K., Makino, M., Hasegawa, K., & Kumasaka, T. (2011). Glycine amide shielding on the aromatic surfaces of lysozyme: Implication for suppression of protein aggregation. *FEBS Letters*, 585(3), 555-560.

Jaenicke, E., & Decker, H. (2003). Tyrosinases from crustaceans form hexamers. *Biochemical Journal*, 371(2), 515 523.

Jehle, F., Macías-Sánchez, E., Sviben, S., Fratzl, P., Bertinetti, L., & Harrington, M. J. (2020). Hierarchically-structured metalloprotein composite coatings biofabricated from co-existing condensed liquid phases. *Nature Communications*, 11(1), 862.

Jenkins, C. L., Siebert, H. M., & Wilker, J. J. (2017). Integrating Mussel Chemistry into a Bio-Based Polymer to Create Degradable Adhesives. *Macromolecules*, 50(2), 561 568.

Jensen, R. and Morse, D. (1988). The bioadhesive of *Phragmatopoma californica* tubes: a silk-like cement containing L-DOPA. *J Comp Physiol B*. 158:317-324.

Johnson, K. A., & Goody, R. S. (2011). The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry*, 50(39), 8264-8269.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5 : Genome-scale protein function classification. *Bioinformatics*, 30(9), 1236-1240.

References

Jonker, J.-L., Von Byern, J., Flammang, P., Klepal, W., & Power, A. M. (2012). Unusual adhesive production system in the barnacle *Lepas anatifera* : An ultrastructural and histochemical investigation. *Journal of Morphology*, 273(12), 1377-1391.

Jonker, J.-L., Abram, F., Pires, E., Varela Coelho, A., Grunwald, I., & Power, A. M. (2014). Adhesive Proteins of Stalked and Acorn Barnacles Display Homology with Low Sequence Similarities. *PLoS ONE*, 9(10), e108902.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.

Junqueira, L. C. U., Bignolas, G., & Brentani, R. R. (1979). Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections. *The Histochemical Journal*, 11(4), 447 455.

Kamino, K., Odo, S., & Maruyama, T. (1996). Cement Proteins of the Acorn-Barnacle, *Megabalanus rosa*. *The Biological Bulletin*, 190(3), 403-409.

Kamino, K., Inoue, K., Maruyama, T., Takamatsu, N., Harayama, S., & Shizuri, Y. (2000). Barnacle Cement Proteins. *Journal of Biological Chemistry*, *275*(35), 27360-27365.

Kamino, K. (2001). Novel barnacle underwater adhesive protein is a charged amino acid-rich protein constituted by a Cys-rich repetitive sequence. *Biochemical Journal*, *356*(2), 503-507.

Kamino, K. (2006). Barnacle underwater adhesive: complexity from multi-functionality in a multi-protein complex. *Biomaterials from Aquatic and Terrestrial Organisms*, 549-570.

Kamino, K. (2008). Underwater Adhesive of Marine Organisms as the Vital Link Between Biological Science and Material Science. *Marine Biotechnology*, 10(2), 111-121.

Kamino, K. (2010). Molecular design of barnacle cement in comparison with those of mussel and tubeworm. *The Journal of Adhesion*, 86(1), 96 110.

Kamino, K., Nakano, M., & Kanai, S. (2012). Significance of the conformation of building blocks in curing of barnacle underwater adhesive. *The FEBS Journal*, 279(10), 1750-1760.

Kamino, K. (2016). Barnacle Underwater Attachment. In A. M. Smith (Éd.), Biological Adhesives (p. 153-176). Springer International Publishing.

Kang, V., Lengerer, B., Wattiez, R., & Flammang, P. (2020). Molecular insights into the powerful mucus-based adhesion of limpets (*Patella vulgata* L.). *Open Biology*, 10(6), 200019.

References

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20(4), 1160 1166.

Kitamura, M., Kawakami, K., Nakamura, N., Tsumoto, K., Uchiyama, H., Ueda, Y., Kumagai, I., & Nakaya, T. (1999). Expression of a model peptide of a marine mussel adhesive protein in *Escherichia coli* and characterization of its structural and functional properties. *Journal of Polymer Science Part A: Polymer Chemistry*, 37(6), 729 736.

Kongsrud, J. A., & Rapp, H. T. (2012). *Nicomache (Loxochona) lokii sp.* nov. (Annelida: Polychaeta : Maldanidae) from the Loki's Castle vent field : an important structure builder in an Arctic vent system. *Polar Biology*, 35(2), 161-170.

Lacombe, D., & Liguori, V. R. (1969). COMPARATIVE HISTOLOGICAL STUDIES OF THE CEMENT APPARATUS OF *LEPAS ANATIFERA* AND *BALANUS TINTINNABULUM*. *The Biological Bulletin*, 137(1), 170-180.

Lacombe, D. (1970). A COMPARATIVE STUDY OF THE CEMENT GLANDS IN SOME BALANID BARNACLES (CIRRIPEDIA, BALANIDAE). *The Biological Bulletin*, 139(1), 164-179.

Laumer, C. E., Bekkouche, N., Kerbl, A., Goetz, F., Neves, R. C., Sørensen, M. V., Kristensen, R. M., Hejnol, A., Dunn, C. W., Giribet, G., & Worsaae, K. (2015). Spiralian phylogeny informs the evolution of microscopic lineages. *Current Biology*, 25(15), 2000-2006.

Le Cam, J.-B., Fournier, J., Etienne, S., & Couden, J. (2011). The strength of biogenic sand reefs: Visco-elastic behaviour of cement secreted by the tube building polychaete *Sabellaria alveolata*, Linnaeus, 1767. *Estuarine, Coastal and Shelf Science*, 91(2), 333 339.

Lebesgue, N., Da Costa, G., Ribeiro, R. M., Ribeiro-Silva, C., Martins, G. G., Matranga, V., Scholten, A., Cordeiro, C., Heck, A. J. R., & Santos, R. (2016). Deciphering the molecular mechanisms underlying sea urchin reversible adhesion: A quantitative proteomics approach. *Journal of Proteomics*, 138, 61-71.

Lee, S. J., Y. H. Han, B. H. Nam, Y. O. Kim and P. Reeves. (2008). A novel expression system for recombinant marine mussel adhesive protein Mefp1 using a truncated OmpA signal peptide. *Mol Cells*, 26(1): 34-40.

Lee, B. P., Messersmith, P. B., Israelachvili, J. N., & Waite, J. H. (2011). Mussel-Inspired Adhesives and Coatings. *Annual Review of Materials Research*, 41(1), 99-132.

Lee, D., Bae, H., Ahn, J., Kang, T., Seo, D.-G., & Hwang, D. S. (2020). Catechol-thiol-based dental adhesive inspired by underwater mussel adhesion. *Acta Biomaterialia*, 103, 92 101.

Lengerer, B., Pjeta, R., Wunderer, J., Rodrigues, M., Arbore, R., Schärer, L., Berezikov, E., Hess, M. W., Pfaller, K., Egger, B., Obwegeser, S., Salvenmoser, W., & Ladurner, P. (2014). Biological adhesion of the flatworm *Macrostomum lignano* relies on a duo-

gland system and is mediated by a cell type-specific intermediate filament protein. *Frontiers in Zoology*, 11(1), 12.

Lengerer, B., Bonneel, M., Lefevre, M., Hennebert, E., Leclère, P., Gosselin, E., Ladurner, P., & Flammang, P. (2018a). The structural and chemical basis of temporary adhesion in the sea star *Asterina gibbosa*. *Beilstein Journal of Nanotechnology*, 9, 2071-2086.

Lengerer, B., & Ladurner, P. (2018). Properties of temporary adhesion systems of marine and freshwater organisms. *Journal of Experimental Biology*, 221(16), jeb182717.

Lengerer, B., Wunderer, J., Pjeta, R., Carta, G., Kao, D., Aboobaker, A., Beisel, C., Berezikov, E., Salvenmoser, W., & Ladurner, P. (2018b). Organ specific gene expression in the regenerating tail of *Macrostomum lignano*. *Developmental Biology*, 433(2), 448-460.

Lengerer, B., Algrain, M., Lefevre, M., Delroisse, J., Hennebert, E., & Flammang, P. (2019). Interspecies comparison of sea star adhesive proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1784), 20190195.

Lerch, K., & Ettlinger, L. (1972). Purification and Characterization of a Tyrosinase from Streptomyces glaucescens. *European Journal of Biochemistry*, 31(3), 427-437.

Letunic, I., Khedkar, S., & Bork, P. (2021). SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Research*, 49(D1), D458-D460.

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293 W296.

Lewit-Bentley, A., & Réty, S. (2000). EF-hand calcium-binding proteins. *Current Opinion in Structural Biology*, 10(6), 637-643.

Li, S., Huang, X., Chen, Y., Li, X., & Zhan, A. (2019). Identification and characterization of proteins involved in stolon adhesion in the highly invasive fouling ascidian Ciona robusta. *Biochemical and Biophysical Research Communications*, *510*(1), 91-96.

Li, X., Li, S., Huang, X., Chen, Y., Cheng, J., & Zhan, A. (2021). Protein-mediated bioadhesion in marine organisms: A review. *Marine Environmental Research*, 170, 105409.

Li, X., Li, S., Cheng, J., Zhang, Y., & Zhan, A. (2024). Deciphering protein-mediated underwater adhesion in an invasive biofouling ascidian: Discovery, validation, and functional mechanism of an interfacial protein. *Acta Biomaterialia*, *181*, 146-160.

Liang, C., Ye, Z., Xue, B., Zeng, L., Wu, W., Zhong, C., Cao, Y., Hu, B., & Messersmith, P. B. (2018). Self-Assembled Nanofibers for Strong Underwater Adhesion: The Trick of Barnacles. *ACS Applied Materials & Interfaces*, 10(30), 25017 25025.

Liang, C., Strickland, J., Ye, Z., Wu, W., Hu, B., & Rittschof, D. (2019). Biochemistry of Barnacle Adhesion: An Updated Review. *Frontiers in Marine Science*, 6, 565.

Liang, Z., Yang, L., Zheng, J., Zuo, H., Weng, S., He, J., & Xu, X. (2019b). A low-density lipoprotein receptor (LDLR) class A domain-containing C-type lectin from *Litopenaeus*

vannamei plays opposite roles in antibacterial and antiviral responses. *Developmental & Comparative Immunology*, 92, 29-34.

Lima, R. N., & Porto, A. L. M. (2016). Recent Advances in Marine Enzymes for Biotechnological Processes. *In Advances in Food and Nutrition Research* (Vol. 78, p. 153 192). Elsevier.

Linnaeus, C. (1767). Systema naturae per regna tria naturae: secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Ed. 12. 1., Regnum Animale. 1 & 2. Holmiae [Stockholm], *Laurentii Salvii.* pp. 1-532 [1766] pp. 533-1327.

Lisco, S. N., Acquafredda, P., Gallicchio, S., Sabato, L., Bonifazi, A., Cardone, F., Corriero, G., Gravina, M. F., Pierri, C., & Moretti, M. (2020). The sedimentary dynamics of *Sabellaria alveolata* bioconstructions (Ostia, Tyrrhenian Sea, central Italy). *Journal of Palaeogeography*, 9(1), 2.

Liu, C., Li, S., Huang, J., Liu, Y., Jia, G., Xie, L., & Zhang, R. (2015). Extensible byssus of *Pinctada fucata:* Ca2+-stabilized nanocavities and a thrombospondin-1 protein. *Scientific Reports*, 5(1), 15018.

Liu, C., & Zhang, R. (2021). Identification of novel adhesive proteins in pearl oyster by proteomic and bioinformatic analysis. *Biofouling*, 37(3), 299-308.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research*, 48(D1), D265-D268.

Lüllmann-Rauch, R. (2008). Histologie. Bruxelles : De Boeck.

Lutz, T. M., Kimna, C., Casini, A., & Lieleg, O. (2022). Bio-based and bio-inspired adhesives from animals and plants for biomedical applications. *Materials Today Bio*, 13, 100203.

Lynn, A. K., Yannas, I. V., & Bonfield, W. (2004). Antigenicity and immunogenicity of collagen. *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, 71B (2), 343 354.

Machold, C., Schlegl, R., Buchinger, W., & Jungbauer, A. (2005). Matrix assisted refolding of proteins by ion exchange chromatography. *Journal of Biotechnology*, 117(1), 83-97.

Mackie, G. O., Mills, C. E., & Singla, C. L. (1988). Structure and function of the prehensile tentilla of Euplokamis (Ctenophora, Cydippida). *Zoomorphology*, 107(6), 319-337.

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., & Butcher, S. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Research*, 52(W1), W521-W525.

Mahmoud, K. (2007). Recombinant protein production: strategic technology and a vital research tool. *Research Journal of Cell and Molecular Biology*, 1(1), 9-22.

References

Manubolu, M., Goodla, L., Pathakoti, K., & Malmlöf, K. (2018). Enzymes as direct decontaminating agents—Mycotoxins. In *Enzymes in Human and Animal Nutrition* (p. 313-330). Elsevier.

Marcia, M., Langer, J. D., Parcej, D., Vogel, V., Peng, G., & Michel, H. (2010). Characterizing a monotopic membrane enzyme. Biochemical, enzymatic and crystallization studies on *Aquifex aeolicus* sulfide:quinone oxidoreductase. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1798(11), 2114-2123.

Märkel, K., Titschack, H. (1965). Das Festhaltevermögen von Seeigeln und die Reißfestigkeit ihrer Ambulacralfüßchen. *Naturwissenschaften* 52, 268.

Martínez Cuesta, S., Furnham, N., Rahman, S. A., Sillitoe, I., & Thornton, J. M. (2014). The evolution of enzyme function in the isomerases. *Current Opinion in Structural Biology*, *26*, 121-130.

Martínez Cuesta, S., Rahman, S. A., & Thornton, J. M. (2016). Exploring the chemistry and evolution of the isomerases. *Proceedings of the National Academy of Sciences*, *113*(7), 1796-1801.

Martinez Rodriguez, N. R., Das, S., Kaufman, Y., Israelachvili, J. N., & Waite, J. H. (2015). Interfacial pH during mussel adhesive plaque formation. *Biofouling*, 31(2), 221-227.

Marumo, K., & Waite, J. H. (1987). Prolyl 4-hydroxylase in the foot of the marine mussel *Mytilus edulis* L.: Purification and characterization. *Journal of Experimental Zoology*, 244(3), 365-374.

Mascolo, J. M., & Waite, J. H. (1986). Protein gradients in byssal threads of some marine bivalve molluscs. *Journal of Experimental Zoology*, *240*(1), 1-7.

May, S. W. (1979). Enzymatic epoxidation reactions. *Enzyme and Microbial Technology*, 1(1), 15-22.

May, S. W. (1999). Applications of oxidoreductases. *Current Opinion in Biotechnology*, 10(4), 370-375.

McDonald, A. G., Boyce, S., & Tipton, K. F. (2009). ExplorEnz: The primary source of the IUBMB enzyme list. *Nucleic Acids Research*, 37(Database), D593-D597.

McIntosh, Carmichael W. (1922) A monograph of the British marine annelids. *London,The Ray society,*1873/1900-1922/23.

McManus, J. F. A. (1948). Histological and Histochemical Uses of Periodic Acid. *Stain Technology*, 23(3), 99-108.

Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., & Ferrin, T. E. (2023). UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, 32(11), e4792.

Merz, R. A. (2015). Textures and traction: How tube-dwelling polychaetes get a leg up. *Invertebrate Biology*, 134(1), 61-77.

References

Mesko, M., Xiang, L., Bohle, S., Hwang, D. S., Zeng, H., & Harrington, M. J. (2021). Catechol-Vanadium Binding Enhances Cross-Linking and Mechanics of a Mussel Byssus Coating Protein. *Chemistry of Materials*, *33*(16), 6530-6540.

Messenger, J. B., Nixon, M., & Ryan, K. P. (1985). Magnesium Chloride as an anaesthetic for cephalopods. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology*, 82(1), 203-205.

Miao, Y., Zhang, L., Sun, Y., Jiao, W., Li, Y., Sun, J., Wang, Y., Wang, S., Bao, Z., & Liu, W. (2015). Integration of Transcriptomic and Proteomic Approaches Provides a Core Set of Genes for Understanding of Scallop Attachment. *Marine Biotechnology*, 17(5), 523-532.

Michaelis, L., and Menten, M. L. (1913). Die Kinetik der Invertinwirkung. *Biochem*. Z. 49, 333-369.

Miserez, A., Li, Y., Cagnon, J., Weaver, J. C., & Waite, J. H. (2012). Four-Stranded Coiled-Coil Elastic Protein in the Byssus of the Giant Clam, Tridacna maxima. *Biomacromolecules*, 13(2), 332-341.

Modaresifar, K., Azizian, S., & Hadjizadeh, A. (2016). Nano/Biomimetic Tissue Adhesives Development: From Research to Clinical Application. *Polymer Reviews*, 56(2), 329-361.

Mondarte, E. A. Q., Wang, J., & Yu, J. (2023). Adaptive Adhesions of Barnacle-Inspired Adhesive Peptides. *ACS Biomaterials Science & Engineering*, 9(10), 5679-5686.

Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S., Grigorenko, A. P., Dailey, C., Berezikov, E., Buckley, K. M., Ptitsyn, A., Reshetov, D., Mukherjee, K., Moroz, T. P., Bobkova, Y., Yu, F., Kapitonov, V. V., Jurka, J., Bobkov, Y. V., … Kohn, A. B. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature*, 510(7503), 109-114.

Müller, W. E. G., Zahn, R. K., & Schmid, K. (1972). The adhesive behaviour in Cuvierian tubules of Holothuria forskali. Biochemical and biophysical investigations. *Cytobiologie*, *5*, 335-351.

Nalbant, D., Youn, H., Nalbant, S. I., Sharma, S., Cobos, E., Beale, E. G., Du, Y., & Williams, S. C. (2005). FAM20: An evolutionarily conserved family of secreted proteins expressed in hematopoietic cells. *BMC Genomics*, 6(1), 11.

Naldrett, M. J. (1993). The importance of sulphur cross-links and hydrophobic interactions in the polymerization of barnacle cement. *Journal of the Marine Biological Association of the United Kingdom*, *73*(3), 689-702.

Newton, I., Marat, J.-P., & Blay, M. (2015). Optique (Nouvelle éd.). *Dunod*.

Nian, R., Kim, D. S., Tan, L., Kim, C., & Choe, W. (2009). Synergistic coordination of polyethylene glycol with ClpB/DnaKJE bichaperone for refolding of heat-denatured malate dehydrogenase. *Biotechnology Progress*, 25(4), 1078-1085.

References

Nicklisch, S. C. T., Spahn, J. E., Zhou, H., Gruian, C. M., & Waite, J. H. (2016). Redox Capacity of an Extracellular Matrix Protein Associated with Adhesion in *Mytilus californianus*. *Biochemistry*, 55(13), 2022-2030.

Nishi, E. (1993). On the Internal Structure of Calcified Tube Walls in Serpulidae and Spirorbidae (Annelida, Polychaeta). *Marine fouling*, 10(1), 17-20.

Noernberg, M. A., Fournier, J., Dubois, S., & Populus, J. (2010). Using airborne laser altimetry to estimate *Sabellaria alveolata* (Polychaeta: Sabellariidae) reefs volume in tidal flat environments. *Estuarine, Coastal and Shelf Science*, 90(2), 93-102.

Nunes, Y. L., De Menezes, F. L., De Sousa, I. G., Cavalcante, A. L. G., Cavalcante, F. T. T., Da Silva Moreira, K., De Oliveira, A. L. B., Mota, G. F., Da Silva Souza, J. E., De Aguiar Falcão, I. R., Rocha, T. G., Valério, R. B. R., Fechine, P. B. A., De Souza, M. C. M., & Dos Santos, J. C. S. (2021). Chemical and physical Chitosan modification for designing enzymatic industrial biocatalysts: How to choose the best strategy? *International Journal of Biological Macromolecules*, *181*, 1124-1170.

Oganesyan, N., Ankoudinova, I., Kim, S.-H., & Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expression and Purification*, 52(2), 280-285.

Ohkawa, K., Nishida, A., Yamamoto, H., & Waite, J. H. (2004). A Glycosylated Byssal Precursor Protein from the Green Mussel *Perna viridis* with Modified Dopa Side-chains. *Biofouling*, 20(2), 101-115.

Olivares, C., García-Borrón, J. C., & Solano, F. (2002). Identification of Active Site Residues Involved in Metal Cofactor Binding and Stereospecific Substrate Recognition in Mammalian Tyrosinase. Implications to the Catalytic Cycle. *Biochemistry*, 41(2), 679-686.

Osumi, N., Kakehashi, Y., Matsumoto, S., Nagaoka, K., Sakai, J., Miyashita, K., Kimura, M., & Asakawa, S. (2008). Identification of the gene for disaggregatase from *Methanosarcina mazei*. *Archaea*, 2(3), 185-191.

Paine, V. L. (1926). Adhesion of the tube feet in starfishes. *Journal of Experimental Zoology*, 45(1), 361-366.

Palmer, T., & Bonner, P. L. (2011). An Introduction to Enzymes. In Enzymes (p. 2-13). Elsevier.

Papov, V. V., Diamond, T. V., Biemann, K., & Waite, J. H. (1995). Hydroxyarginine-containing Polyphenolic Proteins in the Adhesive Plaques of the Marine Mussel *Mytilus edulis*. *Journal of Biological Chemistry*, 270(34), 20183-20192.

Parseghian, M. H., & Luhrs, K. A. (2006). Beyond the walls of the nucleus: The role of histones in cellular signaling and innate immunity. *Biochemistry and Cell Biology*, 84(4), 589-595.

References

Pasche, D., Horbelt, N., Marin, F., Motreuil, S., Macías-Sánchez, E., Falini, G., Hwang, D. S., Fratzl, P., & Harrington, M. J. (2018). A new twist on sea silk: The peculiar protein ultrastructure of fan shell and pearl oyster byssus. *Soft Matter*, 14(27), 5654-5664.

Paul, R. G., & Bailey, A. J. (2003). Chemical Stabilisation of Collagen as a Biomimetic. *The Scientific World JOURNAL*, 3, 138 155.

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., … Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, 51(D1), D418 D427.

Pearson, W.R. (2013). An introduction to sequence similarity ("homology") searching. *Current Protocols in Bioinformatics*, Chapter 3, Unit3.1.

Peng, Y. Y., Yoshizumi, A., Danon, S. J., Glattauer, V., Prokopenko, O., Mirochnitchenko, O., Yu, Z., Inouye, M., Werkmeister, J. A., Brodsky, B., & Ramshaw, J. A. M. (2010). A *Streptococcus pyogenes* derived collagen-like protein as a non-cytotoxic and non-immunogenic cross-linkable biomaterial. *Biomaterials*, 31(10), 2755 2761.

Peng Y. Y., Glattauer. V., Skewes. T. D., White, J. F., Nairn, K. M., McDevitt, A. N., Elvin, C. M., Werkmeister, J. A., Graham, L. D., & Ramshaw, J. A. (2011). Biomimetic materials as potential medical adhesives–composition and adhesive properties of the material coating the Cuvierian tubules expelled by *Holothuria dofleinii*. *Biomaterials: Physics and Chemistry*, 245.

Peng, Y. Y., Glattauer, V., Skewes, T. D., McDevitt, A., Elvin, C. M., Werkmeister, J. A., Graham, L. D., & Ramshaw, J. A. M. (2014). Identification of Proteins Associated with Adhesive Prints from *Holothuria dofleinii* Cuvierian Tubules. *Marine Biotechnology*, 16(6), 695-706.

Petrone, L. (2013). Molecular surface chemistry in marine bioadhesion. *Advances in Colloid and Interface Science*, 195-196, 1-18.

Pfister, D., De Mulder, K., Philipp, I., Kuales, G., Hrouda, M., Eichberger, P., Borgonie, G., Hartenstein, V., & Ladurner, P. (2007). The exceptional stem cell system of *Macrostomum lignano* : Screening for gene expression and studying cell proliferation by hydroxyurea treatment and irradiation. *Frontiers in Zoology*, 4(1), 9.

Pilakka Veedu, A., Nakashima, K., Shiga, H., Sato, T., Godigamuwa, K., Hiroyoshi, N., & Kawasaki, S. (2023). Functional modification of mussel adhesive protein to control solubility and adhesion property. *Journal of Bioscience and Bioengineering*, 136(2), 87 93.

Pjeta, R., Wunderer, J., Bertemes, P., Hofer, T., Salvenmoser, W., Lengerer, B., Coassin, S., Erhart, G., Beisel, C., Sobral, D., Kremser, L., Lindner, H., Curini-Galletti, M., Stelzer, C.-P., Hess, M. W., & Ladurner, P. (2019). Temporary adhesion of the proseriate

flatworm *Minona ileanae*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1784), 20190194.

Pjeta, R., Lindner, H., Kremser, L., Salvenmoser, W., Sobral, D., Ladurner, P., & Santos, R. (2020). Integrative Transcriptome and Proteome Analysis of the Tube Foot and Adhesive Secretions of the Sea Urchin *Paracentrotus lividus*. *International Journal of Molecular Sciences*, 21(3), 946.

Polivares, C., & Solano, F. (2009). New insights into the active site structure and catalytic mechanism of tyrosinase and its related proteins. *Pigment Cell & Melanoma Research*, 22(6), 750 760.

Power, A. M., Klepal, W., Zheden, V., Jonker, J., McEvilly, P., & von Byern, J. (2010). *Mechanisms of adhesion in adult barnacles* (pp. 153-168). Springer Vienna.

Pretzler, M., & Rompel, A. (2018). What causes the different functionality in type-III-copper enzymes? A state of the art perspective. *Inorganica Chimica Acta*, 481, 25 31.

Price, H. A. (1983). STRUCTURE AND FORMATION OF THE BYSSUS COMPLEX IN *MYTILUS* (MOLLUSCA, BIVALVIA*). Journal of Molluscan Studies*, 49(1), 9-17.

Priemel, T., Degtyar, E., Dean, M. N., & Harrington, M. J. (2017). Rapid self-assembly of complex biomolecular architectures during mussel byssus biofabrication. *Nature Communications*, 8(1), 14539.

Priemel, T., Palia, R., Babych, M., Thibodeaux, C. J., Bourgault, S., & Harrington, M. J. (2020). Compartmentalized processing of catechols during mussel byssus fabrication determines the destiny of DOPA. *Proceedings of the National Academy of Sciences*, 117(14), 7613-7621.

Proc, J. L., Kuzyk, M. A., Hardie, D. B., Yang, J., Smith, D. S., Jackson, A. M., Parker, C. E., & Borchers, C. H. (2010). A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin. *Journal of Proteome Research*, 9(10), 5422-5437.

Qin, X., & Waite, J. H. (1995). Exotic Collagen Gradients in the Byssus of the Mussel *Mytilus Edulis*. *Journal of Experimental Biology*, 198(3), 633-644.

Qin, X.-X., Coyne, K. J., & Waite, J. H. (1997). Tough Tendons. MUSSEL BYSSUS HAS COLLAGEN WITH SILK-LIKE DOMAINS. *Journal of Biological Chemistry*, 272(51), 32623-32627.

Qin, L., Wu, Y., Liu, Y., Chen, Y., & Zhang, P. (2014). Dual Effects of Alpha-Arbutin on Monophenolase and Diphenolase Activities of Mushroom Tyrosinase. *PLoS ONE*, 9(10), e109398.

Qin, C. L., Pan, Q. D., Qi, Q., Fan, M. H., Sun, J. J., Li, N. N., & Liao, Z. (2016). In-depth proteomic analysis of the byssus from marine mussel *Mytilus coruscus*. *Journal of proteomics*, *144*, 87-98.

References

Ramazi, S., Allahverdi, A. & Zahiri, J. (2020). Evaluation of post-translational modifications in histone proteins: A review on histone modification defects in developmental and neurological disorders. *Journal of Biosciences*. 45, 135.

Ramazi, S., & Zahiri, J. (2021). Post-translational modifications in proteins: Resources, tools and prediction methods. *Database*, 2021, baab012.

Ramsden, C. A., & Riley, P. A. (2014). Tyrosinase: The four oxidation states of the active site and their relevance to enzymatic activation, oxidation and inactivation. *Bioorganic & Medicinal Chemistry*, 22(8), 2388 2395.

Rasmussen, M., Jacobsson, M., & Björck, L. (2003). Genome-based Identification and Analysis of Collagen-related Structural Motifs in Bacterial and Viral Proteins. *Journal of Biological Chemistry*, 278(34), 32313-32316.

Read, G.; Fauchald, K. (Ed.) (2021). World Polychaeta Database. *Sabellaria alveolata* (Linnaeus, 1767). Accessed through: World Register of Marine Species at: http://www.marinespecies.org/aphia.php?p=taxdetails&id=130866 on 2021-02-24

Réaumur, R-A. (1711) Des différentes manières dont plusieurs espèces d'animaux de mer s'attachent au sable, aux pierres, et les uns aux autres. *Mémoires de mathématique et de physique de l'Académie royale des sciences, Académie royale des sciences*. ffads-00121334.

Reddy K., R. C., Lilie, H., Rudolph, R., & Lange, C. (2005). l -Arginine increases the solubility of unfolded species of hen egg white lysozyme. *Protein Science*, 14(4), 929-935.

Renner-Rao, M., Jehle, F., Priemel, T., Duthoo, E., Fratzl, P., Bertinetti, L., & Harrington, M. J. (2022). Mussels Fabricate Porous Glues via Multiphase Liquid–Liquid Phase Separation of Multiprotein Condensates. *ACS Nano*, 16(12), 20877-20890.

Robinson, P. K. (2015). Enzymes: Principles and biotechnological applications. *Essays in Biochemistry*, 59, 1-41.

Rodrigues, M., Lengerer, B., Ostermann, T., & Ladurner, P. (2014). Molecular biology approaches in bioadhesion research. *Beilstein Journal of Nanotechnology*, 5, 983-993.

Rodrigues, M., Leclère, P., Flammang, P., Hess, M. W., Salvenmoser, W., Hobmayer, B., & Ladurner, P. (2016). The cellular basis of bioadhesion of the freshwater polyp *Hydra*. *BMC Zoology*, 1(1), 3.

Rodríguez-López, J., Serna-Rodríguez, P., Tudela, J., Varón, R., & Garcia-Cánovas, F. (1991). A continuous spectrophotometric method for the determination of diphenolase activity of tyrosinase using 3,4-dihydroxymandelic acid. *Analytical Biochemistry*, 195(2), 369-374.

Romberger, D. J. (1997). Fibronectin. *The International Journal of Biochemistry & Cell Biology*, 29(7), 939-943.

References

Rosenhahn, A., & Sendra, G. H. (2012). Surface Sensing and Settlement Strategies of Marine Biofouling Organisms. *Biointerphases*, *7*(1), 63.

Ross, D. M., & Sutton, L. (1964a). Inhibition of the Swimming Response by Food and of Nematocyst Discharge During Swimming in the Sea Anemone *Stomphia Coccinea*. *Journal of Experimental Biology*, 41(4), 751-757.

Ross, D. M., & Sutton, L. (1964b). The Swimming Response of the Sea Anemone *Stomphia Coccinea* to Electrical Stimulation. *Journal of Experimental Biology*, 41(4), 735-749.

Rouse, G., & Pleijel, F. (2001). Polychaetes. Oxford university press.

Rudolph, R., & Fischer, S. (1990). *U.S. Patent No. 4,933,434. Washington, DC: U.S. Patent and Trademark Office*.

Rzepecki, L. M., Hansen, K. M., & Waite, J. H. (1992). Characterization of a cystine-rich polyphenolic protein family from the blue mussel *Mytilus edulis* L. *The Biological Bulletin*, 183(1), 123-137.

Saad, G. A., AlQurashi, N. A., & Hashimi, S. M. (2017). Comparative studies on the biological glue of some opportunistic adult marine macro-fouling after dislodgement and construction of temporary faunal conglomerations. *J Mar Biol Oceanogr 6*, *1*, 2.

Sagert, J., Sun, C., & waite, J. H. (2006). Chemical subtleties of mussel and polychaete holdfasts. In *Biological adhesives* (pp. 125-143). Berlin, Heidelberg: Springer Berlin Heidelberg.

Sagert, J., & Waite, J. H. (2009). Hyperunstable matrix proteins in the byssus of *Mytilus galloprovincialis. Journal of Experimental Biology*, 212(14), 2224-2236.

Samuel, D., Ganesh, G., Yang, P., Chang, M., Wang, S., Hwang, K., Yu, C., Jayaraman, G., Kumar, T. K. S., Trivedi, V. D., & Chang, D. (2000). Proline inhibits aggregation during protein refolding. *Protein Science*, 9(2), 344-352.

Sanchez, R., Serra, F., Tarraga, J., Medina, I., Carbonell, J., Pulido, L., De Maria, A., Capella-Gutierrez, S., Huerta-Cepas, J., Gabaldon, T., Dopazo, J., & Dopazo, H. (2011). Phylemon 2.0: A suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. *Nucleic Acids Research*, 39(suppl), W470-W474.

Sánchez-Ferrer, Á., Neptuno Rodríguez-López, J., García-Cánovas, F., & García-Carmona, F. (1995). Tyrosinase: A comprehensive review of its mechanism. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1247(1), 1-11.

Sanfilippo, R., Rosso, A., Mastandrea, A., Viola, A., Deias, C., & Guido, A. (2019). *Sabellaria alveolata* sandcastle worm from the Mediterranean Sea: New insights on tube architecture and biocement. *Journal of Morphology*, *jmor*.21069.

References

Santos, R., Haesaerts, D., Jangoux, M., & Flammang, P. (2005). Comparative histological and immunohistochemical study of sea star tube feet (Echinodermata, Asteroidea). *Journal of Morphology*, 263(3), 259-269.

Santos, R., & Flammang, P. (2007). Intra- and interspecific variation of attachment strength in sea urchins. *Marine Ecology Progress Series*, 332, 129-142.

Santos, R., Hennebert, E., Coelho, A. V., & Flammang, P. (2009a). The echinoderm tube foot and its role in temporary underwater adhesion. *Functional Surfaces in Biology: Adhesion Related Phenomena Volume 2*, 9-41.

Santos, R., Da Costa, G., Franco, C., Gomes-Alves, P., Flammang, P., & Coelho, A. V. (2009b). First Insights into the Biochemistry of Tube Foot Adhesive from the Sea Urchin *Paracentrotus lividus* (Echinoidea, Echinodermata). *Marine Biotechnology*, 11(6), 686-698.

Santos, R., Barreto, Â., Franco, C., & Coelho, A. V. (2013). Mapping sea urchins tube feet proteome—A unique hydraulic mechano-sensory adhesive organ. *Journal of Proteomics*, 79, 100-113.

Saroyan, J. R., Lindner, E., & Dooley, C. A. (1970). Repair and reattachment in the Balanidae as related to their cementing mechanism. *The Biological Bulletin*, *139*(2), 333-350.

Schmidt-Rhaesa, A. (Éd.). (2014). Gastrotricha and Gnathifera: DE GRUYTER.

Schmitt, C. N. Z., Winter, A., Bertinetti, L., Masic, A., Strauch, P., & Harrington, M. J. (2015a). Mechanical homeostasis of a DOPA-enriched biological coating from mussels in response to metal variation. *Journal of The Royal Society Interface*, 12(110), 20150466.

Schmitt, C. N. Z., Politi, Y., Reinecke, A., & Harrington, M. J. (2015b). Role of Sacrificial Protein–Metal Bond Exchange in Mussel Byssal Thread Self-Healing. *Biomacromolecules*, 16(9), 2852-2861.

Schröder, K., & Bosch, T. C. (2016). The origin of mucosal immunity: lessons from the holobiont Hydra. *MBio*, *7*(6), 10-1128.

Schultz, M. P. (2007). Effects of coating roughness and biofouling on ship resistance and powering. *Biofouling*, 23(5), 331 341.

Schultz, M. P., Bendick, J. A., Holm, E. R., & Hertel, W. M. (2011). Economic impact of biofouling on a naval surface ship. *Biofouling*, 27(1), 87 98.

Schultzhaus, J. N., Hervey, W. J., Taitt, C. R., So, C. R., Leary, D. H., Wahl, K. J., & Spillmann, C. M. (2021). Comparative analysis of stalked and acorn barnacle adhesive proteomes. *Open Biology*, *11*(8), 210142.

Shao, H., & Stewart, R. J. (2010). Biomimetic Underwater Adhesives with Environmentally Triggered Setting Mechanisms. *Advanced Materials*, 22(6), 729 733.

References

Shcherbakova, T. D., & Tzetlin, A. B. (2016). Fine structure of agglutinated tubes of polychaetes of the family Terebellidae (Annelida). *Doklady Biological Sciences*, 466(1), 16-20.

Shcherbakova, T. D., Tzetlin, A. B., Mardashova, M. V., & Sokolova, O. S. (2017). Fine structure of the tubes of Maldanidae (Annelida). *Journal of the Marine Biological Association of the United Kingdom*, 97(5), 1177-1187.

Shillito, B., Lübbering, B., Lechaire, J.-P., Childress, J. J., & Gaill, F. (1995). Chitin Localization in the Tube Secretion System of a Repressurized Deep-Sea Tube Worm. *Journal of Structural Biology*, 114(1), 67-75.

Shukla, E., D. Bendre, A., & M. Gaikwad, S. (2022). Hydrolases: The Most Diverse Class of Enzymes. *In S. Haider, A. Haider, & A. Catalá (Éds.), Biochemistry* (Vol. 29). IntechOpen.

Silverman, H. G., & Roberto, F. F. (2007). Understanding Marine Mussel Adhesion. *Marine Biotechnology*, 9(6), 661-681.

Simão, M., Moço, M., Marques, L., & Santos, R. (2020). Characterization of the glycans involved in sea urchin *Paracentrotus lividus* reversible adhesion. *Marine Biology*, 167(9), 125.

Smith, A. M. (1991). The Role of Suction in the Adhesion of Limpets. *Journal of Experimental Biology*, 161(1), 151-169.

Smith, A. M. (1992). Alternation between attachment mechanisms by limpets in the field. *Journal of Experimental Marine Biology and Ecology*, 160(2), 205-220.

Smith, A. M., Quick, T. J., & St. Peter, R. L. (1999). Differences in the Composition of Adhesive and Non-Adhesive Mucus From the Limpet *Lottia limatula*. *The Biological Bulletin*, 196(1), 34-44.

Smith, A. M. (2016). The biochemistry and mechanics of gastropod adhesive gels. *Biological adhesives*, 177-192.

So, C. R., Fears, K. P., Leary, D. H., Scancella, J. M., Wang, Z., Liu, J. L., Orihuela, B., Rittschof, D., Spillmann, C. M., & Wahl, K. J. (2016). Sequence basis of Barnacle Cement Nanostructure is Defined by Proteins with Silk Homology. *Scientific Reports*, 6(1), 36219.

So, C. R., Scancella, J. M., Fears, K. P., Essock-Burns, T., Haynes, S. E., Leary, D. H., Diana, Z., Wang, C., North, S., Oh, C. S., Wang, Z., Orihuela, B., Rittschof, D., Spillmann, C. M., & Wahl, K. J. (2017). Oxidase Activity of the Barnacle Adhesive Interface Involves Peroxide-Dependent Catechol Oxidase and Lysyl Oxidase Enzymes. *ACS Applied Materials & Interfaces*, 9(13), 11493-11505.

Solomon, E. I., Sundaram, U. M., & Machonkin, T. E. (1996). Multicopper oxidases and oxygenases. *Chemical Reviews*, 96(7), 2563 2606.

References

Stevens, M. J., Steren, R. E., Hlady, V., & Stewart, R. J. (2007). Multiscale Structure of the Underwater Adhesive of *Phragmatopoma Californica:* A Nanostructured Latex with a Steep Microporosity Gradient. *Langmuir*, 23(9), 5045-5049.

Stewart, R. J., Weaver, J. C., Morse, D. E., & Waite, J. H. (2004). The tube cement of *Phragmatopoma californica*: a solid foam. *Journal of Experimental Biology*, 207(26), 4727-4734.

Stewart, R. J., Ransom, T. C., & Hlady, V. (2011a). Natural underwater adhesives. *Journal of Polymer Science Part B: Polymer Physics*, *49*(11), 757-771.

Stewart, R. J., Wang, C. S., & Shao, H. (2011b). Complex coacervates as a foundation for synthetic underwater adhesives. *Advances in colloid and interface science*, *167*(1-2), 85-93.

Stewart, R. J., Wang, C. S., Song, I. T., & Jones, J. P. (2017). The role of coacervation and phase transitions in the sandcastle worm adhesive system. *Advances in Colloid and Interface Science*, 239, 88-96.

Storch, V. & Lehnert-Moritz, K. (1974). Zur Entwicklung der Kolloblasten von Pleurobrachia pileus (Ctenophora). *Marine Biology*, 28: 215–219.

Sugumaran, M., Soderhall, K., Iwanaga, S., Vastha, G. (1996). Role of insect cuticle in immunity. *In New Directions in Invertebrate Immunology; Eds.; SOS Publications: Fair Haven, NJ, USA*, pp. 355–374.

Suhre, M. H., Gertz, M., Steegborn, C., & Scheibel, T. (2014). Structural and functional features of a collagen-binding matrix protein from the mussel byssus. *Nature Communications*, 5(1), 3392.

Sun, C., Lucas, J. M., & Waite, J. H. (2002). Collagen-Binding Matrix Proteins from Elastomeric Extraorganismic Byssal Fibers. *Biomacromolecules*, 3(6), 1240-1248.

Sun, C., Fantner, G. E., Adams, J., Hansma, P. K., & Waite, J. H. (2007). The role of calcium and magnesium in the concrete tubes of the sandcastle worm. *Journal of Experimental Biology*, *210*(8), 1481-1488.

Sun, C. J., Srivastava, A., Reifert, J. R., & Waite, J. H. (2009). Halogenated DOPA in a Marine Adhesive Protein. *The Journal of Adhesion*, 85(2-3), 126-138.

Sun, J., Han, J., Wang, F., Liu, K., & Zhang, H. (2022). Bioengineered protein-based adhesives for biomedical applications. *Chemistry – A European Journal*, 28(1), e202102902.

Tagliabracci, V. S., Engel, J. L., Wen, J., Wiley, S. E., Worby, C. A., Kinch, L. N., Xiao, J., Grishin, N. V., & Dixon, J. E. (2012). Secreted Kinase Phosphorylates Extracellular Proteins That Regulate Biomineralization. *Science*, 336(6085), 1150-1153.

Tagliabracci, V. S., Pinna, L. A., & Dixon, J. E. (2013a). Secreted protein kinases. *Trends in Biochemical Sciences*, 38(3), 121 130.

References

Tagliabracci, V. S., Xiao, J., & Dixon, J. E. (2013b). Phosphorylation of substrates destined for secretion by the Fam20 kinases. *Biochemical Society Transactions*, 41(4), 1061-1065.

Tagliabracci, V. S., Wiley, S. E., Guo, X., Kinch, L. N., Durrant, E., Wen, J., Xiao, J., Cui, J., Nguyen, K. B., Engel, J. L., Coon, J. J., Grishin, N., Pinna, L. A., Pagliarini, D. J., & Dixon, J. E. (2015). A Single Kinase Generates the Majority of the Secreted Phosphoproteome. *Cell*, 161(7), 1619-1632.

Tajima, N., Takai, M., & Ishihara, K. (2011). Significance of Antibody Orientation Unraveled: Well-Oriented Antibodies Recorded High Binding Affinity. *Analytical Chemistry*, 83(6), 1969-1976.

Tamarin, A., & Keller, P. J. (1972). An ultrastructutal study of the byssal thread forming system in *Mytilus*. *Journal of Ultrastructure Research*, 40(3-4), 401-416.

Tamarin, A., Lewis, P., & Askey, J. (1976). The structure and formation of the byssus attachment plaque in *Mytilus. Journal of Morphology*, 149(2), 199-221.

Tanur, A. E., Gunari, N., Sullan, R. M. A., Kavanagh, C. J., & Walker, G. C. (2010). Insights into the composition, morphology, and formation of the calcareous shell of the serpulid *Hydroides dianthus. Journal of Structural Biology*, 169(2), 145-160.

Taylor, C. M., & Wang, W. (2007). Histidinoalanine: A crosslinking amino acid. *Tetrahedron*, 63(37), 9033-9047.

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., Von Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, 40(7), 1023-1025.

Thalmann, C., & Lötzbeyer, T. (2002). Enzymatic cross-linking of proteins with tyrosinase. *European Food Research and Technology*, 214(4), 276 281.

Theopold, U., Schmidt, O., Söderhäll, K., & Dushay, M. S. (2004). Coagulation in arthropods: Defence, wound closure and healing. *Trends in Immunology*, 25(6), 289 294.

Torrence, S. A., & Cloney, R. A. (1981). Rhythmic contractions of the ampullar epidermis during metamorphosis of the ascidian Molgula occidentalis. *Cell and Tissue Research*, *216*, 293-312.

Townsend, J. P., & Sweeney, A. M. (2019). Catecholic Compounds in Ctenophore Colloblast and Nerve Net Proteins Suggest a Structural Role for DOPA-Like Molecules in an Early-Diverging Animal Lineage. *The Biological Bulletin*, 236(1), 55-65.

Tregouboff, G. & Rose, M. (1957). *Manuel de planctonologiem mediterraneenne*. C. N. R. S., Paris, 587 pp.

References

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, 44(W1), W232 W235.

Trincone, A. (2010). Potential biocatalysts originating from sea environments. *Journal of Molecular Catalysis B: Enzymatic*, 66(3 4), 241 256.

Trincone, A. (2017). Enzymatic Processes in Marine Biotechnology. *Marine Drugs*, 15(4), 93.

Tunnicliffe, V., Desbruyères, D., Jollivet, D., & Laubier, L. (1993). Systematic and ecological characteristics of *Paralvinella sulfincola* Desbruyères and Laubier, a new polychaete (family Alvinellidae) from northeast Pacific hydrothermal vents. *Canadian Journal of Zoology*, 71(2), 286 297.

Tyler, S. (1976). Comparative ultrastructure of adhesive systems in the turbellaria. *Zoomorphologie*, 84(1), 1-76.

Tyler, S., & Rieger, G. E. (1980). Adhesive organs of the gastrotricha: I. Duo-gland organs. *Zoomorphologie*, 95(1), 1-15.

Ueki, T., Koike, K., Fukuba, I., & Yamaguchi, N. (2018). Structural and Mass Spectrometric Imaging Analyses of Adhered Tunic and Adhesive Projections of Solitary Ascidians. *Zoological Science*, *35*(6), 535-547.

Újvári, A., Aron, R., Eisenhaure, T., Cheng, E., Parag, H. A., Smicun, Y., Halaban, R., & Hebert, D. N. (2001). Translation Rate of Human Tyrosinase Determines ItsN-Linked Glycosylation Level. *Journal of Biological Chemistry*, 276(8), 5924 5931.

Ullrich, R., & Hofrichter, M. (2007). Enzymatic hydroxylation of aromatic compounds. *Cellular and Molecular Life Sciences*, 64(3), 271 293.

Van Gelder, C. W. G., Flurkey, W. H., & Wichers, H. J. (1997). Sequence and structural features of plant and fungal tyrosinases. *Phytochemistry*, 45(7), 1309 1323.

Van Holde, K. E., & Miller, K. I. (1995). Hemocyanins. *In Advances in Protein Chemistry* (Vol. 47, p. 1 81). Elsevier.

Varadi, M., Bertoni, D., Magana, P., Paramval, U., Pidruchna, I., Radhakrishnan, M., Tsenkov, M., Nair, S., Mirdita, M., Yeo, J., Kovalevskiy, O., Tunyasuvunakool, K., Laydon, A., Žídek, A., Tomlinson, H., Hariharan, D., Abrahamson, J., Green, T., Jumper, J., … Velankar, S. (2024). AlphaFold Protein Structure Database in 2024: Providing structure coverage for over 214 million protein sequences. *Nucleic Acids Research*, 52(D1), D368-D375.

Ventura, I., Harman, V., Beynon, R. J., & Santos, R. (2023). Glycoproteins Involved in Sea Urchin Temporary Adhesion. *Marine Drugs*, 21(3), 145.

Vermeer, L. M., Higgins, C. A., Roman, D. L., & Doorn, J. A. (2013). Real-time monitoring of tyrosine hydroxylase activity using a plate reader assay. *Analytical Biochemistry*, 432(1), 11-15.

References

Vicente-Manzanares, M., Ma, X., Adelstein, R. S., & Horwitz, A. R. (2009). Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nature Reviews Molecular Cell Biology*, 10(11), 778-790.

Vinn, O., Kirsimäe, K., & Ten Hove, H. A. (2009). Tube ultrastructure of *Pomatoceros americanus* (Polychaeta, Serpulidae): Implications for the tube formation of serpulids. *Estonian Journal of Earth Sciences*, 58(2), 148.

Vinn, O. (2011). The role of an internal organic tube lining in the biomineralization of serpulid tubes. *Carnets de Géologie* CG2011_L01.

Vinn, O. (2013). Occurrence, Formation and Function of Organic Sheets in the Mineral Tube Structures of Serpulidae (Polychaeta, Annelida). *PLoS ONE*, 8(10), e75330.

Von Byern, J., Mills, C. E., & Flammang, P. (2010). *Bonding tactics in ctenophores—morphology and function of the colloblast system* (pp. 29-40). Springer Vienna.

Vovelle J. (1965). Le tube de *Sabellaria alveolata* (L.) : Annélide polychète Hermellidae et son ciment. Etude écologique, expérimentale, histologique et histochimique. *Arch. Zool. Exp. Gén.*, 106 : 1-187.

Vreeland, V., Waite, J. H., & Epstein, L. (1998). MINIREVIEW—POLYPHENOLS AND OXIDASES IN SUBSTRATUM ADHESION BY MARINE ALGAE AND MUSSELS. *Journal of Phycology*, 34(1), 1-8.

Waite, J. H., & Tanzer, M. L. (1981). Polyphenolic Substance of *Mytilus edulis:* Novel Adhesive Containing L-Dopa and Hydroxyproline. *Science*, 212(4498), 1038-1040.

Waite, J. H. (1983a). ADHESION IN BYSSALLY ATTACHED BIVALVES. *Biological Reviews*, 58(2), 209-231.

Waite, J. H. (1983b). Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the mussel, *Mytilus edulis* L. *Journal of Biological Chemistry*, *258*(5), 2911-2915.

Waite, J. H. (1985). Catechol oxidase in the byssus of the common mussel, *Mytilus Edulis* L. *Journal of the Marine Biological Association of the United Kingdom*, 65(2), 359 371.

Waite, J. H. (1987). Nature's underwater adhesive specialist. *International Journal of Adhesion and Adhesives*, 7(1), 9 14.

Waite, J. H. (1990). The phylogeny and chemical diversity of quinone-tanned glues and varnishes. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 97(1), 19 29.

Waite, J. H., Jensen, R. A., & Morse, D. E. (1992). Cement precursor proteins of the reef-building polychaete *Phragmatopoma californica* (Fewkes). *Biochemistry*, 31(25), 5733 5738.

Waite, J. H., Qin, X.-X., & Coyne, K. J. (1998). The peculiar collagens of mussel byssus. *Matrix Biology*, 17(2), 93-106.

References

Waite, J. H., & Qin, X. (2001). Polyphosphoprotein from the Adhesive Pads of *Mytilus edulis*. *Biochemistry*, 40(9), 2887-2893.

Waite, J. H., Vaccaro, E., Sun, C., & Lucas, J. M. (2002). Elastomeric gradients: A hedge against stress concentration in marine holdfasts? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1418), 143-153.

Waite, J. H. (2017). Mussel adhesion – essential footwork. *Journal of Experimental Biology*, 220(4), 517-530.

Walker, G. (1970). The histology, histochemistry and ultrastructure of the cement apparatus of three adult sessile barnacles, *Elminius modestus*, *Balanus balanoides* and *Balanus hameri*. *Marine Biology*, 7(3), 239-248.

Walker, G. (1987). Marine organisms and their adhesion. *Synthetic adhesives and sealants*, 112-125.

Walmsley, S. J., Rudnick, P. A., Liang, Y., Dong, Q., Stein, S. E., & Nesvizhskii, A. I. (2013). Comprehensive Analysis of Protein Digestion Using Six Trypsins Reveals the Origin of Trypsin As a Significant Source of Variability in Proteomics. *Journal of Proteome Research*, 12(12), 5666-5680.

Wang, C. S., & Stewart, R. J. (2012). Localization of the bioadhesive precursors of the sandcastle worm, *Phragmatopoma californica* (Fewkes). *Journal of Experimental Biology*, *215*(2), 351-361.

Wang, C. S., & Stewart, R. J. (2013). Multipart Copolyelectrolyte Adhesive of the Sandcastle Worm, *Phragmatopoma californica* (Fewkes): Catechol Oxidase Catalyzed Curing through Peptidyl-DOPA. *Biomacromolecules*, 14(5), 1607-1617.

Wang, J., & Scheibel, T. (2018). Recombinant Production of Mussel Byssus Inspired Proteins. *Biotechnology Journal*, 13(12), 1800146.

Wang, J., Suhre, M. H., & Scheibel, T. (2019). A mussel polyphenol oxidase-like protein shows thiol-mediated antioxidant activity. *European Polymer Journal*, 113, 305 312.

Wang, L., Teng, L., Zhang, X., Liu, X., Lyu, Q., Yang, Y., & Liu, W. (2020). Discovery and characterization of tyrosinases from sea anemone pedal disc. *Journal of Adhesion Science and Technology*, 34(17), 1840-1852.

Wang, L., Zhang, X., Xu, P., Yan, J., Zhang, Y., Su, H., ... & Liu, W. (2022). Exploration of sea anemone-inspired high-performance biomaterials with enhanced antioxidant activity. *Bioactive Materials*, *10*, 504-514.

Waterhouse R.M., Seppey M., Simão F.A., Manni M., Ioannidis P., Klioutchnikov G., Kriventseva E.V. and Zdobnov E.M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, 35, 543-548.

References

Webb, E. C. (1992). Enzyme nomenclature 1992. Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes.

Whaite, A., Klein, A., Mitu, S., Wang, T., Elizur, A., & Cummins, S. (2022). The byssal-producing glands and proteins of the silverlip pearl oyster *Pinctada maxima* (Jameson, 1901). *Biofouling*, 38(2), 186-206.

Whittington, I. D., & Cribb, B. W. (2001). Adhesive secretions in the Platyhelminthes. *Advances in Parasitology*, *48*, 101-224.

Wichers, H. J., Recourt, K., Hendriks, M., Ebbelaar, C. E. M., Biancone, G., Hoeberichts, F. A., Mooibroek, H., & Soler-Rivas, C. (2003). Cloning, expression and characterisation of two tyrosinase cDNAs from *Agaricus bisporus*. *Applied Microbiology and Biotechnology*, 61(4), 336 341.

Williams, A. F., & Barclay, A. N. (1988). The Immunoglobulin Superfamily—Domains for Cell Surface Recognition. *Annual Review of Immunology*, 6(1), 381-405.

Williams GC, Van Syoc RJ. (2007) Methods of preservation and anesthetization of marine invertebrates. In: Carleton JT. (Ed.) The Light and Smith Manual: intertidal invertebrates from central California to Oregon. *University of California Press, Berkeley, CA*, 33–41.

Wunderer, J., Lengerer, B., Pjeta, R., Bertemes, P., Kremser, L., Lindner, H., Ederth, T., Hess, M. W., Stock, D., Salvenmoser, W., & Ladurner, P. (2019). A mechanism for temporary bioadhesion. *Proceedings of the National Academy of Sciences*, 116(10), 4297-4306.

Xu, Z., Liu, Z., Zhang, C., & Xu, D. (2022). Advance in barnacle cement with high underwater adhesion. *Journal of Applied Polymer Science*, 139(37).

Yamaguchi, H., & Miyazaki, M. (2014). Refolding Techniques for Recovering Biologically Active Recombinant Proteins from Inclusion Bodies. *Biomolecules*, 4(1), 235-251.

Yao, L., Wang, X., Xue, R., Xu, H., Wang, R., Zhang, L., & Li, S. (2022). Comparative analysis of mussel foot protein 3B co-expressed with tyrosinases provides a potential adhesive biomaterial. *International Journal of Biological Macromolecules*, 195, 229-236.

Yentsch, C. S., & Pierce, D. C. (1955). « Swimming » Anemone from Puget Sound. *Science*, 122(3182), 1231-1233.

Yoo, H. Y., Iordachescu, M., Huang, J., Hennebert, E., Kim, S., Rho, S., Foo, M., Flammang, P., Zeng, H., Hwang, D., Waite, J. H., & Hwang, D. S. (2016). Sugary interfaces mitigate contact damage where stiff meets soft. *Nature Communications*, *7*(1), 11923.

Young, G. A., Yule, A. B., & Walker, G. (1988). Adhesion in the sea anemones *Actinia equina* L. and *Metridium senile* (L.). *Biofouling*, 1(2), 137-146.

Younus, H. (2019). Oxidoreductases: Overview and Practical Applications. In Q. Husain & M. F. Ullah (Éds.), *Biocatalysis* (p. 39-55). Springer International Publishing.

References

Yu, M., Hwang, J., & Deming, T. J. (1999). Role of l -3,4-Dihydroxyphenylalanine in Mussel Adhesive Proteins. *Journal of the American Chemical Society*, 121(24), 5825-5826.

Yu, J., Wei, W., Danner, E., Ashley, R. K., Israelachvili, J. N., & Waite, J. H. (2011). Mussel protein adhesion depends on interprotein thiol-mediated redox modulation. *Nature Chemical Biology*, 7(9), 588 590.

Yu, Z., An, B., Ramshaw, J. A. M., & Brodsky, B. (2014). Bacterial collagen-like proteins that form triple-helical structures. *Journal of Structural Biology*, 186(3), 451-461.

Zeng, F., Wunderer, J., Salvenmoser, W., Ederth, T., & Rothbächer, U. (2019). Identifying adhesive components in a model tunicate. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1784), 20190197.

Zeugolis, D. I., Paul, G. R., & Attenburrow, G. (2009). Cross-linking of extruded collagen fibers—A biomimetic three-dimensional scaffold for tissue engineering applications. *Journal of Biomedical Materials Research Part A*, 89A (4), 895 908.

Zhang, X., Huang, H., He, Y., Ruan, Z., You, X., Li, W., Wen, B., Lu, Z., Liu, B., Deng, X., & Shi, Q. (2019). High-throughput identification of heavy metal binding proteins from the byssus of chinese green mussel (*Perna viridis*) by combination of transcriptome and proteome sequencing. *PLOS ONE*, 14(5), e0216605.

Zhao, H., Sun, C., Stewart, R. J., & Waite, J. H. (2005). Cement proteins of the tube-building polychaete *Phragmatopoma californica*. *Journal of Biological Chemistry*, *280*(52), 42938-42944.

Zhao, H., & Waite, J. H. (2006). Linking Adhesive and Structural Proteins in the Attachment Plaque of *Mytilus californianus*. *Journal of Biological Chemistry*, 281(36), 26150-26158.

Zhao, H., Robertson, N. B., Jewhurst, S. A., & Waite, J. H. (2006). Probing the Adhesive Footprints of *Mytilus californianus* Byssus. *Journal of Biological Chemistry*, 281(16), 11090-11096.

Zhao, H., Sagert, J., Hwang, D. S., & Waite, J. H. (2009). Glycosylated Hydroxytryptophan in a Mussel Adhesive Protein from *Perna viridis. Journal of Biological Chemistry*, 284(35), 23344-23352.

Zhao, Q., Lee, D. W., Ahn, B. K., Seo, S., Kaufman, Y., Israelachvili, J. N., & Waite, J. H. (2016). Underwater contact adhesion and microarchitecture in polyelectrolyte complexes actuated by solvent exchange. *Nature Materials*, 15(4), 407 412.

Zhao, Y., Wu, Y., Wang, L., Zhang, M., Chen, X., Liu, M., ... & Wang, Z. (2017). Bio-inspired reversible underwater adhesive. *Nature communications*, *8*(1), 2218.

Zhong, C., Gurry, T., Cheng, Downey, J., Deng, Z., Stultz, C.M., & Lu, T. K. (2014). Strong underwater adhesives made by self-assembling multi-protein nanofibres. *Nature Nanotech*, 9, 858–866 (2014).

Glossary

Glossary

Glossary

Acidic isoelectric point
The pH at which a molecule, typically a protein, has no net charge due to the balance between positive and negative charges. Proteins with an acidic isoelectric point are more likely to interact with basic surfaces.

Active site
The active site of an enzyme is the specific region where substrate molecules bind and undergo a chemical reaction.

Adhesive
A substance or protein capable of forming bonds with a surface or another material. In marine organisms, these adhesives help them to attach to substrates like rocks, shells, or other organisms.

Aldehyde
An organic compound containing a functional group with a carbonyl group (C=O) attached to a hydrogen atom.

Aldolase
An enzyme that catalyzes the cleavage of sugar molecules into smaller fragments. Aldolase plays a role in metabolic processes, and its activity can be part of the biochemical pathways in marine organisms.

Antifouling
The process of preventing the attachment of unwanted organisms (such as barnacles, algae, etc.) to submerged surfaces, often involving the production of substances that inhibit adhesion.

Baseplate
In marine organisms like barnacles, a baseplate is the part that anchors the organism to the substrate. It often serves as the initial point of contact and adhesion in the settling process.

BIC scores
Bayesian Information Criterion scores, used in statistical models to evaluate the fit of a model, often in bioinformatics and sequence alignment.

BLAST
Basic Local Alignment Search Tool, a bioinformatics algorithm used to compare a query sequence against a database to find regions of local similarity, essential in sequence alignment for identifying homologous genes or proteins.

BLOSUM62
A substitution matrix used in protein sequence alignment to score substitutions between amino acids based on observed evolutionary rates.

Bootstrap
A statistical method used to estimate the reliability of a dataset by resampling the data with replacement. In bioinformatics, it is used to evaluate the robustness of phylogenetic trees.

BUSCO
Benchmarking Universal Single-Copy Orthologs, a tool used in bioinformatics to assess the completeness of genome assemblies by evaluating the presence of core genes.

BPTI/Kunitz domain
The BPTI/Kunitz domain is a protein domain known for its role in inhibiting serine proteases, characterized by a compact, disulfide-rich structure that provides high stability and specificity.

Catechol
Catechol is an organic compound consisting of a benzene ring with two hydroxyl groups in adjacent positions, known for its role as a precursor in the biosynthesis of various biomolecules and as a key component in adhesive and redox processes.

Catecholase
An enzyme that catalyzes the oxidation of catechols to quinones, which can be involved in the hardening or cross-linking of adhesive molecules.

Chitobiose
Chitobiose is a disaccharide composed of two β-(1→4)-linked N-acetylglucosamine (GlcNAc) units, serving as a key building block in the structure of chitin and a substrate for chitin-degrading enzymes.

Chordin-like domain
A protein domain associated with proteins that regulate growth factors and may be involved in developmental processes.

CLANS
A tool for clustering large protein datasets based on sequence similarity, often used in bioinformatics to analyze large datasets of protein sequences, such as those involved in adhesion.

Coacervation
Coacervation is a phase separation process in which a homogeneous solution separates into two liquid phases: a dense, polymer-rich phase (the coacervate) and a dilute, polymer-depleted phase, often driven by electrostatic interactions or changes in environmental conditions.

Cohesive
Describing a substance or protein that exhibits the ability to stick to itself.

Coenzyme
A coenzyme is a non-protein organic molecule that binds to an enzyme and assists in catalyzing a chemical reaction, often by transferring electrons, atoms, or functional groups between molecules.

Co-expression
In recombinant protein technology, co-expression refers to the simultaneous expression of multiple genes or proteins in a host organism, often to produce a multi-subunit protein complex or to enhance the production and functionality of a recombinant protein.

Concomitant production
Concomitant production refers to the simultaneous production of two or more substances or products in a biological system, often occurring together in response to a single stimulus or process.

Cresolase
Cresolase is an enzyme that catalyzes the methylation of phenolic compounds, specifically converting phenol into cresol, a process important in the degradation of aromatic compounds in various biological systems.

Cryotome
A device used for cutting frozen samples into thin sections for microscopy.

Disulfide bond
A covalent bond formed between the sulfur atoms of two cysteine residues in a protein, stabilizing its structure.

ECM (Extracellular Matrix)
ECM (Extracellular Matrix) refers to the complex network of proteins, polysaccharides, and other molecules that provide structural and biochemical support to surrounding cells, playing a crucial role in tissue and organ development, function, and repair.

Edman degradation
A method of sequencing proteins by removing one amino acid at a time from the N-terminus of a peptide, allowing for the determination of the amino acid sequence.

Exocytosis
A process by which cells transport molecules, including adhesive proteins, out of the cell via vesicles that fuse with the cell membrane.

FASTA
A format used for representing nucleotide or protein sequences, often used for sequence alignment and analysis in bioinformatics.

FPKM (Fragments Per Kilobase of exon per Million fragments mapped)
A method used in RNA sequencing to measure gene expression levels.

Globular proteins
Globular proteins are proteins that fold into compact, spherical shapes, typically functioning as enzymes, antibodies, or transport molecules, and are soluble in water due to their hydrophilic exterior.

Glycan chains
Chains of sugar molecules that are often attached to proteins or lipids. They can play a role in the adhesion properties of marine invertebrates.

Glycocalyx
A glycoprotein and glycolipid-rich layer that surrounds the membrane of cells.

Glycoconjugates
Molecules consisting of carbohydrates covalently linked to proteins or lipids.

Inclusion bodies
Dense aggregates of proteins that accumulate inside cells, often occurring when proteins are overexpressed.

Intertidal zone
The intertidal zone is the coastal area that is exposed to air at low tide and submerged under water at high tide, creating a dynamic environment with varying conditions for organisms that live there.

Kinetics
Kinetics refers to the study of the rates at which chemical reactions occur and the factors that influence these rates, including temperature, concentration, and the presence of catalysts or inhibitors.

Ketone
A chemical compound containing a carbonyl group (C=O) bonded to two carbon atoms.

Kielin domain
The Kielin domain is a protein domain found in certain extracellular matrix proteins, involved in the regulation of signaling pathways, particularly those related to bone morphogenetic proteins (BMPs), which play a role in development and tissue homeostasis.

Kringle-domain-containing protein
It is a type of protein characterized by the presence of one or more Kringle domains, which are compact, looped structures stabilized by disulfide bonds. These proteins often participate in binding to other molecules, such as in blood clotting or in regulating cell adhesion and migration.

Lectin
Lectins are a type of protein that bind specifically to carbohydrates, often on the surface of cells, playing key roles in cell recognition, signaling, and immune responses.

Metamere
A segment of an organism's body, often referring to the repeating segments of some invertebrates.

Michaelis-Menten
A model that describes the kinetics of enzyme-catalyzed reactions, useful for understanding the speed at which adhesive enzymes function.

MAFFT algorithm
A sequence alignment algorithm used to align multiple sequences, important for comparing protein sequences involved in adhesion.

Meiofauna
Microscopic animals that live in the spaces between sediment particles in marine environments.

Mesopsammon
Organisms living in the upper layer of the sandy substrate in marine environments.

Mucocytes
Cells that secrete mucus or mucins.

Nectin
Nectin is a type of adhesion protein involved in cell-cell junctions, playing a key role in tissue formation and immune responses by facilitating the adhesion between neighboring cells.

Nematocysts
Specialized cells in cnidarians (such as jellyfish and sea anemones) that contain venomous structures used for adhesion, defense, and prey capture.

Omics technologies
Technologies used to study the entire complement of molecules in biological samples, including genomics, proteomics, and metabolomics.

ORF (Open Reading Frame)
A sequence of DNA that has the potential to be translated into a protein.

Orthologue
Genes in different species that evolved from a common ancestral gene and typically retain the same function.

Peritoneocytes
Peritoneocytes are specialized cells that line the peritoneal cavity, the space within the abdomen. They play a role in immune defense, fluid homeostasis, and tissue repair.

Peroxinectin
A type of adhesion protein found in certain invertebrates, involved in immune responses and wound healing.

Peroxidase
An enzyme that breaks down hydrogen peroxide and other peroxides.

Potter-Elvehjem tissue
A type of tissue homogenization method used to break down tissues into smaller fragments, useful in protein extraction.

Proteoglycans
Glycoproteins that are heavily glycosylated and play important roles in cell adhesion and extracellular matrix interactions.

Reverse-transcription
The process of converting RNA into complementary DNA (cDNA).

Sclerotization
The process by which proteins or chitin become hardened, often involved in the stabilization of adhesive structures in invertebrates.

Sessile
Describes organisms that are fixed in one place and do not move, such as barnacles and mussels.

Spurr resin
A type of resin used in electron microscopy to embed biological samples for study.

Support cells
Cells that provide structural support and assist in various biological functions.

Transaldolase
Transaldolase is an enzyme in the pentose phosphate pathway that catalyzes the transfer of a

three-carbon unit from a ketose donor to an aldose acceptor, playing a key role in cellular metabolism and the balance of sugars for nucleotide and amino acid synthesis.

Transketolase
An enzyme involved in the pentose phosphate pathway that transfers a two-carbon unit from a ketose donor to an aldose acceptor, contributing to the interconversion of sugars for energy and biosynthesis.

UFBoot
A statistical method used in phylogenetic analysis to assess the reliability of tree branchings, often used in the study of evolutionary relationships between proteins.

Vector system
A system used to introduce foreign genes into cells, often used in the production of recombinant proteins.

Yield
The amount of product, such as proteins or molecules, produced in a biochemical or biological process.

Annexes

Annexes

This section includes abstracts of communications presented at international conferences, as well as other relevant scientific publications in their published format.

Annex 1: Poster presentation at the 4th International Conference on Biological and Biomimetic Adhesives (ICBBA 2021) (University of Aveiro, Portugal, 18-19 February 2021)

Annex 2: Poster presentation at the international conference of the Royal Belgian Zoological Society (Zoology 2022) (KU Leuven, Campus Kortrijk, Belgium, 22-23 September 2022)

Annex 3: Oral communication at the 47th Annual Meeting of the Adhesion Society (Savannah, Georgia, United States, 11-14 February 2024)

Annex 4: Oral communication at the 16th International Conference on the Science and Technology of Adhesion and Adhesives (Adhesion 2024) (Oxford University, Oxford, United Kingdom, 4-6 September 2024)

Annex 5: Renner-Rao, M., Jehle, F., Priemel, T., Duthoo, E., Fratzl, P., Bertinetti, L., & Harrington, M. J. (2022). Mussels Fabricate Porous Glues via Multiphase Liquid–Liquid Phase Separation of Multiprotein Condensates. *ACS Nano*, *16*(12), 20877-20890. https://doi.org/10.1021/acsnano.2c08410

Annex 6: Nonclercq, Y., Lienard, M., Lourtie, A., Duthoo, E., Vanespen, L., Eeckhaut, I., Flammang, P., & Delroisse, J. (2024). *Opsin-based photoreception in Crinoids*. https://doi.org/10.1101/2024.08.14.607903

Annex 7: Duthoo, E., Delroisse, J., Maldonado, B., Sinot, F., Mascolo, C., Wattiez, R., Lopez, P. J., Van De Weerdt, C., Harrington, M. J., & Flammang, P. (2024). Diversity and evolution of tyrosinase enzymes involved in the adhesive systems of mussels and tubeworms. *iScience*, 111443. https://doi.org/10.1016/j.isci.2024.111443

Annexes

ANNEX 1

A tube inside a tube: The inner organic layer of *Sabellaria alveolata*

Emilie Duthoo [1*], Jérome Delroisse [1], Ruddy Wattiez [2] and Patrick Flammang [1]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Mons, Belgium
[2] Proteomics and Microbiology Unit, Research Institute for Biosciences, University of Mons, Mons, Belgium
*Contact: emilie.duthoo@student.umons.ac.be

Sabellaria alveolata, the honeycomb worm, is a gregarious tubeworm of the family Sabellariidae. This worm builds its tube thanks to a building organ that selects particles from the environment and binds them together with a strong cement. Because of those constructions, the honeycomb worm has been extensively studied. However, some aspects of the tube construction remain little known. Notably, *S. alveolata* secretes a thin fibrous layer covering the inside surface of its tube and the information available on this structure is limited to a few histological and histochemical analyses on the organ that secretes it: the parathoracic shield. Several experiments This study is therefore to better characterize its molecular composition and biosynthesis.

SEM images showed that this organic sheath is a fibrous structure which appears as a superposition of layers of parallelly-organized fibres embedded in a matrix. A histochemical study performed on the parathoracic shield showed two main types of secretory cells: collagen-secreting and acidic mucopolysaccharide-secreting cells. The combination of proteomic and transcriptomic analyses led to the identification of 33 putative proteins within the organic sheath. Interestingly, the two main proteins in MS analysis do not have homology with other proteins from public databases. Moreover, *in situ* hybridization showed that the mRNAs encoding these two proteins are well expressed in the parathoracic shield but in distinct areas. Two other proteins identified by the analysis are a collagen and a tyrosinase.

ANNEX 2

MATURATION OF MARINE ADHESIVE PROTEINS: INVOLVEMENT OF TYROSINASES IN MUSSELS AND TUBEWORMS

Duthoo Emilie[1*], Jérôme Delroisse[1], Barbara Maldonado[2], Matthew J. Harrington[3] and Patrick Flammang[1]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Mons, Belgium
[2] Molecular Biology and Genetic Engineering, GIGA-R, University of Liège, Liège, Belgium
[3] Department of Chemistry, McGill University, Montreal, Quebec H3A 0B8, Canada

*Corresponding author. E-mail: emilie.duthoo@umons.ac.be

Many benthic invertebrate organisms have developed attachment strategies to cope with the dynamic ocean environment. Permanent adhesion through the secretion of an underwater glue is one of these strategies and it allows sessile organisms such as mussels, tubeworms, and barnacles to remain attached at the same place throughout their lifetime and not being dislodged by waves. The blue mussel (*Mytilus edulis*) and the honeycomb worm (*Sabellaria alveolata*) have long been studied for their adhesion mechanisms known for their superior strength and durability compared to man-made materials. Mussels attach to surfaces by producing a proteinaceous holdfast called byssus, which consists of many load-bearing threads each terminated by a plaque that mediates adhesion. Honeycomb worms are tube-dwelling and build their tube by collecting particles with their tentacles and then applying dots of cement on these particles to bind them together. Both organisms have developed an underwater adhesive system based on numerous adhesive proteins rich in post-translationally modified amino acids, notably DOPA (3,4-dihydroxyphenylalanine) which plays important interfacial adhesive as well as bulk cohesive roles. DOPA is produced by the post-translational hydroxylation of tyrosine residues via a tyrosinase enzyme. However, the variability and specificity of the tyrosinases involved in the modification of adhesive proteins in both species are poorly understood.

By performing a BLAST search in the transcriptomes of *S. alveolata* and *M. edulis* with tyrosinase sequences known to be involved in the maturation of adhesive proteins in related species, several tyrosinases candidates were retrieved. These candidates were then analyzed *in silico* and localized by *in situ* hybridization in mussel and tubeworm tissues to confirm their expression in the adhesive glands of the two studied species. These findings provide new insights into the specificity and diversity of tyrosinases in these organisms with relevance for understanding the materials performance of this biological adhesion role model.

POSTER

ANNEX 3

Author Name: Emilie DUTHOO

Job Title: PhD student in the Biology of Marine Organisms and Biomimetics unit at the University of Mons, Belgium

Company Affiliation: Biology of Marine Organisms and Biomimetics unit, University of Mons, Belgium

Personal Biography:

After graduating from high school, I pursued a bachelor's degree in Biological Science, followed by a master's degree in Biochemistry, Molecular, and Cell Biology at the University of Mons in Belgium. During my master's thesis, I had the opportunity to conduct a research internship in the Harrington lab at McGill University in Canada, where I became deeply fascinated by the adhesive system of marine invertebrates. Back in Belgium, this passion subsequently guided me to a doctoral thesis, focusing on the enzymes responsible for the maturation process of adhesive proteins in various marine invertebrate species in collaboration with the Harrington lab.

Contact Information: emilie.DUTHOO@umons.ac.be

Abstract Title: Unraveling the complexity of the tyrosinase(s) involved in the blue mussel adhesive system

Abstract description:

The blue mussel, Mytilus edulis, secretes numerous proteinaceous threads, collectively called byssus, each ending in a plaque that allows adhesion to the substrate. The byssus therefore enables the animal to withstand the hydrodynamic forces of the waves in the intertidal zone. Because it is one of the strongest adhesive secretions among marine invertebrates, mussel adhesion has been studied for several decades. It is well established that the blue mussel has a DOPA-based adhesive system. DOPA (3,4-dihydroxy-L-phenylalanine) is produced by the post-translational hydroxylation of tyrosine residues within the adhesive proteins by tyrosinase enzymes. However, contrary to the adhesive proteins, the enzymes responsible for their modification, the tyrosinases, have been poorly studied.

Tyrosinases are oxygen-transferring, copper metalloproteins that catalyze the o-hydroxylation of monophenols (e.g., tyrosine) into o-diphenols (e.g., DOPA) and further oxidize o-diphenols into o-quinones. Our research integrated transcriptomic analyses with mass spectrometry analyses to identify the tyrosinase enzymes potentially involved in DOPA synthesis in the adhesive proteins of the blue mussel. Using this approach, several tyrosinases have been identified. Their roles were subsequently confirmed through their localization via in situ hybridization. To gain deeper insights, one of these enzymes was recombinantly produced, and we evaluated its enzymatic activity by measuring DOPA production in some mussel adhesive proteins. For this purpose, a comparison was made with the well-established mushroom tyrosinase. Our investigation may significantly contribute to advancing the design of innovative DOPA-based adhesive materials.

ANNEX 4

Investigating the complexity of tyrosinase(s) in the adhesive system of blue mussels

Emilie Duthoo[1], Barbara Maldonado[1], Elise Hennebert[2], Ruddy Wattiez[3], Mathieu Rivard[4], Matthew James Harrington[4] & Patrick Flammang[1]

[1] Biology of Marine Organisms and Biomimetics Unit, Research Institute for Biosciences, University of Mons, Place du Parc 23, B-7000 Mons, Belgium
[2] Laboratory of Cell Biology, Research Institute for Biosciences, Research Institute for Health Sciences and Technology, University of Mons, Place du Parc 23, B-7000 Mons, Belgium
[3] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of
Mons, Place du Parc 23, B-7000 Mons, Belgium
[4] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec H3A 0B8, Canada

The blue mussel, *Mytilus edulis*, uses its foot to secrete numerous protein threads collectively known as byssus, each terminating in a plaque that allows adhesion to the substrate. This byssus enables the animal to resist the hydrodynamic forces of waves in the intertidal zone and is considered one of the most powerful adhesive secretions among marine invertebrates, which explains why mussel adhesion has been studied for several decades. It is well established that the blue mussel has a catechol-based adhesive system. Catechol groups are produced by the enzymatic post-translational hydroxylation of tyrosine residues into DOPA (3,4-dihydroxy-L-phenylalanine) within the byssal adhesive proteins. DOPA has crucial roles in both interfacial adhesive and bulk cohesive interactions within their adhesive secretions. However, the enzymes responsible for its biosynthesis, the tyrosinases, have been poorly studied.

Tyrosinases are oxygen-transferring, copper metalloproteins that catalyse the o-hydroxylation of monophenols (e.g., tyrosine) into o-diphenols (e.g., DOPA) and further oxidise o-diphenols into o-quinones. Our research integrates foot transcriptomics with mass spectrometry analyses of byssal threads to identify the tyrosinase enzyme(s) potentially involved in DOPA synthesis in the adhesive proteins of the blue mussel. Using this approach, several tyrosinases have been identified. Their role in byssus formation was subsequently confirmed through their localisation in the foot glands via *in situ* hybridization. In parallel, proteomics confirmed the presence of one of the tyrosinase candidates in the vesicles of the plaque gland. Therefore, to gain deeper insights into this candidate, this enzyme was recombinantly produced. We evaluated its enzymatic activity by measuring DOPA production in comparison with the well-established mushroom tyrosinase. Our investigations contribute to our understanding of the molecular basis of adhesion in marine organisms, particularly helping to advance innovative DOPA-based adhesive materials.

Annexes