Università di Genova

Doctoral School of Marine Sciences and Technologies

Curriculum: Marine Ecosystem Sciences – Marine Biotechnology

XXXVI cycle (A.A. 2020/2021)

Molecular characterization of the fibrillar collagens from *Chondrosia reniformis* Nardo, 1847 and their biotechnological potential

Author: Dr. Eleonora Tassara

Supervisors: Prof. Marina Pozzolini and Prof. Marco Giovine

Table of Contents

Premise – *background and goals* 5

Section A – General Introduction 7

1 Chapter I – COLLAGEN 8

1.1 The collagen molecule 8

1.1.1 Chemical structure 8

1.1.2 Biosynthesis 9

1.1.3 Collagen types and their functions 11

1.1.4 Structure of collagen genes 17

1.2 The collagens of Invertebrates 19

1.3 Biotechnological usage of collagen 28

1.3.1 Properties and requirements of collagen as a biomaterial 28

1.3.2 Applications of collagen 29

1.3.3 Sources of collagen for biotechnological purposes 34

1.3.4 Employments of marine collagen 36

1.4 Extraction, purification, and characterization of collagen 38

1.4.1 Main methods of extraction and purification 38

1.4.2 Analyses for in vitro characterization 41

2 Chapter II – THE MARINE DEMOSPONGE *Chondrosia reniformis* (Nardo, 1847) 46

2.1 Ecology 46

2.1.1 Habitat 46

2.1.2 Reproduction strategies 47

2.1.3 Microbiome 47

2.2 Morphology 47

2.2.1 The collagen of *Chondrosia reniformis* 48

2.3 A unique skill: selective engulfment and dissolution of quartz particles 53

2.4 Biotechnological appeal 54

2.5 Sustainable production of *Chondrosia reniformis* 55

Section B – Research Activity 57

3 Chapter III – MOLECULAR CHARACTERIZATION OF THE FIBRILLAR COLLAGENS IN *Chondrosia reniformis* (Nardo, 1847) 58

3.1 Materials and methods 58

3.1.1 Acquisition of Fibrillar Collagen Gene Sequences 58

3.1.2 Analysis of Fibrillar Collagen Gene Sequences 61

3.1.3 Gene expression profile of fibrillar collagen genes in *Chondrosia reniformis* 61

3.2 Results 62

3.2.1 Transcriptomic analysis 62

3.2.2 Gene sequences and their translation products 63
3.2.3 Amino acid composition of the pre-proproteins 73
3.2.4 Comparison with the C-terminal domain of other fibrillar collagens 75
3.2.5 Potential regulatory sequences in the genes' upstream regions 75
3.2.6 Gene expression analysis 75
3.3 Discussion 77
***4* Chapter IV - FACTORS INFLUENCING COLLAGEN PRODUCTION IN *Chondrosia reniformis* (Nardo, 1847)** 82
EFFECT OF PHOTOBIOMODULATION ON THE REGENERATION PROCESS IN *CHONDROSIA RENIFORMIS* 82
4.1 Materials and methods 83
4.1.1 Sponge harvesting 83
4.1.2 Experimental design 83
4.1.3 Histology and Immunofluorescence 84
4.1.4 Gene Expression Evaluation 85
4.1.5 Oxidative Stress Enzyme Assays 86
4.1.6 Symbiont Bacteria Quantification 86
4.2 Results 86
4.2.1 Sponge Morphology during Regeneration 86
4.2.2 Histology and Archaeocyte Identification by Immunostaining 87
4.2.3 Gene Expression Profile 88
4.2.4 Sample Temperature Analysis 92
4.2.5 Oxidative Stress Evaluation 92
4.2.6 Prokaryotic Symbiont Quantification 92
4.3 Discussion 93
5 Chapter V - NOVEL EMPLOYMENTS OF BIOMATERIALS OBTAINED FROM THE COLLAGEN OF *Chondrosia reniformis* (Nardo, 1847) 95
2D COLLAGEN MEMBRANES FROM MARINE DEMOSPONGE *CHONDROSIA RENIFORMIS* (NARDO, 1847) FOR SKIN-REGENERATIVE MEDICINE APPLICATIONS: AN *IN VITRO* EVALUATION. 95
5.1 Materials and methods 95
5.1.1 Sponge sampling 96
5.1.2 *Chondrosia reniformis* Fibrillar Collagen Extraction and Purification 96
5.1.3 Production of 2D Collagen Membranes 96
5.1.4 2D Collagen Membranes Permeability Tests 96
5.1.5 Bacteria Infiltration Tests 98
5.1.6 Bacteriostaticity Test 99
5.1.7 *In Vitro* Biodegradability Test 99
5.1.8 Effect of 2D Collagen Membranes on Fibroblasts 100

5.2 Results 101
5.2.1 *Chondrosia reniformis* Fibrillar Extract Extraction and Purification 101
5.2.2 2D Collagen Membrane Permeability Tests 101
5.2.3 Bacteria Infiltration Tests 102
5.2.4 Bacteriostaticity Test 103
5.2.5 In Vitro Biodegradability Test 103
5.2.6 Effect of the 2D Collagen Membranes on Fibroblasts 104
USE OF COLLAGENASE-RESISTANT MARINE COLLAGEN FROM *CHONDROSIA RENIFORMIS* TO PRODUCE 2D-FILMS CONTAINING COLLAGENASE IMMOBILIZED IN BIO-SILICA NANOPARTICLES 105
5.3 Materials and methods 106
5.3.1 Collagenase encapsulation and immobilization on collagen fibrils 106
5.3.2 Production of composite collagen-silica-collagenase 2D membranes 106
5.3.3 Characterization of composite collagen-silica-collagenase 2D membranes 106
5.4 Results 107
5.4.1 Collagenase Activity Assay on Composite Membranes 107
5.4.2 Environmental Scanning Electron Microscope (ESEM) Observation 108
5.4.3 Thermogravimetric analysis (TGA) 110
5.4.4 Fourier transform infra-red (FT-IR) analysis 111
5.4.5 Absorbance Capacity 111
5.5 Discussion 112
Conclusions 116
References 118

Premise – *background and goals*

This doctoral project focuses on the molecular characterization of fibrillar collagens obtainable from the marine demosponge *Chondrosia reniformis* and on the factors that regulates their biosynthesis in the animal, prospecting an enhancement of their biotechnological applications.

The term "collagen" encompasses a diverse family of extracellular matrix (ECM) proteins characterized by the common feature of possessing at least one triple helix with a coiled-coil structure. Collagen finds application in a wide array of biotechnological fields, spanning from tissue engineering to various industries, including nutraceuticals, pharmaceuticals, and cosmetics. Currently, the predominant source of this molecule are mammals such as cows and pigs, which poses drawbacks such as the risk of spreading zoonotic diseases (like bovine spongiform encephalopathy), the ethical concerns related to intensive farming practices, and the limitations for several consumers due to religious constraints. On the other hand, the use of recombinant collagen –which could potentially resolve the majority of these issues– is restrained due to high costs of production, low yields of extraction and obstacles in recreating the original structure. In this scenario, a valid answer to the ever-growing demand for collagen could be found in collagen from marine origin, as it that offers numerous advantages: first, it retains the positive attributes of mammalian collagen, such as biocompatibility, biodegradability, and the presence of beneficial biological activities, along with comparable thermal stability; additionally, it shares the same potential applications, while eliminating the risk of zoonotic transmission and adapting to a wider consumer base. To date, the primary sources of collagen from marine origin –as well as the only ones that can be economically exploited– include bycatch, processing waste from the fishing industry, and blooms of gelatinous plankton. Within the marine collagen types described by now, some of the most intriguing features are frequently observed in those derived from the Porifera *phylum*. Indeed, studies have shown that collagen extracted from marine sponges holds promising potential for various applications, including the development of scaffolds for cell regeneration, membranes, hydrogels, drug delivery systems, and supports for immobilizing enzymes. Among poriferans, *Chondrosia reniformis* stands out as one of the species richest in collagen, being this protein the main component of its body. In addition to being very abundant, *C. reniformis*' collagen also has unique properties, responding to mechanical *stimuli* (resembling the mutable collagen of holothurians) and reorganizing into long filaments during sponge reproduction by budding. For these reasons, this animal has attracted attention in the field of biotechnology, both for the applications of its collagen and for successful integration into aquaculture systems.

C. reniformis and its collagen have been the subject of many studies since the past century; however, a widespread employment of this biomaterial is still limited because scaling up its production at an industrial level for meeting market demands is challenging. One of the most significant gaps still existing in literature regards an exhaustive molecular characterization of the collagen of this sponge; moreover, is necessary to expand the current knowledge regarding efficient methods to artificially enhance the collagen biosynthesis in the animal. Hence, the goals of this research project can be summarized into three main points:

1. **Achievement of an as comprehensive as possible molecular characterization of the fibrillar collagen(s) of *Chondrosia reniformis***: the aim is to obtain the sequences of all the fibrillar collagens genes present within the genome of this sponge, to define which of those genes are expressed in the adult animal in average conditions and to determine which gene products are

maintained in the collagenous extracellular matrix extract that constitutes the biomaterial obtainable from the animal.

2. **Evaluation of factors stimulating collagen synthesis and wound healing in *Chondrosia reniformis***: here, the purpose is to subject the animal or its regenerated explants to various types of environmental *stimuli* in order to enhance not only the collagen biosynthesis, but also to improve the ability of the sponge to recover and regain a normal conformation after fragmentation.
3. **Investigation of new ways of applications of the biomaterial obtainable from *Chondrosia reniformis***: starting from what is already known, the research activity plans to explore innovative ways to employ the collagenous extract of *C. reniformis* in the biotechnological field.

The goals were set with the idea that an in-depth understanding of the structure and of the expression patterns of collagen genes, responsible for coding for a great part of the proteins which constitute the extracellular matrix extract derived from *C. reniformis*, serves as a valuable tool for achieving an optimal manipulation of this biomaterial within a biotechnological framework. This information is envisioned to not only enhance the handling of the biomaterial itself, but also to offer insights for potential biomimetic applications. The discovery of specific *stimuli* able to influence the collagen production, improving the regeneration ability and, therefore, promoting sponge growth, is crucial in order to develop efficient aquaculture implants: this would lead to the creation of systems that would integrate an optimized diet and abiotic conditions, tailored for this species, with new additional devices, aiming to establish a sustainable production chain for *C. reniformis*' collagen. Finally, despite having been explored yet for several biotechnological applications, the full potential of the biomaterial obtainable from C. reniformis has not been explored yet. the project aims to enrich the landscape of innovative biomaterials by building upon ongoing studies conducted at the molecular biology laboratory associated with DiSTAV (University of Genoa) particularly in the field of collagen-derived membranes. This will expand their potential applications as marine biomaterials and, at the same time, contribute to improving the production efficiency of sponge aquaculture.

Section A – General Introduction

Chapter I – COLLAGEN

1.1 The collagen molecule

The structural integrity of multicellular organisms, spanning from algae to vertebrates, is provided by a non-cellular, three-dimensional macromolecular network, generically called extracellular matrix (ECM), primarily constituted by glycoproteins, proteoglycans and glycosaminoglycans [Exposito et al., 2002; Theocharis et al., 2016]. This framework is very dynamic and plays a pivotal role in development processes in both animals and plants; however, its composition differs significantly between the two kingdoms. One particular component exclusively identifies the extracellular matrix of Metazoans, from poriferans to humans, and that is collagen, the most abundant protein by weight in animals [Exposito et al., 2002; Patino et al., 2002]. The word "collagen" (from the Greek "κολλα", *glue*, and "Υενος", *origin*) literally means "something that produces glue" [Van Der Rest and Garrone, 1991] and refers to heterogeneous family of glycoproteins which carry out their functions solely in the ECM. Many types of collagens exist, each one of them with a more or less specific localization within the body, even if in the same ECM is not rare to observe two or more different typologies together [Ricard-Blum and Ruggiero, 2005].

1.1.1 Chemical structure

Despite the proven existence of many different types of collagens (see 1.1.3), all of them are bound by a common feature: their hallmark is a triple helix motif composed by three parallel left-handed polypeptide strands, called α-chains, held together by hydrogen bonds and associated in a typical right-handed, coiled-coil structure, which length can change considerably among the various collagens [Patino et al., 2002; Gelse et al., 2003; Shoulders and Raines, 2009]. This structure is the basic molecular unit of collagen and goes by the name of tropocollagen. A molecule of collagen can be a homotrimer, if all the three α-chains are identical in their aminoacidic composition, or a heterotrimer, if at least one of them differs from the other two [Gelse et al., 2003]; in any case, the aminoacid sequence of the α-chains is always characterized by the repeated motif $(Gly\text{-}Xaa\text{-}Yaa)_n$ which characterizes the "collagenous" domains of all collagens and where Xaa and Yaa can be any aminoacid, even if more often they are proline (Pro), hydroxyproline (HyPro, which is the hydroxylated counterpart of proline and can be exclusively found in collagen), or less frequently lysine (Lys) and hydroxylysine (HyLys) [Patino et al., 2002;Ricard-Blum and Ruggiero, 2005]. The occurrence of glycine every three aminoacid residues promotes the formation of the triple helix, as it represents the only aminoacid small enough to fit at the centre of the triple helix itself. Glycine residues, in fact, form the central axis of the triple helix, around which the three α-chains are supercoiled, staggered by one residue from each other; the larger side chains of other aminoacids, instead, occupy the outer positions, allowing a snug packing along the central axis of the molecule [Gelse et al., 2003]. Generally, the stability of the triple helical conformation is further improved by intra- and intermolecular bonds, which formation depends especially on the presence of hydroxyproline and hydroxylysine; the distribution and the percentage of these chemical interactions can significantly change among different collagens. Besides the triple helix, the "non-collagenous" domains flanking the central helical part, called telopeptides, are also an important structural component; in fact, they participate in both the covalent cross-linking of collagen molecules and the connection to other molecular structures within the surrounding matrix [Gelse et al., 2003].

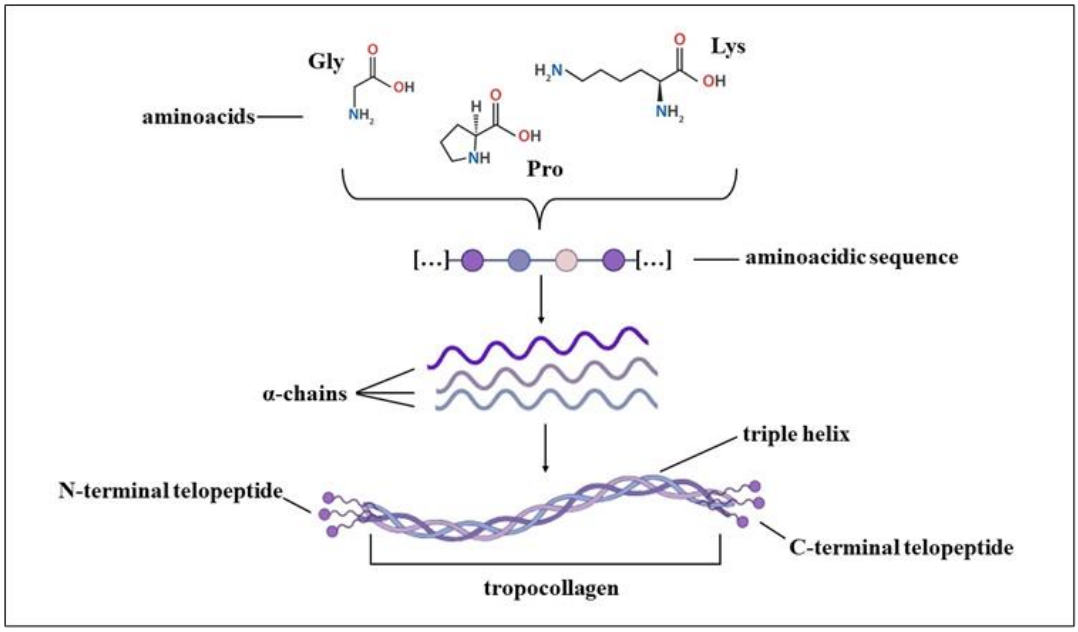


***Figure 1**: schematic representation of the chemical structure of collagen. Image realized with Biorender.com.*

1.1.2 Biosynthesis

The biosynthesis of collagen is carried out by different types of cells depending on the organism and on the type of tissue; for instance, in mammals, it is produced by fibroblasts in connective tissue and by osteoblasts in bone. Most of the knowledge about the biosynthesis of collagens regards the fibril-forming collagens (see 1.1.3); however, the fundamental processes involved in the formation and management of the triple helix are likely applicable to other types of collagens as well [Gelse et al., 2003]. The path from gene transcription to the formation of the final product is a complex process, in which the single steps can be divided, basing on where they occur, into intracellular and extracellular events [Patino et al., 2002]:

1. **Intracellular events**
   a. ***Transcription and translation***: the initial stage in the collagen synthesis process is the formation of collagen specific messenger-RNA. Even if the specific regulation of this event can vary between organisms and cell types, it often depends on the action of cytokines and growth factors like the transforming growth factor-β (TGF-β), the fibroblast growth factor (FGF), or insulin-like growth factors [Rossert and Crombrugghe, 2002; Gelse et al., 2003]. The transcription process yields a functional mRNA (typically around 3000 bases in length, though variations can occur), which then exits the nucleus and undergoes translation on membrane-bound polysomes attached to the rough endoplasmic reticulum (rER) [Patino et al., 2002].
   b. ***Post-translational modifications***: the collagen polypeptides synthesis in the rER is accompanied by essential co-translational events [Patino et al., 2002]. First, a signal peptidase removes the signal peptide, leading to the production of a so-called procollagen, which besides the triple helix also contains N- and C-terminal globular domains. Subsequently, the procollagen encounters a series of post-translational modifications. Several proline residues are hydroxylated by prolyl 4-hydroxylase and/or by prolyl 3-hydroxylase, which catalyse the hydroxylation at the level of the fourth or at the third carbon, respectively, of the pyrrolidine ring in the lateral chain of proline. 4-hydroxyproline is crucial for the formation of intramolecular hydrogen bonds, playing a key role in enhancing the thermal stability of the

triple helical domain; in fact, animals living at lower environmental temperatures usually have a lower extent of proline hydroxylation within their collagen [Cohen-Solal et al., 1986]; given its importance in maintaining molecular integrity, 4-hydroxyproline is particularly relevant in fibril-forming collagens, where averagely the 50% of the proline residues is hydroxylated on the fourth carbon of the ring [Gelse et al., 2003]. The function of 3-hydroxyproline, instead, is not completely clarified yet; however, it was hypothesized that its presence could help modulating the local stability of triple helices. This is supported by the fact that it is especially found in basement membranes collagens, engaged in more diverse and intricate interactions with each other and other biomolecules than those in fibrils; hence, 3-hydroxyproline could be needed in regions that necessitate finely tuned conformational stability [Jenkins et al., 2003]. In fibrillar collagens, 3-Hyp is present with only up to six occurrences per fibrillar collagen chain (depending on collagen type, tissue, species and age), in contrast with the more abundant and regularly distributed 4-hydroxyproline, but its role appears to be crucial [Parry et al., 2017]; in fact, 3-Hyp residues display their hydroxyl groups pointing out from the triple helix, making possible to establish hydrogen bonds with neighbouring molecules and possibly representing sites of molecular recognition. Alterations in prolyl 3-hydroxylation could induce changes in subsequent modifications in collagen, potentially modifying the phenotype [Parry et al., 2017]. It was in fact proven that this aminoacid is indispensable in embryogenesis, as mutations in the gene coding for prolyl 3-hydroxylase cause osteogenesis imperfecta in humans and mice, and prolyl 3-hydroxylase knock-out mice embryos encounter death during the early phases of development [Pokidysheva et al., 2014]. Lysine residues can also be hydroxylated by lysyl hydroxylase, in different amount based on the tissue and on the collagen type; the role of hydroxylysine residues is to form strong intermolecular cross-linking of collagen molecules in fibrils and to provide attachment sites for carbohydrates [Gelse et al., 2003]. All the three afore-mentioned enzymes require molecular oxygen, Fe^{2+} ions, 2-oxoglutarate and ascorbate as cofactors; the level of ascorbic acid, in fact, is strictly associated with the quantitative and structural integrity of collagen [Boo, 2022]. Furtherly, the enzymes hydroxylysyl-galactosyltransferase and galactosylhydroxylysyl-glucosyltransferase respectively transfer galactosyl- and glucosyl- residues on the hydroxyl groups of hydroxylysines [Gelse et al., 2003]. In conjunction with these enzymatic processes, the proper alignment and assembly of pro-α-chains into a triple helix take place by means of the C-propeptides [Patino et al., 2002; Gelse et al., 2003]; in particular, the transition from α-chains to triple helix is mediated by the alignment of the C-terminal domains of three α-chains, thus juxtaposing cysteine residues and promoting the creation of disulfide bridges that will link the single pro-α chains at the C-terminus. Then, the triple-helix formation progresses toward the N-terminal domains. The correct folding of the procollagen chains is also assisted by more enzymes, like peptidyl-prolyl *cis-trans*-isomerase (PPI) and collagen-specific chaperones, like HSP47 [Lang et al., 1987; Gelse et al., 2003].

c. ***Secretion***: at this point, procollagen molecules are transferred from the rough endoplasmic reticulum (rER) to the Golgi apparatus within the microsomal lumen; there, they are enclosed in secretory vesicles and transported to the cell surface, where they are released into the extracellular environment through exocytosis [Patino et al., 2002].

2. Extracellular events

d. ***Processing and modification***: after secretion, procollagen trimers undergo type-specific processing, involving the removal of N-propeptides and C-propeptides by specific proteases belonging to the Zn^{2+}-dependent metalloproteinase family: procollagen N-proteinase and procollagen C-proteinase, respectively. Additionally, lysyl oxidase, a Cu-dependent enzyme,

initiates cross-linking by catalysing the formation of aldehydes from lysine and hydroxylysine residues in telopeptides. Subsequent spontaneous reactions lead to intermolecular cross-links, essential for the physical and mechanical properties of collagen fibrils and stable network formation. Bifunctional cross-links undergo further intra and intermolecular reactions, resulting in diverse mature, trifunctional cross-links. This cross-link diversity contributes to significant variations between the different types of collagens [Patino et al., 2002; Gelse et al., 2003].

To maintain the normal tissue homeostasis, the ECM is subject to a slow but constant and dynamic turnover, in which collagen is degraded and then newly synthesised. The rate of turnover change significantly with respect to the tissue and the collagen type [Onursal et al., 2021]. Not only the turnover, but also the further management of collagens in the ECM is assigned to matrix metalloproteinases (MMPs), that are Zn^{2+}-dependent endopeptidase often acting specifically on different types of collagens [Ricard-Blum, 2011].

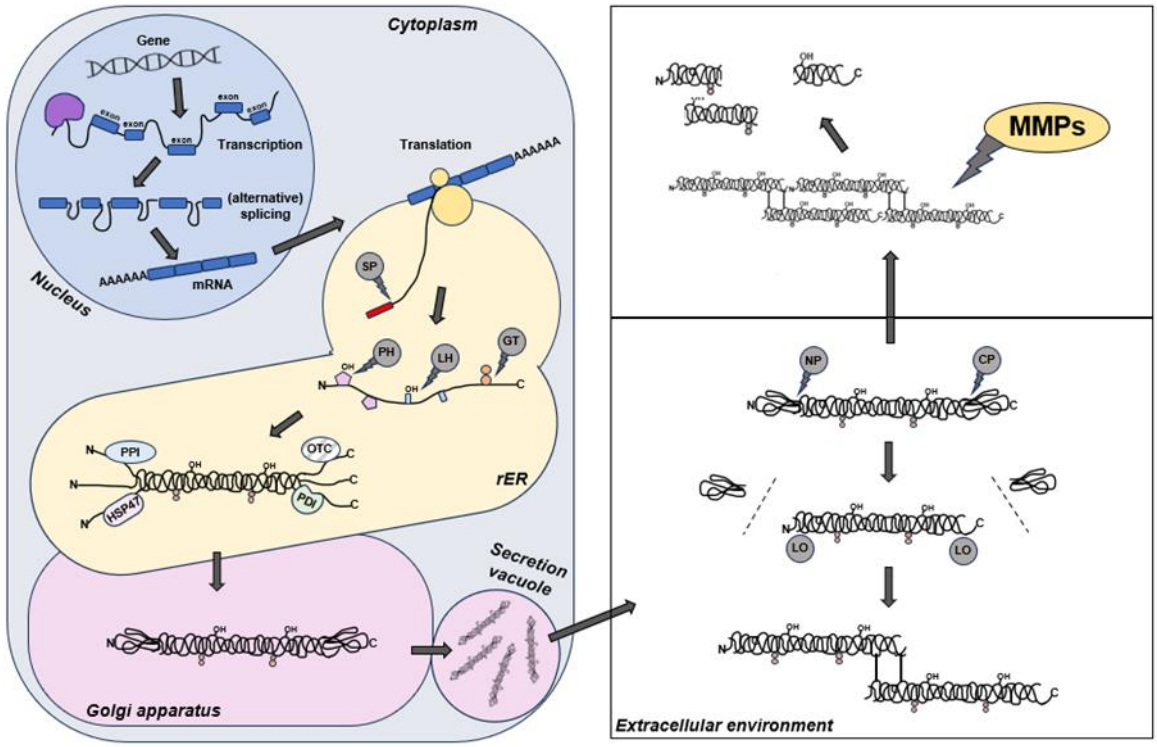


***Figure 2**: schematic representation of the collagen biosynthesis process. The involved enzymes are indicated with the following acronyms. SP: signal peptidase; PH: prolyl-hydroxylase; LH: lysyl-hydroxylase; GT: hydroxylysyl-galactosyltransferase/ galactosylhydroxylysyl - glucosyltransferase; PPI: peptidyl prolyl cis-trans isomerase; OTC: oligosaccharyl transferase complex; HSP47: heat shock protein 47; PDI: protein disulphide isomerase; NP: procollagen N-proteinase; CP: procollagen C-proteinase; LO: lysyl oxidase; MMPs: matrix metalloproteinases.*

1.1.3 Collagen types and their functions

Until about some decades ago, collagens were usually divided only in fibril-forming or network/basement membrane-forming proteins. Thanks to more recent advancements in molecular biology, it was possible to identify numerous molecules, now classified in every respect as collagens, that however deviate structurally and functionally from the traditional fibril and basement membrane-forming types [Fukai et al., 1994]. To date, 29 different collagens have been described, respectively designated with Roman numerals from I to XXIX accordingly the order in which they were discovered; based on their primary and/or supramolecular structures, most authors agree to classify

them in two main classes: the fibrillar (or fibril forming) collagens and the non-fibrillar (or non-fibril forming) collagens. The latter class can be furtherly divided into other sub-categories: basement membrane collagens, anchoring collagens, microfibrillar collagens, network-forming collagens, FACITs (Fibril Associated Collagens with Interrupted Triple helices), MACITs (Membrane Associated Collagens with Interrupted Triple helices, previously called also transmembrane collagens) and multiplexins [Gelse et al., 2003; Owczarzy et al., 2020] (see ***Figure 3***).

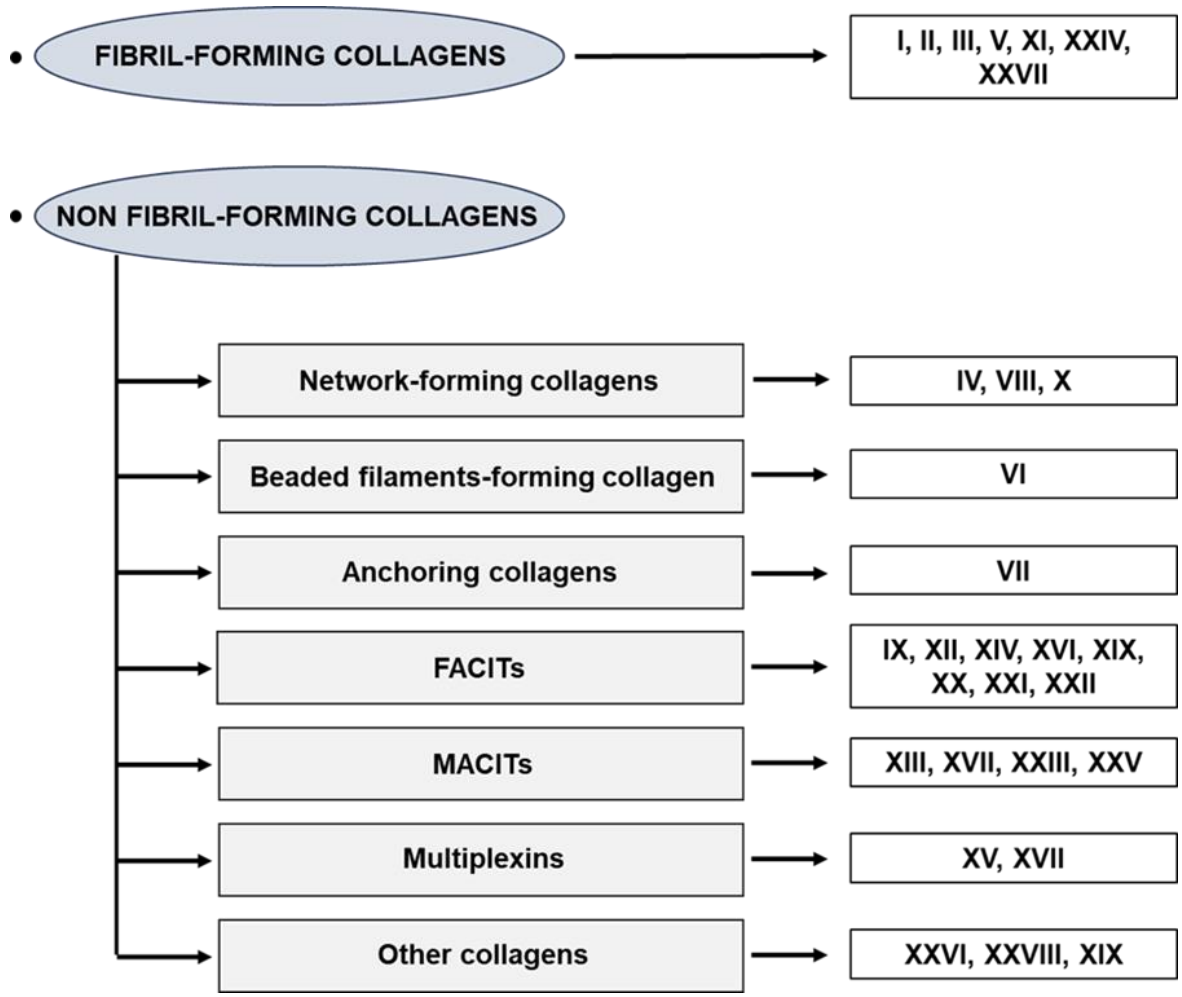


Figure 3*: the family of collagen proteins [Onursal et al., 2021]. Note that this classification may change over time and that not all authors agree, as some of the nonfibrillar collagens may fit in more than one subcategory, or not fit completely; moreover, subcategories may be made following different criteria.*

Such diversity in collagen categories reflects an as wide array of functions. Collagens are defined as multi-domain proteins; each domain can contribute to structural reinforcement and/or confer various biological activities to the molecule. In addition to the representative triple helix domain, commonly occurring domains in collagens include Fibronectin-III, Kunitz, Thrombospondin-1, and von Willebrand domains. To a large extent, collagens have a mechanical function and are essential in the maintenance of the structural integrity of tissues and organs; however, has become increasingly evident in recent years their involvement in highly specialized functions beyond ECM architecture. For instance, collagens mediate interactions with specific receptors, influencing adhesion, differentiation, growth, cellular reactivity, and cell survival through signaling pathways. Furthermore, collagens contribute to the entrapment, local storage, and controlled release of growth factors and

cytokines, playing crucial roles in organ development, wound healing, and tissue repair [Gelse et al., 2003; Ricard-Blum, 2011].

1.1.3.1 Fibrillar collagens

The first collagens to be ever discovered were the fibril-forming collagens, more commonly known as fibrillar collagens: for this reason, they are still considered today as the "classical" collagens. At least in Vertebrates, they constitute the principal products synthesized by cells within connective tissues (including fibroblasts, osteoblasts, and chondrocytes), where their role is primarily mechanical-structural, as they create and maintain three-dimensional frameworks for tissues and organs [Ricard-Blum and Ruggiero, 2005; Parry et al., 2017]. Fibril-forming collagens owe their name to the ability of their molecules to aggregate into highly organized fibrillar structures with a banded pattern at an ultrastructural level. In particular, mature tropocollagen molecules (see paragraph 1.1.2) assemble into bundles through intermolecular bonds with D-staggered interactions, thus forming collagen fibrils; in the D-staggered arrangement, the individual tropocollagen molecules are slightly displaced with respect to each other along the long axis of the fibril, creating a typical banded pattern observable with scanning electron microscopy. This kind of interaction is guided and regulated by cell-surface and extracellular proteins (for instance, other collagens, integrins, fibronectin, and proteoglycans) and is crucial for the stability and strength of collagen fibrils [Parry et al., 2017]. In turn, such fibrils combine into fibers, ready to carry out their structural functions in tissues and organs (see ***Figure 4***).

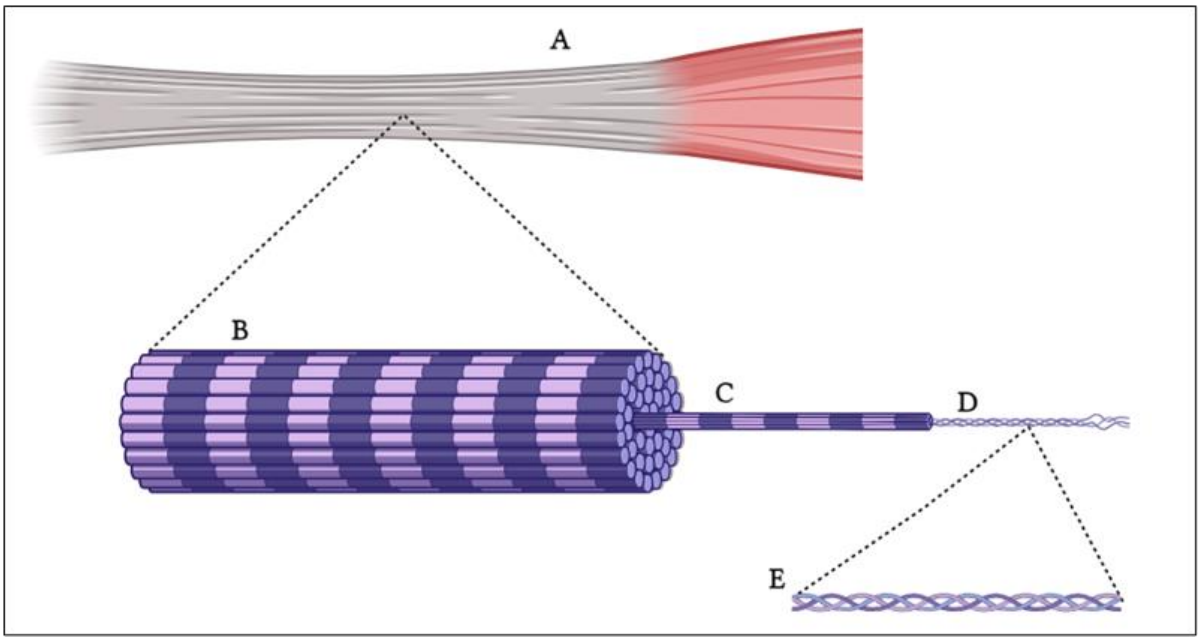


Figure 4*: hierarchical organization of fibrillar collagens. A: tendon; B: collagen fiber; C: collagen fibril; D: collagen molecule; E: triple helix. Image realized with Biorender.com.*

When fibrils contain a mixture of major and minor fibrillar collagens, they are called heterotypic; the structure and arrangement of the matrix network, as well as the biomechanical properties of the tissues, are dictated by the proportional composition and quantities of various collagen types that form such heterotypic fibrils [Ricard-Blum and Ruggiero, 2005; Parry et al., 2017].

This subfamily further comprehends seven members: types I, II, and II (major fibrillar collagens), types V and XI (minor fibrillar collagens), and types XXIV and XXVII (other fibrillar collagens,

more recently described). A fibrillar collagen has a relatively simple structure, and is considered as such if it contains a central uninterrupted triple helical domain (COL1 domain), which is averagely 1014–1020 amino acids long and represents the major part of the polypeptide, flanked by relatively short telopeptides; in particular, the C-terminal domain (NC1 domain) is trimeric as the central region, while the N-terminal domain (N-propeptide) can have various conformations in the different clades of fibrillar collagens (see below), but usually it is constituted by one or more short triple helical sequences (40-60 amino acids) and is therefore called minor helix domain or COL2 domain [Ricard-Blum and Ruggiero, 2005; Parry et al., 2017; Onursal et al., 2021]. From evolutionary studies [Exposito et al., 2008], three fibrillar collagen clades (named A, B, and C) have been defined:

- **Clade A** includes the collagen types I, II, and III; here, the N-propeptide, which often possesses a short von Willebrand factor type C domain (VWC), is completely removed by an endopeptidase cleavage that takes place between the minor (COL2) and the major (COL1) triple helices. Following the removal of the C-propeptide as well, mature clade A collagens are essentially constituted by extended COL1 domains flanked by short telopeptides, which provide sites for cross-linking during fibril assembling [Parry et al., 2017]. Each member of this clade, however, has its distinctive features. Type I collagen is the most abundant type in mammals, constituting up to 99% of collagen present in skin, tendons, bones, arteries, and cornea; It is composed of two identical α-chains, α-1(I), and one different chain, α-2(I). During their assembling, type I collagen fibrils result less compactly arranged in the gap regions than in the overlapped regions, because of lower levels of proline and hydroxyproline; in such regions, the fibril undergoes kinks. This conformation makes this collagen one of the main stress-carrying protein structures in mammals [Patino et al., 2002]. Type II collagen is less abundant than type I and is mainly found in cartilaginous tissues, being synthesized in chondrogenic stages of mesoderm development. It is a homotrimer, with three identical α-chains, α-1(II), that contain a high amount of hydroxylysine and display high levels of hydroxylysine glycosylation, as well [Patino et al., 2002]. Type III collagen is a homotrimer too and contains three identical α-1(III) chains; contrarily to type II, it has a low percentage of hydroxylysine residues, while is high in hydroxyproline. This collagen is correlated with the ability of a tissue to extend, in fact it is especially present, often in association with type I, in tissues and organs that need elastic properties, such as skin, lungs, heart, uterus, eye structures, and so on [Patino et al., 2002].
- **Clade B** comprehends type V and XI collagens, that form elongated, cable-like striated fibrils and have a well-defined structural role of mechanical support and dimensional stability. In this group, the C-propeptide is usually completely cleavaged, as in clade A, while the N-propeptide is retained partially. The retained part of the N-propeptide often contains a thrombospondin domain (TSPN), also known as laminin globular (LamG), and a minor triple helix, like what is observed in several nonfibrillar collagen sequences; this architecture works as a regulator for the fibril diameter. The minor and the major triple helices are divided by a short flexible region [Parry et al., 2017]. Type V collagen is the most heterogeneous typology among the fibrillar collagens [Ricard-Blum and Ruggiero, 2005], is mainly found in vascular tissues, and has a variable α-chain composition: the most common isoform contains two α-1(V) chains and one α-2(V) chain, but it has also been found composed by three different α-chains, α-1(V), α-2(V), and α-3(V), or as a homotrimer, with three identical α-1(V) chains. Compared to the other types of collagens, type V is more soluble and has significantly larger N-terminal domains. It has a low content of alanine, but a high percentage of hydroxylysine, which is only partially glycosylated [Patino et al., 2002]. Type XI collagen is a component of cartilaginous tissues, where is thought to be involved in the regulation of the diameter and of the growth of type II collagen. Its predominant isoform contains

three different α-chains, α-1(XI), α-2(XI), and α-3(XI). Type V and XI collagen can form tissue-specific heterotypic molecules, with mixed chains of the two different types [Parry et al., 2017].

- **Clade C** encompasses the less abundant type XXIV collagen and type XXVII collagen, which display several important structural differences if compared with clades A and B. First, they have a shorter COL1 domain, counting less than 1000 amino acids, which surprisingly contains two interruptions; moreover, their N-terminal propeptide is devoid of the minor triple helix and consists of a 'mutable' domain and a thrombospondin domain. These structural characteristics influence the diameter of the fibrils, which are much thinner than in the other fibrillar collagens. Regarding their collocation, type XXIV is predominately expressed during the formation of bone tissue, even if minor amounts of it can be found in soft tissues too, such as in brain, muscles, kidneys, spleen, liver, lung, and gonads; type XXVII collagen, despite being expressed in a variety of tissues during development stages, can be most prominently found in cartilaginous tissues [Plumb et al., 2011; Parry et al., 2017; Nielsen et al., 2024].

1.1.3.2 Non-fibrillar collagens

As the name says, non-fibril forming collagens include every type of collagen which may form any supramolecular structure other than fibrils [Parry and Squire, 2017]; actually, the greatest part of collagens discovered to date belongs to this group. For this reason, the author believes that a brief, but exhaustive description of these molecules is necessary, although the present work is mainly focused of fibrillar collagens. With respect to their structure and function, many sub-categories can be distinguished:

- **Network forming collagens** (**type IV, VIII and X collagens**): the prototypical network forming collagen, as well as the first to be included in this category, is type IV collagen [Onursal et al., 2021]. Together with laminins and heparan sulphate proteoglycans, type IV collagen is one of the major components of basement membranes, i. e. that thin layer of highly specialized ECM which is located basolaterally to all cell monolayers [Patino et al., 2002; LeBleu et al., 2007]. Basement membranes secure cells to their surrounding ECM and potentially function as barriers for macromolecules and storage sites of growth factors, meanwhile maintaining the structural integrity and facilitating the differentiation of tissues [Exposito and Lethias, 2013]. To date, six different α-chains, but only three isoforms, have been identified: $[a1(IV)]_2$ a2(IV), a3(IV) a4(IV) a5(IV), and $[a5(IV)]_2$ a6(IV) [LeBleu et al., 2007; Birk and Brückner, 2011]. Type IV collagen is composed by a) a N-terminal domain, called 7S, which consists of interrupted Gly-Xaa-Yaa repeats, b) a triple helical collagenous domain also containing several interruptions of the Gly-Xaa-Yaa sequence, and c) a C-terminal non collagenous, globular domain. The 7S domains connect to each other in an antiparallel way through interchain disulfide bonds and lysine-associated cross-links [Exposito and Lethias, 2013], while the C-terminal domains of different collagen molecules connect "head-to-head". The triple helical domains connect laterally. The resulting supramolecular organization is a chicken-wire-like network, where the mesh sizes are variable. The short, but numerous, interruptions of the type IV collagen domains give great flexibility to the whole structure [LeBleu et al., 2007; Birk and Brückner, 2011]. It has been demonstrated that analogous network systems can be formed also by type VIII and type X collagens, even if the molecular mechanisms of their assembling are still not known in such details [Onursal et al., 2021].

- **Anchoring collagen (type VII collagen)**: as type IV, type VII collagen is an integral and fundamental part of basement membranes, as it represents one of the major components of the so-called anchoring fibrils, essential in the dermo-epidermal junction of the skin [Patino et al., 2002;

Onursal et al., 2021]. This collagen is composed by two flanking triple helices, with a long non collagenous N-terminal domain (with von Willebrand domains and fibronectin type III repeats) and a shorter C-terminal domain, also non collagenous. In order to form the anchoring fibrils, the C-terminal domains align in an antiparallel fashion, thus forming a dimer that undergoes a cleavage of the C-terminal domains themselves and a stabilization through disulfide bonds [Onursal et al., 2021]. This supramolecular organization allows the extension of the N-terminal domains from both ends of the fibrils, hence mediating the attachment of the basement membrane to the dermis [Burgeson, 1993].

- **Beaded filaments-forming collagen (type VI collagen)**: type VI collagen forms a unique microfibrillar network at the interface between the basement membrane and interstitial matrix. The monomers, characterized by N-and C-terminal globular domains and a central triple helical domain, can dimerize in antiparallel mode, interacting with each other both through the terminal domains and the triple helices; furtherly, dimers connect with each other thanks to the protruding globular domains, thus creating a tetramer stabilized by disulfide bonds. Tetramers, finally assemble end-to-end, form a peculiar "beaded" supramolecular structure. Besides its structural and binding roles, this collagen also acts as a signalling protein, as its C-terminal propeptide is the hormone endotrophin, that has a role in metabolic syndrome in humans. Moreover, is thought that type VI collagen could have the potential to serve as an early detector in the injury/repair response, influencing fibrogenesis, promoting the proliferation of mesenchymal cells, and inhibiting cell apoptosis [Onursal et al., 2021; Sun et al., 2024].

- **Fibril Associated Collagen with Interrupted Triple helix (FACITs – type IX, XII, XIV, XVI, XIX, XX, XXI and XXII collagens)**: FACITs represent the largest subcategory of the collagen superfamily [Ricard-Blum and Ruggiero, 2005]. Different types of FACITs can be expressed in different tissues, or display molecular diversity, for instance in their level of glycosylation [Patino et al., 2002; Ricard-Blum, 2011], but they are all united by the fact that they are always associated to the surface of pre-existing collagen fibrils, without forming fibrils themselves [Patino et al., 2002; Ricard-Blum, 2011], containing even less than 10% of "typical" collagenous domains [Onursal et al., 2021]. Their N-terminal domain possess a thrombospondin subdomain and sometimes also a von Willlebrand factor-type A-like domain (both mainly involved in cell interactions and cell adhesion [Whittaker and Hynes, 2002; Tucker and Adams, 2024]). FACITs are flexible molecules, and they are most likely involved in in the stabilization and integrity of the extracellular matrices in association with fibrillar collagens [Ricard-Blum and Ruggiero, 2005].

- **Membrane Associated Collagen with Interrupted Triple helix (MACITs) or Transmembrane Collagens (type XIII, XVII, XXIII and XXV collagens)**: membrane-associated collagens can occur in two different forms, respectively as transmembrane proteins or as soluble proteins [Ricard-Blum, 2011]. The former function as cell surface receptors and display a N-terminal domain located inside the cell (called endodomain), a hydrophobic transmembrane domain and multiple extracellular collagenous domains (called ectodomains); it is curious that, in the case of type XIII and type XVII collagens, the folding of the triple helix in the ectodomains proceeds in opposite direction to that of the fibrillar collagens, going from the N- to the C-terminus [Ricard-Blum and Ruggiero, 2005]. The ectodomains can undergo proteolytical shedding and released from the cell surface, becoming soluble proteins involved in cell behaviour regulation [Ricard-Blum, 2011; Ricard-Blum and Ruggiero, 2005]; in particular, they are classified as matrikines or matricryptins, which are bioactive fragments released from the ECM proteins, responsible of major biological processes [Ricard-Blum and Salza, 2014].

- **Multiplexins or Endostatin-producing Collagens (type XV and XVII collagens)**: multiplexins contain multiple triple helix domains frequently interrupted by non-collagenous domains and have a N-terminal thrombospondin domain; their C-terminal domains can be freed from the parent molecule by proteolytic cleavage and then act as molecular modulators in the ECM. More specifically, collagen type XV generates restin and collagen type XVIII produces endostatin, respectively; both of these molecules have major antiangiogenic properties. It has to be underlined that multiplexins can also carry glycosaminoglycan chains (notably, chondroitinsulphate in type XV and heparansulphate in type XVIII), thus belong not only in the collagen family, but are also recognized as proteoglycans [Ricard-Blum and Ruggiero, 2005; Onursal et al., 2021].
- **Other collagens (type XXVI, XXVIII and XXIX collagens)**: some of the most recently discovered collagens do not neatly align with the previous described categories, or exhibit features that can be found in more than one subfamily, making it difficult to assign them a specific classification. Type XXVI is a small collagen of only 438 aminoacids with two short collagenous domains, not very similar to any of the other described collagens; in humans, it is especially expressed in genitalia. Type XXVIII has a long triple helix domain flanked by von Willlebrand factor-type A domains, and is thought to be related to type VI collagen, as it shares some sequence homologies with it [Gordon and Hahn, 2010]. Conversely, type XXIX has a short triple helical domain surrounded by six N-terminal and three C-terminal von Willebrand factor–type A domains [Söderhäll et al., 2007].

1.1.4 Structure of collagen genes

About 40 collagen genes have been discovered in higher Vertebrates genomes, which can combine and give rise to the 29 distinct collagen types to date identified; also considering the presence of many collagen-like proteins (i.e., proteins structurally related to collagens), the typical vertebrate collagenome expands to more than 80 collagen and collagen-like genes [Parry et al., 2017] The first collagens to be identified, that is fibrillar collagens, were also the first to be described at a genetic level. It's intriguing to notice that the common molecular structure of the fibrillar collagen α-chains is reflected by the corresponding genomic pattern [Exposito et al., 2008]. In particular, sequences encoding for the major triple helix can be observed on the gene as an in-frame "cassette" type of organization [Ramirez et al., 1990]: in fact, the coding exons responsible for the triple-helical structure adhere to a precise organizational pattern characterized by a strict maintenance of both their quantity and dimensions. Such exons mostly measure 54 bp, multiples of 54 bp, or multiples of 54 – 9 bp (e.g., 45 bp, 99 bp, 108 bp, 162 bp and so on). This strongly indicates that the evolutionary development of the triple-helical coding domain occurred through events of duplications, fusions, and deletions involving an ancestral 54 bp unit responsible for encoding six Gly-Xaa-Yaa triplets, equivalent to 1% turns of the triple helix. An exception can be found in the COL1A1 gene, in which two exons are fused together [Ramirez et al., 1990; Exposito et al., 2008]. Regardless of their length, each one of these exons starts with an intact glycine codon and ends with an intact Y codon: this strongly suggest that the primordial fibrillar collagen gene probably originated thanks to repeated duplications of an ancestral exon of 54-bp. An unequal cross-over event could have caused the emerging of a novel modular unit of 45 bp [Exposito et al., 2008]. The discovery of viral and bacterial collagen-related structural motifs, with a homotrimeric conformation stabilized by six Gly-Xaa-Yaa repeats and with threonine as a substitute for hydroxyproline, supports this hypothesis [Ramirez et al., 1990; Ricard-Blum and Ruggiero, 2005]. Unlike the triple-helical domain, where the 54 bp exon

module appears to have been chosen as an optimal unit for chain construction, the evolutionary paths of the N- and C-propeptides seem to have taken separate and more flexible trajectories [Ramirez et al., 1990]. While the C-propeptide coding sequences are, however, usually divided into four exons, the sequences responsible for coding the N-propeptide exhibit more variable organizational patterns, resulting in differences in both the quantity and dimensions of the corresponding exons. The diversity in the numerical aspects of the N-terminal coding units appears to be more correlated with the final heterotrimeric assembly of their products than with the specific structure of each individual N-propeptide [Ramirez et al., 1990].

This kind of genetic organization has been found in invertebrates collagen genes, as well, except for insects (model: *Drosophila melanogaster*) and nematodes (model: *Caenorhabditis elegans*), in which the collagen genes contain very few introns [Runnegar, 1985]. Hence, it is probable that all collagen genes share a common ancestry, and the differences observed among various vertebrate and invertebrate collagen types likely arise from divergent evolutionary pathways that followed their initial divergence, with most of the diversification occurring during the early history of Metazoans [Runnegar, 1985]; the differences in invertebrate collagens could likely reflect differences in the evolution of the Protostomia and Deuterostomia genomes [Ramirez et al., 1990]. The collagen triple helix is an example of the diverse array of modules that emerged during the metazoan radiation, which made easier the assembly of extensive multi-modular proteins. The retention of the exon-intron organization within the region encoding the major triple helix of metazoan fibrillar α-chains suggests a robust selection pressure to maintain both the size and the Gly-Xaa-Yaa repeat, as crucial elements for fibril formation. Additionally, the relatively conserved C-propeptide coding sequence in both vertebrates and invertebrates indicates that the role of their products (intra- and intermolecular junctions, as well as N-linked glycosylations), constitute fundamental structural and functional elements in the molecular architecture of both vertebrates and invertebrates. An advantageous feature of the collagen module is its participation in oligomerization, thus functioning as a structural organizer that not only exhibits resilience against proteases but also facilitates the creation of multivalent supramolecular networks [Ramirez et al., 1990; Exposito et al., 2008].

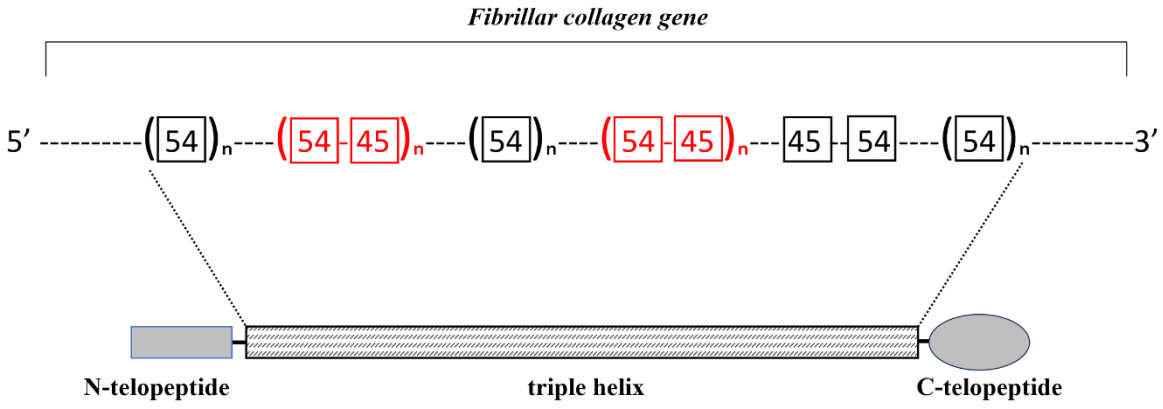


***Figure 5**: schematic representation of a generic fibrillar collagen gene.*

On the contrary, nonfibrillar collagens genes exhibit more diversity. Phenotypically, they give rise to various supramolecular structures and display α-chains with interrupted triple helical regions. Accordingly, the genes responsible for encoding these collagens deviate from the previously described fibrillar collagen gene exon model: in particular, they lack an evident conservation of

intron/exon organization. In general, the coding regions for the triple helix in these genes differ from those of fibrillar collagen genes by not having a significant number of 54 bp, 45 bp or multiples thereof exons; moreover, they frequently include exons that start with split codons for glycine [Vuorio and de Crombrugghe, 1990; Christiano et al., 1994; Boyd, 2012].

1.2 The collagens of Invertebrates

The greatest part of the knowledge regarding the molecular and functional characteristics of collagen primarily focuses on Vertebrates, particularly mammals. Nevertheless, approximately 97% of all life forms on Earth is represented by Invertebrates. This extensive biodiversity is likely to be mirrored in molecular diversity as well. It is reasonable to consider that the diverse forms of collagen, the hallmark of all Metazoans, would manifest in numerous distinct variations among Invertebrates, corresponding to the environmental adaptations these animals have evolved over their evolutionary history. Not every phylum of Invertebrates has been explored in this context yet. Collected information regarding the collagen types of the major phyla of Invertebrates is provided below.

***Phylum* Porifera**

Poriferans, commonly known as sponges, are the most primitive and basic group of multicellular animals. They separated from other Metazoans in the evolutionary timeline earlier than any other known animal *phylum*, whether currently existing or extinct. The lack of organs and highly organized tissues is made up for by specialized cells, surrounded by extracellular matrix. The Porifera *phylum* is categorized into four classes: Calcarea (which body support is guaranteed by $CaCO_2$ spicules), Hexactinellida, Demospongiae, and Homoscleromorpha (with a skeletal framework primarily consisting in silica spicules); in the subclass Keratosa, the mineralized skeleton is even absent and is substituted by macroscopically visible rigid fibers [Reitner and Mehl, 1996; Ehrlich et al., 2018]. The work published by [Ehrlich et al., 2018] provides a comprehensive review of the different categories of collagen that have been found in sponges till now:

- ***Spongin Short-Chain Collagens***: as mentioned above, the Demosponges belonging to the Keratosa subclass have a "mineral-free" fibrous skeleton, which is made of a halogenated scleroprotein known as spongin, involved in the attachment to the substratum and secreted by spongioblasts (a cell type deriving from the epithelium of sponges) [Ehrlich et al., 2018; Fassini et al., 2021]. Already in 1956 [Gross et al., 1956], studying the extracellular matrix of marine sponges, recognized the existence of the so-called "Spongin A" and "Spongin B": the former refers more generally to intercellular collagen fibrils, much more classic and common to all sponge species (see below), while the latter was named to the flexible organic skeleton made up of reticulate fibres typical of horny sponges, such as *Spongia officinalis*. The chemical and molecular composition of spongin is still under investigation [Ehrlich et al., 2018; Fassini et al., 2021]; however, it has – at least, partially – a collagenous nature, as residues of hydroxyproline were identified within it [Gross et al., 1956] (note that this result may vary, not only basing on the method of extraction, but also because of the variety of the Demosponges). Many authors agree on considering as the collagenous constituents of spongin a family of short, nonfibrillar collagens (thus renamed spongin short-chain collagens or SSCCs), not identified in Vertebrates, but found in the majority of the other Invertebrates [Aouacheria et al., 2006]. SSCCs exhibit a collagenous domain of about 120 Gly-Xaa-Yaa triplets, occasionally interrupted, and a non-collagenous C-terminal region (NC1) that shares similarities with both nematode cuticular collagen and the type IV collagen family [Fassini et al., 2021]. Given their resemblance with type IV collagens, which

are primarily associated with basement membranes, and given that basement membrane-like structures have been found only in the most recent group of marine sponges (i. e., Homoscleromorpha), SSCCs may share the same ancestral origin [Exposito et al., 2002].

- ***Collagen type IV-related proteins***: together with fibrillar collagens, type IV collagens are the only ones, among the 29 different types to date identified, to have been observed from sponges to humans. As more extensively described in previous paragraphs, this collagen is one of the main components of basement membranes in higher animals; basement membrane-like structures, in the Porifera *phylum*, have been described only in the Homoscleromorpha class, but type IV collagenous sequences were found also in the classes Calcarea and Demospongiae, devoid of the basement membrane structure [Aouacheria et al., 2006; Ehrlich et al., 2018]. Hence, both collagen type IV and its ancestral form, SSCCs, are thought to be linked to the shift towards multicellularity and that their presence subsequently facilitated the development of multicellular epithelial tissues [Fidler et al., 2017].
- ***Fibrillar collagens***: as type IV collagens, also fibrillar collagens are present almost in all *phyla*, from Porifera to higher Vertebrates, performing essential physiological functions. [Exposito et al., 2002] proposed that the genes responsible for producing an ancestral form of fibrillar collagen could have emerged at the onset of metazoans, prior to the divergence of sponges and eumetazoan lineages; following this hypothesis, the duplication events leading to the formation of fibrillar collagen clades A, B, and C occurred prior to the radiation of eumetazoans. Notably, the resemblance in the modular structure between sponges and humans is only conserved in the B clade of fibrillar collagens: this is coherent with the hypothesis that type V/XI fibrillar collagens played a foundational role in initiating collagen fibril formation. Many sponges provided of fibrillar collagens within their body have been identified in all the four classes of this *phylum*. Fibrillar bundles appear to be secreted only by lophocytes, which are highly specialized and polarized cells typical of poriferans. Collagen fibrils are involved in the formation and in the maintenance of the mesohyl structure, as well as in the mediation of cell-matrix interactions [Ehrlich et al., 2018]. The structural role of fibril-forming collagens can be declined differently depending on the sponge species: for instance, the demosponge *Chondrosia reniformis* relies all of its body support on collagen fibrils, being completely devoid of a mineral skeleton; in spicule-producing sponges, instead, collagen fibrils have been observed in the sponge ECM, strictly associated to the spicules themselves, where they can encase the components of the inorganic skeleton or even function as a template for the biomineralization process. The Hexactinellida class can produce very thin and long silica spicules, characterized by great flexibility, suggesting a role of collagen in the organic components within the spicule. A high fibrillar collagen content can also occur in the external asexual buds of some species (*Tethya lyncurium*, *Tethya sychellensis*) [Ehrlich et al., 2018].

Phylum Ctenophora

Together with poriferans, ctenophores are one of the most ancient metazoans; they are planktonic animals characterized by a gelatinous body provided ciliated comb rows used for locomotion [Churches et al., 2012; Draper et al., 2019]. Many species of ctenophores are adapted to live in deep sea environment, where low oxygen levels, scarcity in food and elevate hydrostatic pressure may have led the evolutive pressure into forming unique kinds of ECMs [Draper et al., 2019]. The information regarding the composition of ctenophore ECM is still limited; contrarily to poriferans, ctenophores show a considerable lack of ECM-related proteins conservation if compared to the other *phyla*. A collagenous nature is associated to both the mesoglea and the prehensile coiled *tentilla* found in many species of ctenophores [Mackie et al., 1988]; however, surprisingly, to date no fibrillar

collagens, nor SSCCs homologues, were found in this phylum. Studying the genome of *Pleurobrachia bachei*, [Churches et al., 2012] individuated 7 distinct type IV collagen α-chains homologues (five of which probably associated with gene duplication events), with an average length of 2200 amino acids and provided with multiple stretches (mainly, 13) of Gly-Xaa-Yaa repeats. It was possible to phylogenetically divide the 7 α-chains in two groups, respectively similar to that of α1-like and α2-like collagen type IV sub-families in humans. *P. bachei* exhibits the most extensive expansion of type IV collagens among all known animals, with seven different sequencing encoding as many different α-chains complete sequences. The closest comparison comes from vertebrate chordates, which possess 6 type IV α-chains. Moreover, *in situ* experiments revealed collagen expression across multiple cell types, indicating that ctenophores utilize collagen to a greater extent than previously observed in other basal organisms [Churches et al., 2012].

Phylum **Cnidaria**

The Cnidaria *phylum* is extremely biodiverse and encompasses five classes: Hydrozoa, Anthozoa, Scyphozoa, Cubozoa and Myxozoa. Cnidarians have a body wall structured with two layers of epithelial cells separated by an acellular material, the mesoglea, which contains collagen fibers distributed throughout. Despite the differences occurring across various cnidarian species, their collagens possess biochemical characteristics resembling either fibrillar or basement membrane collagens and are key molecules in the structural composition of the mesoglea, as well as for Ctenophora [Bairati, 2013; Orgel et al., 2017]. The basement membrane-associated collagens in this group share similar features with type IV collagen, reaffirming the ubiquitous presence of this collagen type across animals [Orgel et al., 2017]. For what concerns fibrillar collagens, cnidarian α-chains have the typical modular structure of eumetazoan fibrillar collagens, with a central major triple helix flanked by the N- and C-propeptides [Exposito et al., 2008]; in addition to the minor triple helix, the cnidarian N-propeptide is provided of a specific whey acid protein module (WAP), sometimes in association with a von Willebrand A domain [Fassini et al., 2021]. Besides that, two noteworthy and unique collagenous structures can be found in cnidarians: cord-like collagen fibrils (CLCFs) and nematocysts minicollagens. CLCFs were firstly isolated by [Orgel et al., 2017] from the reef-dwelling octocoral *Sarcopython ehrenbergi*: these fibers are extremely long, as they can measure up to tens of centimetres; in addition, they display a proteoglycan-rich organization and are not distributed within the mesoglea but are instead allocated in the mesenteries. Usually, the mesoglea of cnidarians contains short collagen fibrils surrounded by a thick matrix, allowing the mesoglea itself to withstand significant strains thanks to the sliding of the fibrils one on another; the presence of CLCFs, more similar to a vertebrate tendon or ligament in their structure, confers to *Sarcopython ehrenbergi* a much tighter organization, but their role in the octocoral physiology is still unclear. More widely known, instead, is the cnidarian nematocyst. Nematocysts are capsule-shaped secretory organelles, exclusive of cnidarians, used for predation and defence through the release of toxins [David et al., 2008]. The walls of the mature nematocyst are strong enough to contain a 150 bar osmotic pressure; the discharge of the nematocysts' content is extremely rapid, with an expulsion strength comparable to that of firearms and able to penetrate even the cuticle of crustaceans [Özbek et al., 2011]. The greatest part of the nematocyst capsule walls is constituted by a family of uniquely short collagens of 8–16 Gly-Xaa-Yaa repeats, thus named minicollagens. The domain architecture of minicollagens is well conserved among the Cnidaria and is constituted by a central short triple helix comprised between variable polyproline stretches and typical terminal cysteine-rich domains (CRD), which show a CX_3-CX_3-CX_3-CX_3-CC cysteine pattern, at both the N- and the C-terminus. The CRDs domains promote the formation of -S-S- intermolecular links, hence they are likely responsible for the high stability of the nematocysts' capsule. Despite having the same cysteine patterns, the N- and C-terminal CRD domains give origin to different structures: in the mature molecule, the N-terminal CRD maintains its

original fold, while the C-terminal CRD usually gains a new fold, creating a molecular polarity that allows the macromolecular assembly of the capsule. The CRD domains, as variable elements, can be used to study the phylogeny of Cnidarians from a molecular point of view; this led to the definition of three primary categories: Group I, with preserved N- and C-terminal CRDs (such as in Hydra vulgaris); Group II, where C-terminals CRDs are duplicated, and the second CRD domains has a different cysteine arrangement than the first; Group III, in which both N- and C- terminals domain have altered cysteine arrangements [Özbek et al, 2011]. It is possible to notice an increase of nematocyst complexity from Anthozoa to Medusozoa, which is reflected by an increase of the number and complexity of minicollagens proteins, as well. This aspect primarily involves the diversification of the CRD domains: novel CRD motifs, in fact, lead to a modification of the minicollagens network and, consequently, to the formation of new and more diversified capsule structures [David et al., 2008].

Phylum Platyhelminthes

The Platyhelminthes *phylum* encompasses four classes of acoelomates flatworms: Turbellaria (free-living), Monogenea, Trematoda and Cestoda (parasites). The number of studies regarding the molecular characterization of the platyhelminths' collagen is definitely lower if compared with other groups. In general, evidence supports the presence of collagens analogous to the Vertebrates' fibrillar (I, II, II, V, XI) and type IV collagens across all the classes [NCBI]. Two remarkable facts can be highlighted in two of the most studied species of platyhelminths, i.e., *Cystycercus cellulosae* (the larval form of the human parasite *Taenia solium*) and the planarian *Schmidtea mediterranea*, known for its extraordinary regenerative abilities. In 1981, Torre-Blanco and Toledo isolated and analyzed a fibril-forming collagen from *Cystycercus cellulosae*, that displayed a composition similar to that of higher *phyla*; however, surprisingly, no amount of hydroxyproline was detected, a condition observed only in a few other cases, such in the collagenous silk of a sawfly; in spite of this, its thermal stability is comparable to that of other platyhelminths collagen, like *Fasciola sp.*'s – containing normal levels of hydroxyproline – , probably because of a high degree of cross-linking [Torre-Blanco and Toledo, 1981]. In the case of *S. mediterranea*, most focus was given to its type IV collagen, not because of its molecular architecture (it has the typical characteristics of mammalian type IV collagen α-chains), but especially because of its physiological role. *S. mediterranea* maintains a wide population of multipotent stem cells, called neoblasts, even in its adult form, making it possible to regenerate its tissues even when cut in multiple fragments, thus creating new individuals. This process is regulated primarily by type IV collagen, which affects differentially the promotion or the suppression of neoblast proliferation [Chan et al., 2021].

Phylum Nematoda

Nematodes have thread-like, unsegmented bodies, and a generally simple body structure; nevertheless, this *phylum* display an astonishing biodiversity [Kiontke and Fitch, 2013]. A representative feature of the *phylum* is the cuticle, i. e. an exoskeleton which forms a protective layer between the animal and the outer environment, defining the shape of the nematode [Johnstone, 1994]; despite exhibiting a great variability in its morphology, cuticles are primarily constituted by small collagen-like proteins characterized by an extensive amount of nonreducible cross-links, more likely tyrosine-tyrosine cross-links [Cox, 1992]. The Rhabditidae *Caenorhabditis elegans* stands out as one of the most studied nematodes and is considered as a model organism especially in developmental biology; therefore, the greatest part of information we have inherently this *phylum* comes from this species. In general, nematode collagens exhibit significant structural and size divergences from those found in Vertebrates or other Invertebrates *phyla*, thus representing a specialized evolutionary adaptation unique to nematodes. Notably, in *C. elegans* 185 collagen genes have been found

distributed throughout its genome [Teuscher et al., 2019], of which 173 encode for cuticle collagens; this is remarkable, considering that the total number of collagen genes identified in the human genome is 44, almost four times less [Haq et al., 2019]. In order to better classify the high number of collagens in *C. elegans* (thus considered representative of the *phylum*), four major groups can be identified [Tauscher et al., 2019]:

- ***Group 1 – conserved vertebrate-like collagens***: a few collagens in the genome of *C. elegans* can be considered homologues to their vertebrate counterpart. While fibrillar collagens appear to have been lost, type IV collagens are instead well conserved in this animal. Another collagen present in the basement membrane of *C. elegans*, but also and especially around its neuronal system, is CLE-1, similar to collagen types XV and XVIII, with which shares a sequence identity of 14% and 18%, respectively; this collagen type possesses one or two fibronectin-type-III-like domains, a laminin G-like domain, a short, interrupted triple helix, and an endostatin domain. Finally, the so-called COL-99 of *C. elegans* is similar to the vertebrate MACITs, being hence classified as a transmembrane collagen, as it comprehends a smaller cytoplasmic domain and a larger extracellular region composed by 10 collagenous domains.
- ***Group 2 – collagen domain-containing proteins with mammalian orthologues***: in addition to group 1, four other collagenous proteins with mammalian orthologues were identified: cof-2 and unc-122 (resembling gliomedins), clec-222 and clec-223 (resembling collectins). The cluster of gliomedin-like proteins, composed by cof-2 and unc-122, have a N-terminal transmembrane domain, followed by a triple helix domain of 15 triplets, and a C-terminal olfactomedin-like domain. The cluster of collectins, comprehending clec-222 and clec-223, have short collagenous domains and 1 or 3 C-type lectin domains.
- ***Group 3 – non-cuticular collagens with no clear orthology to mammalian collagens***: belong to this group four collagen types (MEC-5, COL-55, ROL-8 and COL-135) that set apart from all the other categories, as they don't cluster neither with type IV collagens, nor with vertebrate-like or cuticular collagens. Even if their role has not been clarified yet, they may have a specialized function in some extracellular matrix not related to the basement membrane ore the cuticle. MEC-5 is a medium sized collagen, characterized by a short collagen domain, that resembles the N-terminal minor helix typical of fibrillar collagens, followed by a major triple helix with multiple interruptions, with no other predicted domains. COL-55 and ROL-8, despite being quite similar to the cuticular collagens (Group 4), lack some features, such the minor N-terminal triple helix or the typical cysteine knots; instead, they display a predicted transmembrane domain overlapping with a predicted N-terminal cuticular domain. Finally, COL-135 has a peculiar composition for a collagen: it contains a signal peptide, three minor triple helices and a wide major collagenous domain; while proline residues are under-represented, there's an over representation of lysine and aspartate residues, and in fact 109 out of the 198 triplets are Gly-Xaa-Lys repeats.
- ***Group 4 – cuticle collagens***: roughly 80% of the nematode cuticle is constituted by collagenous proteins. If compared to vertebrates fibrillar collagens, nematode cuticle collagens are considerably smaller and less complex in their architecture; the assembly of these collagens is also different from the majority of other organisms: while certain post-translational modifications, such as proline hydroxylation, appear conserved between vertebrates and nematodes, other processes, including N-terminal proteolytic processing and extracellular crosslinking mechanisms, are entirely distinct. Cuticular collagens usually display some conserved features, but their most typical characteristic is a major collagenous domain of 37 to 43 Gly-Xaa-Yaa repeats, flanked by non-trimeric N- and C-terminal peptides, typically measuring around 100 and 20-30 amino acids in length, respectively, and equipped with cysteine knots. A minor N-Prohelix

of about 10 Gly-Xaa-Yaa triplets, stabilized by a cysteine knot as well, is also present. Cuticular collagens can be furtherly classified into five main clusters (from A to E) based on the number of interruptions in the main triple helix (respectively, either 0, 1, 2, 3, or n3 interruptions) [Novelli et al., 2006; Tauscher et al., 2019].

Phylum Mollusca

Mollusca is a very broad and biodiverse *phylum* and comprehends eight classes, three reunited under the Aculifera clade – namely, Solenogastres, Caudofaveata, and Polyplacophora – and four belonging to the Conchifera clade – Monoplacophora, Gastropoda, Bivalvia, Scaphopoda and Cephalopoda –, basing on the absence or presence of a proper conch or shell, respectively [Schierwater and DeSalle, 2021]. However, from what concerns the composition and the features of the ECM, the most studied species are those with a commercial value, especially the edible ones (e. g., mussels, scallops and octopuses), from which derives the greatest part of information about this knowledge area. In general, the fibril-forming interstitial collagens (type I-, II-, III-, and V-like) and basement membrane collagen (type IV-like), are present throughout the entire phylum [Engel, 1997]. As in the other organisms that produce a biomineralized skeleton (such as sponges and corals), also in molluscs which possess an external or internal shell collagen serves as a template for biomineralization and can be part of the organic matrix of the shell itself [Ehrlich, 2010; Marin, 2020]. The ECM composition of some species has a remarkable economic relevance, for instance in the case of the abalone (*Haliotis* sp.): it has been demonstrated that the collagen biosynthesis in this species has a defined seasonality, with higher peaks during winter and lower levels in summer [Yoneda et al., 2000]; the toughness of abalone meat is strictly correlated with its collagen content, that is therefore a determining factor in its organoleptic profile [Øiseth et al., 2013]. A type of remarkable form of mollusc collagen is that which is found in octopuses: the unique muscle texture of these animals is harder if compared with other marine organisms, because of the close union between cells; this is a consequence of the high crosslinking levels of the collagen fibers and of the consequent accumulation of pyridinoline (Pyr). Pyridinoline derives from the oxidative deamination of lysine and hydroxylysine by lysyl-oxidase and is a crucial crosslinking molecule in collagen fibers. Due to the high concentration of Pyr, octopuses' collagen results more resistant to degradation than that from other marine species [Tapia-Vasquez et al., 2019]. However, the most noteworthy collagen typology among the Mollusca phylum is certainly the one that constitute the byssus threads, secreted by marine mussels to adhere to solid surfaces. Byssal threads are 2-4 cm long with diameters of 0.10 to 0.15 mm and can be considered as a sort of "tendons" in which one extremity inserts into the byssal retractor muscles at the base of the foot of the mussel, while the other extremity is projected outside the animal and is strictly attached to a hard surface thanks to an adhesive plaque. Each thread is characterized by a gradient of tensile capacity: in particular, the proximal end is highly elastic and can extend by up to 160% (this is outstanding considering that, normally, collagenous structures have an extensibility of less than 10%), making it a strong shock absorber; on the other hand, the distal end is way stiffer and more compact. From the proximal to the distal extremity, each thread can be subdivided into four morphologically distinct areas: a stem, which denotes the point of attachment to the animal, an elastic proximal thread, a stiff distal thread and an adhesive plaque. Overall, the byssus' collagen is extensively crosslinked and contains high levels of hydroxyproline; it is possible to isolate from the byssus two collagen-like proteins, namely Col-P and Col-D, both homotrimers, which are complementary distributed along the length of the thread, with Col-P predominating at the proximal end and Col-D prevailing at the distal end: likely, this distribution is directly correlated with the physical properties of the byssal thread. In fact, Col-P presents a collagen-like domain flanked by elastin-like structures, being the first known protein to contain both collagenous and elastin-like domains; considerably, before the discovery of Col-P, elastins were thought to be absent in

Invertebrates. Instead, Col-D shows silk-fibroin-like domains and it is therefore presumed that these sections contribute to the stiffness of the threads by forming crystalline β-sheet structures, similar to those found in natural silk fibers. Finally, another striking feature of the byssus is the presence, by the level of the adhesive plaque, of significant concentration of the amino acid 3,4-dihydroxyphenyl-l-alanine (DOPA), which residues are essential to the chemisorption of the thread to surfaces underwater, thanks to the formation of covalent bonds between the adhesive plaque and the substrate [Coyne et al, 1997; Deming, 1999].

***Phylum* Annelida**

The Annelida *phylum* encompasses segmented worms and is divided into three main classes: Polychaeta (mostly, marine sandworms), Oligochaeta (mostly, earthworms) and Hirudinea (leeches). In each of these classes are present at least three types of collagenous structures: the type IV-like collagens of the basement membrane, common to all Metazoans, the "classical" banded collagen fibers, morphologically identical to those found in vertebrates and many invertebrates species, and non-banded cuticle fibers which can be also observed in Nematodes [Bairati, 2013]. Once again, cuticular collagens are the most noteworthy ones in this *phylum*, like in the case of Nematoda, as they don't have a counterpart in Vertebrates. The collagenous cuticles of annelids appear like highly specialized structures and provide a flexible, however strong, exoskeleton to the animal. Differently from the nematodes, which shed their cuticles as they grow, in the case of annelids cuticles undergo continuous growth through a process of degradation and deposition. These cuticles consist of a layered, plywood-like organized extracellular structure of large collagen fibers that run along the worm's longitudinal axis, alternating directions at nearly right angles to each other. The angle between fibers in adjacent layers of the cuticle likely adjusts to accommodate local changes in diameter during the animal's movement. Evenly spaced between each row of fibers it is possible to notice some sort of *microvilli*, which are thought to play a role in transferring spatial information from cells to cuticle, in assisting the maintenance of the regular orientation of the fibers themselves, and in facilitating the transport of molecules, such as newly synthetized collagen or breakdown products from aged fibres. Surprisingly, this fibre arrangement allows to create an enduring skeleton and consents the worm's body to expand and contract, while at the same time lacking in molecular crosslinking [Murray et al., 1981]. The cuticle collagens from annelids show differences respect to other groups of collagens also from a biochemical point of view: in fact, they have a very high Hyp/Pro ratio, are highly glycosylated (mainly with the mono-, di-, and trimeric form of galactose and especially at the level of serine or threonine residues), have a low number of hydrophobic amino acid residues, and lack of cysteine and hydroxylysine; generally, polychaetes cuticles contain higher levels of polar amino acids and lower levels of hydrophobic amino acids compared to oligochaetes cuticle collagens. Apart from their unusual chemical compositions, annelids' cuticle collagens also exhibit significant differences in molecular size if compared to other collagens, as they are five/six times longer than classical interstitial collagens (in particular, 1500 nm in polychaetes and 1800 nm in oligochaetes, respectively), making them one of the longest collagen protomers known so far. An additional distinctive characteristic of annelid cuticle collagen fibers is their absence of banding when observed under an electron microscope [Murray et al., 1981], suggesting the presence of straight and apparently non-coiled collagen fibrils, despite the ability of annelids of producing banded fibrils in other anatomical contexts. One exception to this last feature is represented by the marine polychaete *Alvinella pompejana*, which cuticle includes instead supercoiled collagen fibrils [Gaill and Bouligand, 1987]. *A. pompejana* has been subjected to many studies, because is probably the most heat-tolerant metazoan organism known to date [Cary et al., 1998], as it lives by the hottest part of the hydrothermal ecosystems, on the wall of chimney-like structures. Besides heat, the hydrothermal vents environment is considered extreme also because of the pressure (about 200-300 atmospheres),

the toxicity of the fluid which comes out of the vents and the low levels of dissolved oxygen [Sicot et al., 2000]. To survive into such conditions, *A. pompejana* developed several adaptations, for instance concerning its ECM. At an ultrastructural level, the collagen network of its cuticle shows fibrils with different degrees of helicoidal coiling [Gaill and Bouligand, 1987], as they can be either rectilinear or strongly helicoidal, with all the possible intermediate states. This kind of structure, scarcely present or absent in other annelids, is called "supercoil" and, given its important biomechanical implications (such the improved extensibility in physiological conditions), it could be an adaptation to barometric stress. For what concerns thermal stability, all the known stabilizing factors are amplified in *A. pompejana*: in fact, it was proven a higher presence of proline content, frequence of stabilizing Gly-Xaa-Yaa triplets with hydroxyproline in the Y position, and glycosylation levels, especially on threonine residues. Thanks to these post-translational modifications, the isolated collagen of *A. pompejana* remains stable at 45°C and is one of the most thermostable fibrillar collagen known up to now [Pradillon and Gaill, 2007]. Moreover, enzymes involved in collagen synthesis, such as prolyl-hydroxylases, exhibit contrasting functionality in *Alvinella pompejana* if compared to other metazoans. While typically the activity of this enzyme increases with higher oxygen levels, in this species it only remains active in hypoxic conditions, suggesting that this polychaete not only thrives in the highest temperatures recorded for marine invertebrates but also possesses metabolic adaptations suited for low-oxygen environments [Le Bris and Gaill, 2006].

***Phylum* Arthropoda**

The *phylum* Arthropoda is the widest group within the entire animal kingdom, of which represent nearly the 80% of all species. It comprehends four main classes: Crustacea (such as shrimps, crabs, and isopods), Hexapoda (insects), Arachnida (spiders and sea spiders) and Myriapoda (millipedes and centipedes) [Kershaw, 1983]. Despite being present in several regions of the arthropod body, in this group collagen does not perform the ubiquitous connective tissue role as is evident in both Vertebrates and other Invertebrates and, in general, connective elements are scarce [Bairati, 2013]. For these reasons, the collagen of this group is not as investigated as in other *phyla*, also because of its low biotechnological interest and inefficient yields of extraction [Prajaputra et al., 2024]. The greatest part of information available regards mostly edible species (i.e., many crustaceans) or model organisms (such the fruit fly *Drosophila melanogaster*). Generally, it is possible to recognize both fibril-forming and non-fibrillar collagens in all arthropods. Arthropods' fibrillar collagens occur in two forms: a) typically banded fibrils, with a clear periodicity; or b) thin filaments with faint or indistinct banding, described to date in horseshoe crabs, crustaceans and insects. These collagens have the classical amino acid composition of their Vertebrates' counterparts, with some peculiarities that appear also in other Invertebrates' collagens. In crustaceans and insects, remarkable aspects are the presence of half-cystine, the low content of alanine, and the high content of hydroxylysine (several times as large as those of Vertebrates). Two forms of fibril-forming collagens were moreover observed in crustaceans, named AR-I and AR-II collagens (notably, AR-II is richer in Val, Leu and Cys), which are similar to vertebrate type V collagen and widely distributed among different species [Mizuta et al., 1994]. As many other invertebrate collagens, insects' and crustaceans' ones are plenty of carbohydrate moieties, with glucose and galactose which are present as galactosyl- and glucosylgalactosyl- residues linked to the hydroxyl group of hydroxylysine. Cockroaches' fibrillar collagens have a higher melting point if compared to other Arthropoda and resemble type I vertebrate collagen [Francois et al., 1980]. Interestingly, the fibres of opilionids are reinforced by encrustations of crystalline calcium oxalate, making these arachnids the only known case of natural calcification of connective tissue in the entire Arthropoda *phylum* [Bairati et al., 2013]. On the other hand, the features of arthropods' non-fibrillar collagens are less known and studied; however, they seem to be similar to type IV vertebrate collagens (especially in *Locusta migratoria* and *Drosophila melanogaster*) [Bairati et al., 2013].

Phylum Echinodermata

The members of the Echinodermata phylum are morphologically characterized by a pentameric symmetric adult body plan and are distributed into five classes: Crinoidea (sea lilies and feather stars), Asteroidea (sea stars), Ophiuroidea (brittle stars), Echinoidea (sea urchins), and Holothuroidea (sea cucumbers) [Arnone et al., 2015]. In each *subphylum* are present collagens homologous to all groups found in Vertebrates, except for type IV collagen [Dolmatov and Nizhnichenko, 2023], generally with a high level of glycosylation [Bairati, 2013]. A recent analysis has demonstrated the presence of a) at least from four to six genes encoding for fibril-forming collagens, homologous to type I, II, III, V and XI collagens and sometimes characterized by the presence of a N-terminal von Willebrand factor type C (vWC) domain, absent in Vertebrates; b) one FACIT gene, which product may be involved in the processes that allow the change of the mechanical properties of the mutable collagenous tissue, typical of echinoderms (see later); c) two network-forming collagen genes; d) one microfibrillar collagen gene with an EMI domain; e) one multiplexin gene; moreover, the same authors have discovered also collagens that evidently differ from those of Vertebrates and cannot be associated to any of the known groups [Dolmatov and Nizhnichenko, 2023]. As mentioned above, a remarkable aspect of echinoderms is the ability to drastically, but reversibly alter the mechanical properties of their extracellular matrix in a span of a few seconds, under control of the nervous system in response to traumas, pressure, or other environmental stimuli: this is particularly evident in sea cucumbers, that are normally flabby but that can rapidly become stiff to the touch. This reaction occurs in collagenous tissues devoid of muscular structures, which have therefore been named "mutable collagenous tissues" (MCTs). MCTs are absent in any other animal phyla, likely being an Echinodermata synapomorphy; however, there are remarkable homologies between MCTs and the collagenous mesohyl of some demosponges, such *Chondrosia reniformis* (see *2.2.1.3*) which become rigid if manipulated, changing the tensile properties of their tissues under a non-neural, physiological mechanism. Even if the molecular events that make possible such kind of ECM alteration are still under investigation, is clear that the occurring of these events is facilitated by a crucial characteristic of MCTs that sets it apart from vertebrate collagenous tissues: while interfibrillar crosslinking in Vertebrates remains permanently stable, in MCTs it is labile and subject to physiological regulation, also with the assistance of many other ECM-related molecules besides collagens [Wilkie et al., 2021].

Phylum Tunicata

Tunicates are marine filter-feeding animals traditionally divided into three classes: Appendicularia, Thaliacea, and Ascidiacea [Stolfi and Brown, 2015]. For what concerns collagen, ascidians (especially *Ciona intestinalis*) are the most studied subphylum within Tunicata. The presence of homologues to vertebrate type I, IV, V, and IX (FACIT) has been documented by several authors. In particular, ascidians' fibrillar collagens are generally in 3-hydroxyproline, hydroxylysine, and glutamic acid, but with modest levels of lysine and alanine if compared with classical vertebrate collagens; they resemble the collagens from arthropods because of their extremely high contents of carbohydrate and glycosylated hydroxylysines. Over 90% of the carbohydrate units attached to hydroxylysine were identified as the disaccharide Glc-Gal, while less than 10% were found to be the monosaccharide Gal. This feature (i. e., the combination Glc-Gal-Hyl), frequently occurs in marine invertebrate collagens, apart from the annelid cuticle [Kimura et al., 1972]. In 2022, Mizuta et al. extracted from the muscle layer of the ascidian *Halocynthia roretzi* two molecular species of collagen, that the authors named AS-I and AS-II collagen, respectively. Both AS-I and AS-II lack in disulfide bonds and display a considerably lower percentage of alanine than other marine invertebrate collagens; however, the overall difference in the amino acid composition between the two supports a genetical distinction; type AS-II collagen may be homologous to vertebrates' type V collagen or

crustacean type AR collagens [Mizuta et al., 2002]. Finally, the FACIT-like collagen identified by [Vizzini et al., 2008], named *Ci*-type IX-Col by the authors, may play a role within a network of non-fibrillar collagens, contributing to the organization of the extracellular matrix and responses related to defence.

1.3 Biotechnological usage of collagen

The progressive advancement in understanding the chemistry and the biology of collagen has gone hand in hand with the increase in its biotechnological application. Nowadays it is possible to exploit collagen not just as a biomaterial in its entirety, either by itself or in combination with other polymers, but also its derived bioactive components.

1.3.1 Properties and requirements of collagen as a biomaterial

A biomaterial is a substance, of artificial or natural origin, that can be employed either independently or as a component within a complex system, usually to interface with biological systems, for instance, to influence the course of therapeutic, diagnostic, or cosmetical procedures [Hudecki et al., 2019]; an extended definition of biomaterials includes substances deriving from a biological source and exploited not only for biomedical purposes but also in industrial contexts, such as for the creation of food packaging. Collagen, as well as gelatine – its denatured form –, is an extremely versatile natural polymer, possessing a plethora of biochemical and biophysical properties, which makes it a unique biomaterial [Patino et al., 2002], adaptable to many biotechnological purposes. Depending on the designed application, the specific requirements of a collagen-based biomaterial may vary; however, this molecule has some intrinsic features that make it prone to be used in many contexts. Collagen is a primary component of extracellular matrices, where it creates a complex framework and functions as a natural substrate for cells' growth, in conjunction with many other ECM molecules; for this reason, it is in general highly biocompatible. In particular, it is non-toxic and has low immunogenicity, especially in its purified and undenatured form: this is because the primary antigenic loci of collagen are located at both the C- and N-terminal regions of the molecule at the level of the telopeptides, which are cleavaged in the mature protein (see 1.1.2), and because of the ability of its helical structure to mask potential antigenic determinants [Patino et al., 2002]. The Gly–Pro–Hyp and Gly–Phe–Hyp motifs within collagen triple helices are recognized by cell receptors, improving affinity with adherent cells [Ghomi et al., 2021]; moreover, the presence in the collagen molecules of RGD fibronectin-binding sites, recognized by integrin receptors, are involved in the activation of intracellular signaling pathways able to induce epithelial-mesenchymal transition phenomena [Heino and Kapyla, 2009; Yuan et al., 2015], related to cellular adhesion, growth, differentiation, and migration [Patino et al., 2002]. Another intrinsic characteristic of collagen is its good biodegradability: normally, this biomaterial is susceptible to degradation within the human body thanks to the activity of matrix metalloproteinases (MMPs) at physiological conditions of pH and temperature. The peptides deriving from collagen degradation (notably, from type I and type III collagen) have been proven to act as chemotactic stimuli for neighbouring fibroblasts, which are thus recruited when the repairing of damaged tissues is needed [Ghomi et al., 2021]. Thanks to its chemical structure, collagen is equipped with high tensile strength and generally low extensibility, depending on the number of intra-/intermolecular cross-links and interactions with glycoproteins and proteoglycans, which altogether provide stability to the molecule [Patino et al., 2002]. Depending on the utilization purpose, a different level of biodegradability can be required; this property is usually inversely proportional to stability, which may vary relatively to the collagen source: in fact, as

explained before, collagens can be more or less thermally and mechanically stable depending on their chemical features. In the biotechnological field, collagen can be used in its native form (i.e., extracted maintaining the intact structure of fibrils) or reconstituted by in vitro fibrillogenesis after extraction. Compared to the native form, the self-assembled collagen in vitro has much less resistance not only because of the absence of natural lysyl oxidase-mediated cross-linking but also because the in vivo process also revolves around the presence of so-called organizers (e.g., fibronectin and integrin) and nucleators (e.g., collagen types V and XI) which can control and regulate the assembling process [Lee et al., 2020]. In any case, a collagen-based biomaterial must be sufficiently stable not to undergo denaturation or degradation during extraction, processing, and before having fulfilled its role during utilization. The biodegradability of collagen can be reduced or even suppressed, and consequently its stability can be improved, by artificially modulating the cross-linking degree [Patino et al., 2002]; cross-links can be introduced in vitro biologically (i.e., through the action of enzymes such as transglutaminase), physically (using UV-irradiation at 254 nm or dehydrothermal treatment, DHT), or chemically (with formaldehyde, glutaraldehyde, carbodiimides, polyepoxy compounds, acyl azides, or hexamethylene diisocyanate). An alternative approach foresees the stabilization of collagen biomaterials by mixing them with polycationic molecules like chitosan [Chattopadhyay and Raines, 2014]. Each one of these methods has a different outcome in terms of tensile strength, enzymatic resistance towards different enzymes, and cytotoxicity of the final product; therefore, the artificial cross-linking method must be carefully chosen depending on the utilization purposes. Finally, collagen is particularly appreciated for producing biomaterials due to its high hydrophilicity, capacity for extensive swelling and deswelling [Panduranga Rao, 1996], and haemostatic properties [Dinescu et al., 2019]; platelets can adhere to the surface of collagen, thus triggering clotting and haemostasis. This process primarily relies on the degree of cross-linking, on the positive charges on side chain base groups, and on conformational structure of collagen: in fact, chemically blocking the free carboxylic groups leads to a significant reduction in haemostatic activity, while the denatured form of collagen (i.e., gelatine) fails to activate platelet aggregation [Dinescu et al., 2019].

1.3.2 Applications of collagen

1.3.2.1 Collagen as a biomaterial: fields of employment and formulations

The intrinsic properties of collagen make this polymer one of the most widely used biomaterials, especially in the biomedical field, precisely for its hydrophilicity, biocompatibility, low antigenicity, flexibility, chemotaxis ability, and biodegradability [Lin et al., 2019], as better explained in the previous paragraph. Specifically, collagen-based biomaterials have become increasingly prevalent in the following areas:

a) Medical Applications:

- *Regenerative medicine and tissue engineering*. The goal of regenerative is not only to heal tissues after traumatic injury, but to get to an as complete as possible restoration of the tissue's native function, often through the stimulation of the body's natural healing processes [Pawelec et al., 2016]. Tissue engineering is a possible approach to regenerative medicine that focuses on designing and developing substitutes that can replace or support damaged or diseased tissues or organs; it involves the use of scaffolds, cell cultures, and bioactive molecules to recreate *in vitro* functional tissues that can subsequently be implanted into the body to promote tissue repair and/or regeneration, with the ultimate goal to restore the healthy physiology of the tissue itself. Being collagen the main structural protein of ECMs, collagen-based scaffolds are widely used in this field, as they mimic the natural ECM environment and thus provide support for cell adhesion,

proliferation, and differentiation; it is therefore exploited for various applications, such as skin grafts, bone grafts, and cartilage repair; it can also be combined with other biomaterials, growth factors or cells to improve the finals result [Shekhter et al., 2019]. To date, using this kind of approaches, collagen-based devices have been successfully employed in bone and cartilage repair, muscle tissue repair, nerve tissue regeneration, vascular grafts, corneal reconstruction, dental surgery, and many more [Dinescu et al., 2019].

- *Wound healing*. Wound healing is the process by which the body repairs damaged or injured tissue to restore its structural integrity and function; this process can be assisted by collagen more than any other biomaterial: being naturally a key molecule in the dermis, it is able to provide a suitable substrate for fibroblasts, thus promoting cell proliferation and a higher wound healing rate. However, besides proliferation, collagen is beneficial also for other stages of wound healing, including haemostasis, inflammation, and tissue remodeling, thanks to its hydrophilic and chemotactic properties. Currently, there are numerous commercially available dressings designed to provide a conducive environment for tissue regeneration in the treatment of many types of wounds [Ghomi et al., 2021].
- *Surgical procedures devices*. Collagen is extensively employed in surgical interventions due to its biocompatibility and capacity to integrate with the body's tissues without causing immune responses; in this case, it serves as a biomaterial not only for the realization of devices involved in the final goal of the surgery (such as grafts, implants, and dressings), but also for the creations of objects necessary during the intervention itself, assisting the surgeon's actions. This biopolymer is in fact often the primary component of haemostatic agents, due to its higher efficiency if compared to traditional devices made from cellulose or non-woven textile materials [Fronchek et al., 2020], and of surgical sutures [Hu et al., 2020] which exploit collagen's tensile strength.

In each of these perspectives, collagen can be used alone or combined with other polymers (such as chitosan, polycapronolactone, or alginate [Lin et al., 2019]), with inorganic elements (such as hydroxyapatite or silica [Heinemann et al., 2013; Dinescu et al., 2019]), and/or with bioactive molecules (e.g., antioxidants, antibiotics [Shekhter et al., 2019], and growth factors, such as the transforming growth factor-β (TGFβ), the vascular endothelial growth factor (VEGF), the platelet-derived growth factor (PDGF), and the basic fibroblast growth factor (bFGF) [Hu et al., 2020]). The variety of possible combinations is extremely wide, and the final product can be very specific for its intended application.

b) Cosmetic applications: collagen is one of the main components of the skin and its presence is linked to skin's conditions and appearance; therefore, it is a very popular ingredient in cosmetic products. When used in its original conformation, collagen is not able to penetrate into the dermis because of its molecular dimensions, and it only remains at the level of the epidermis; however, given its ability to bind water, its effect after external application is positively related to the maintenance of the hydration levels: the film-forming properties of collagen are in fact used for covering skin or hair, reducing trans-epidermal water loss, and preventing damage by surfactants [Sionkowska et al., 2020]. Nowadays, one of the most commercially successful employments of collagen in the cosmetic industry is the subcutaneous injection of collagen and the use of collagen-based prosthetic implants for the correction of aesthetical and dermatological defects and to reduce the symptoms of aging. Also in this case, many different formulations are possible, for instance in combination with elastin [Sionkowska et al., 2020].

c) ***Pharmaceutical applications***: collagen results particularly eligible for drug-delivery systems (DDS). A drug delivery system is a method designed to transport therapeutic substances – drugs or any bioactive substance – to specific target sites within the body, in a controlled and precise manner, thanks to an adequate carrier. DDSs are designed to enhance the efficacy and the bioavailability of the employed substances, while minimizing side effects and improving patient compliance. Thanks to its versatility, collagen can act as a carrier for drug delivery, basically in every possible formulation previous described, i.e. can occur in inserts, shields, particles, aqueous injections, sponges, films, monolithic devices, micelles, and micro-/nanoparticles. These devices can be loaded with antibiotics, antimycotics, anti-inflammatory drugs, painkillers, hormones, growth factors, antitumoral, and anaesthetic drugs. The association between the carrier formulation and the designed drug depends on the specific purpose of the DDS; in fact, the release rate can be modulated by intervening on the mesh size of the device and on its diffusivity capacity, which relies on a combined effect of both electrostatic and hydrophobic interactions, depending in turn on the differences in amino acid composition of the chosen collagen source [Friess, 1998].

d) ***Other biotechnological fields***: besides the biomedical field, collagen finds use in many areas, of which one of the most important is the food and beverage industry; here, collagen is primarily used both as a food supplement and an additive in drinks and solid foods, for instance in desserts and confectionery, to improve texture, stability, and nutritional value [Hashim et al., 2015]. Very diffused are collagen-based edible films and coatings, which are applied on or within foods in thin layers by wrapping, soaking, brushing, or spraying. This practice consent to create a barrier against the migration of oxygen, moistures, and solutes, provides structural integrity and vapour permeability to the food products and prolongs the shelf life of products. Besides, this kind of coating can function also as a carrier of bioactive substances (such as antioxidant and antimicrobials), as well as artificial flavours and colours [Hashim et al., 2015]. Finally, collagen can be extracted from various sources (i.e., calf skin or rat tail) and highly purified to be employed in biotechnological research activities, for instance as a control sample or as a substrate for cell cultures and enzymatic activities.

The multifunctional nature of collagen-based biomaterials is because they can occur in a multitude of physical forms, which differ in their degree of hydration, tridimensionality, density of fibrous network, and so on. The most common physical forms in which collagen is employed can be summarized as follows:

- ***Hydrogels***: the term "collagen hydrogel" refers to a water-based colloidal gel in which collagen is the main solute – with variable possible concentrations –, while water is the diluting medium. Collagen hydrogels can be realized in many ways, with different outcomes in terms of physical-chemical properties. Traditionally, a collagen hydrogel can be prepared *via* a self-assembling process, neutralizing with PBS and NaOH an acidic solution of collagen monomers, thus promoting the aggregation of molecules into fibrils, and subsequently triggering a fibrillogenic-induced gelation by warming the solution to physiological temperatures (37° C). Another method is the ion leaching technique, which includes the addition of NaCl into the collagen solution and its subsequent removal in order to create a porous network within the hydrogel. Some more advanced methodologies foresee the gelation process to occur into customized micro-moulds or the bioprinting technology, with which is possible to precisely manage the morphology of the hydrogel by directly manufacturing its internal 3D structure. In any case, the biggest advantage of a collagen hydrogel is that cells and/or bioactive compounds can be homogeneously incorporated directly into it during the fabrication process, before gelation. The orientation of the collagen fibers within the biomaterial can be used to adjust transparency and mechanical

properties, and/or to guide the direction of cell growth; in fact, stem cells or functional components are often loaded within collagen hydrogels for tissue regenerations, to mimic the desired elements of native tissue function. An effective method to control collagen fibrils orientation is to mix magnetic nanoparticles together with the collagen hydrogel, as they can be thus controlled remotely applying an external magnetic field. Collagen hydrogels are particularly employed in DDSs, surgery procedures (notably, vitreous substitutions), coats for prostheses and moisturizers [Lin et al., 2019].

- ***Sponges***: a collagen sponge is a porous, three-dimensional structure made from purified collagen fibers, usually by freeze-drying a collagen suspension in which were previously been injected air bubbles, or obtained by 3D printing; differently from hydrogels, in this case the hydration levels are very low or even null, and they can vary depending on factors such as its manufacturing process, intended use, and storage conditions. Usually, these devices are conserved in a dehydrated or lyophilized state to provide a better stability and shelf-life and hydrated only when needed for use. Collagen sponges can usually incorporate and retain significant amounts of liquids, thus increasing their volume and recreating a moisturized environment which is particularly suitable for wound healing, 3D-cultures, or as scaffolds in tissue regeneration applications. Thanks to their hydrophilicity, they are also often used as a haemostatic agent or to absorb tissue exudates if used in adherence to wounds. The main parameter that may vary among different collagen sponge devices is the porosity range, on which depend the transport of nutrients and cell attachment, growth, and migration [Ghomi et al., 2021].
- ***Microspheres***: collagen microspheres are spherical particles, filled or hollow, employed for tissue regeneration and pharmaceutical purposes. Indeed, they can function as a stable and injectable delivery device for different types of cells, drugs, and bioactive factors, while providing large attachment surface for cells and easily accelerating oxygen and nutrients diffusion through the spaces between the microspheres themselves. Some off the most common methods to prepare collagen microspheres are direct aliquoting (in which droplets of a collagen solution are disperses onto a non-adherent surface and the temperature of the system is raised to induce the gelation of the droplets), emulsification (in which droplets of a collagen solution are dispersed into a continuous oil phase solution and then recovered by centrifugation), and freeze-drying (collagen solutions are dropped into liquid nitrogen for a quick solidification); however, a better way to precisely control the microspheres' size and shape is to apply microfluidic technology. Cells can be added within the collagen solution and be directly encapsulated into the microspheres [Lin et al., 2019].
- ***Membranes***: collagen membranes are two-dimensional, flexible collagen sheets with a thickness usually comprised between 0.01 and 0.5 mm, which can primarily serve as scaffolds for cell attachment, proliferation, and tissue regeneration, as they provide a supportive environment for cells to grow and differentiate. It is possible to obtain collagen membranes by mould-casting, freeze-drying, or electrospinning a purified and generally cross-linked collagen solution; these different methods of production allow to obtain products with various thickness, mesh tightness, and pore size depending on the intended application. Collagen membranes are very versatile as they can be furtherly modified and functionalized after their production to improve their biocompatibility, bioactive properties, and resistance; the functionalization of a membrane foresees the incorporation within its collagen matrix of bioactive molecules, drugs (like antibiotics, hormones, antitumoral compounds, and so on), or enzymes, which are intended for local action [Chak et al., 2013]. To date, the factors that represent the main limit in collagen membranes utilization is their intrinsic low stiffness and rapid degradation *in vivo*, as after

implantation they are rapidly degraded by collagenases, bacterial proteases, and macrophage-derived enzymes [Ren et al., 2022]. Even if their harshness and resistance towards enzymes can be improved by physical, chemical, or enzymatic cross-linking, on the other hand these procedures can have a negative impact on biocompatibility and reduce the range of molecules that can be incorporated within the membrane itself.

- ***Hollow fiber tubes***: collagen hollow fiber tubes are tubular structures that can vary in diameter, length, and thickness and offer a unique architecture able to efficiently support cell growth, facilitate nutrient diffusion, and mimic the natural environment of living tissues; these devices allowed great developments in various biomedical applications, especially in the three-dimensional engineering of artificial tissues and organs. They can be fabricated primarily through electrospinning, 3D-printing, or exploiting the gelation ability of collagen, strictly controlling pH and temperature conditions, as well as the collagen concentration in the starting material [Prade et al., 2022; Iwamoto et al., 2022]. To improve the mechanical properties of the tubes, the starting collagen material (usually in a hydrogel form) can be enriched with other materials (as Ca-alginate) and/or cross-linked. It is possible to furtherly functionalize the final product with the addition of bioactive molecules, such as growth factors [Shen et al., 2015; Ishibashi et al., 2023]. Collagen hollow tubes are especially suitable for the biomimicry of blood vessels and renal tubes and facilitate the creation of pre-vascularised engineered tissues as well as the promotion of neovascularization around the artificial tissues themselves [Ma et al., 2014].

1.3.2.2 Collagen-derived molecules

The biotechnological exploitation of collagen goes besides its usage as a whole molecule. A large portion of its market, in fact, regards collagen-derived molecules which are largely employed especially in the pharmaceutical, nutraceutical, and cosmetical field. Collagen-derived molecules are primarily represented by collagen hydrolysates/peptides, matrikines, and matricryptins.

The enzymatic hydrolysis of collagen or gelatine during food processing, gastric digestion, or through the action of endogenous and exogenous proteases can produce peptide fragments, usually between 2 and 20 amino acids long, that are inactive within their parent proteins, but that exhibit a wide plethora of bioactive properties once released, also thanks to their high bioavailability: in fact, it has been proven that collagen peptides, especially those with C-terminal Pro or Hyp residue, can be transported through the intestinal epithelial monolayer, enter the bloodstream and there exhibit their bioactive properties [Fu et al., 2019]. Some of the main potentialities of collagen/gelatine-derived peptides are listed below:

- ***Antioxidant activity***: while the precise mechanism behind how they exhibit antioxidant activity remains incompletely understood, it has been demonstrated that collagen-derived peptides are able to scavenge free radicals, neutralize singlet oxygen, or chelate metal ions [Kitts and Weiler, 2003; Chen et al., 2018]; is therefore possible that the total antioxidant effect of peptides could be dependent on the synergic effects of those mechanisms. The antioxidant properties are influenced by the sequence, composition, conformation, presence and positioning of certain amino acids such as amino acids, such as Tyr, Phe, Trp, His, Met and Pro [Banerjee and Shanthi, 2016].
- ***Bone and joint health improvement***: *in vitro* studies show that collagen peptides can stimulate osteoblast proliferation and differentiation, can up-regulate the expression of the bone-related genes in osteoblasts (including type I collagen, alkaline phosphatase, osteocalcin and osteopontin) and can improve calcium absorption [Fu et al., 2019].

- ***Skin health benefits***: some *in vitro* experiments showed that collagen-derived peptides can stimulate proliferation and motility of fibroblasts and their hyaluronic acid production [Matsuda et al., 2006; Chai et al., 2010], and even increase the density and diameter of collagen fibers [Matsuda et al., 2006]. The bioefficacy of these molecules has also been proven in clinical trials which demonstrated positive properties such the enhancement of hydration [Choi et al., 2014] and the elasticity of the skin [Proksch et al., 2013]. Differently from its whole counterpart, hydrolysed collagen in much more soluble in water, and it can be added to many cosmetic formulations [Sionkowska et al., 2020].
- ***Anti-hypertensive effect***: thanks to their amino acid composition, especially because of the presence of C-terminal hydrophobic residues, such as Pro, Phe, and Val, some collagen-derived peptides possess a strong ACE (angiotensin-I converting enzyme) inhibitory effect [Aluko, 2015] and can therefore be considered as a potential safer alternative to other drugs for lowering blood pressure [Fu et al., 2019].
- ***Dipeptidyl peptidase-IV (DPP-IV) inhibitory effect***: a wide amount of *in vitro* studies report that collagen peptides can inhibit the action of dipeptidyl peptidase-IV (DPP-IV), thus retarding the degradation of glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide-1 (GLP-1), and therefore leading to insulin secretion; this would potentially consent, in an *in vivo* system, to control the haematic glucose level in the treatment of type-2 diabetes [Fu et al., 2019].
- ***Immunostimulatory activity***: some studies report an immunostimulatory capacity in several cell lines such as RAW 264.7 macrophages [Sae-leaw et al., 2016] and Caco-2 endo-epithelial cells [Chen et al., 2017].
- ***Wound-healing promotion***: research conducted on animals suggests that collagen hydrolysates have the potential to accelerate wound-healing [Sionkowska et al., 2016]. In particular, Sato and coworkers [Sato et al., 2020] highlighted that the collagen-derived, low molecular weight dipeptide Pro-Hyp is a growth-initiating factor for fibroblasts, hence playing a significant role in wound healing and showing therapeutic potential for the treatment of chronic wounds.
- ***Other bioactive properties***: in addition to the above-mentioned bioactivities, collagen peptides have been found to display various other physiological effects, such as inhibiting renin activity, lipid-lowering activity, neuroprotective effects, probiotic effects, and anti-microbial properties [Fu et al., 2019].

In vivo, many ECM components can be actively cleavaged by controlled proteolysis, leading to the generation of fragments called matrikines, that serves as signaling molecules; meanwhile, during ECM remodeling, some otherwise "cryptic" peptides can be exposed or released, allowing them to exert various physiological or physio-pathological effects, including the stimulation of angiogenesis, tissue remodeling, wound healing, inflammation, tumor growth, and metastasis; these molecules are identified as matricryptins [Ricard-Blum and Ballut, 2011]. Collagen-derived matrikines and matricryptins (from both fibrillar and non-fibrillar collagens) have been recently taken into consideration for their potential pharmaceutical uses: in particular, they have highly promising efficacy in terms of anti-angiogenic and anti-tumoral properties [Ricard-Blum and Ballut, 2011; Monboisse et al., 2014].

1.3.3 Sources of collagen for biotechnological purposes

Among collagens, fibrillar ones (especially type I) are the most exploited for biotechnological purposes, due to their abundance and their advantages. However, this material can be obtained from

a wide range of different sources, which differ in terms of availability, costs/yield rates, and chemical-physical properties: as explained in previous paragraphs, the amino acid compositions of collagen types can significantly differ between species, especially affecting thermal stability, viscosity, and crosslinking density [Davison-Kotler et al., 2019]. Notably, collagen can be obtained from natural or artificial sources.

- ***Natural sources***: the natural sources of collagen are of course animals, especially the species consumed for food, both from the terrestrial and marine environments. To date, the most common and economically advantageous sources for collagen production are bones and skin obtained from the remains of the industrially bred terrestrial animals after slaughter, especially bovine and swine species, but also poultry (in this case, collagen can be obtained not only from the adult animal's tissues, but also from the chicken embryo sternal cartilage and from eggshell membrane [Hashim et al., 2014]). Another common mammalian collagen source is rat tail collagen, because of the abundance of rodents employed in research activities [Davison-Kotler et al., 2019]. Other terrestrial sources of collagen are kangaroo tail, duck feet, equine tendon, alligators' bone and skin, birds' feet, sheep skin, and frog skin [Avila Rodríguez et al., 2017]. Although the mentioned sources are generally cheap and easy to acquire, their use don't come without risks: fallacies during purification could result in infectious diseases transmission, like in the case of bovine spongiform encephalopathy (BSE) transmission via prion-contaminated bovine-derived collagen scaffolds, or the cases of bird flu and swine flu. An additional problem comes from the fact that approximately 4% of the population is allergic to bovine- and porcine-derived collagen and is subjected to immunogenic reactions [Davison-Kotler et al., 2019]. Besides health concerns, this kind of sources could not be suitable for every consumer, because of ethical and socio-cultural reasons. The possible adverse effects linked with animal-derived collagen could be overcame by using human-derived collagens; however, the physical characteristics of human feasible sources (such as placental and skin-derived collagens) are influenced by factors like ethnicity, age, environmental conditions, and the genetic profile of the source tissue, leading to a high variability among samples; for instance, aged collagen is associated with increased intermolecular crosslinking, heightened resistance to collagenase, reduced elasticity, and diminished osmotic swelling capacity. Due to the mechanical and chemical diversity observed in human-derived collagen, along with concerns regarding infection, immunogenicity, and the inconsistent quality inherent in animal-derived collagen, the interest towards exploring alternative options keeps on growing [Davison-Kotler et al., 2019]. In this context, a promising natural source is marine collagen, which solves most of the problems related to mammalian and avian sources, at least in terms of health and religious concerns; in fact, it has no risk of disease transmission, as the FDA considers it as GRAS (i.e., "Generally Recognized as Safe") [Avila Rodríguez et al., 2017] and it is not forbidden by any religion. To date, the main economically profitable sources of marine collagen are the waste coming from fishing industry, fishing bycatch, and occasional blooms of marine animals such as jellyfish. The most common marine species feasible for marine collagen production belong to the fish group, both elasmobranch fish (blacktip shark) and teleost fish (mostly Atlantic salmon, chum salmon, Pacific cod, Baltic cod, black drum, sheepshead seabream, and olive flounder); in this case, collagen is extracted from fins, skin, scales, and bones, and is prevalently a type-I collagen. It is noteworthy that fish scales, together with their inorganic component of calcium carbonate, form a naturally composite collagen [Sionkowska et al., 2020]. However, collagen can be obtained also from many invertebrates, mainly jellyfish (jelly blubber, flame jellyfish, Atlantic Sea nettle, barrel jellyfish) – which collagen mostly resemble types II, IV, and V – but also molluscs (like squids), echinoderms (like sea cucumbers or the invasive species Crown-of-thorns starfish), and sponges [Avila Rodríguez et al., 2017; Davison-Kotler et

al., 2019]. Marine-derived collagen displays different chemical and mechanical properties if compared to mammalian collagen, especially regarding lower melting point, lower viscosity at a given concentration in solution, lower solubility in water, higher fibrillar proportions of glutamic acid and alanine, and lower concentrations of proline. The biological properties of marine collagens are positive, as collagen scaffolds sourced from marine organisms demonstrate high biodegradability, allowing for gradual replacement by regenerated tissue over time. They also exhibit low immunogenicity and high biocompatibility, making them suitable for various biomedical applications. However, it's important to consider that denaturation temperatures significantly lower than the human core body temperature of 37.5°C may require additional crosslinking to ensure stability, especially for applications such as wound healing [Davison-Kotler et al., 2019].

- ***Artificial sources***: instead of recurring to natural sources such as animals or marine organisms, it is possible to recreate collagen-like materials that are artificially produced in laboratory by mimicking the structure and properties of natural collagen through chemical or biotechnological methods. The main advantages of this approach regard the absence of any risks and the customizability of the final product; on the other hand, these methods are still very expensive and not able to provide molecules having the same chemical complexity of the natural sources. There are two main approaches to generate artificial collagens: the first involves chemical synthesis, where collagen-like peptides are chemically synthesized step-by-step; these peptides can be designed to mimic specific regions or motifs of natural collagen, allowing for precise control over their structure and function. An example is the work of O'Leary and coworkers, which managed to achieve the multihierarchal self-assembly of the collagen mimetic peptide $(Pro\text{-}Lys\text{-}Gly)_4(Pro\text{-}Hyp\text{-}Gly)_4(Asp\text{-}Hyp\text{-}Gly)_4$ from triple helix, to nanofiber, to hydrogel and to successfully employing it to recreate a synthetic biomimetic matrix useful in thrombosis [O'Leary et al., 2011]. The second type of approach regards the recombinant DNA technology to produce collagen peptides or fragments in genetically engineered organisms such as bacteria, yeast, or mammalian cells. Recombinant collagen expression models have been effectively showcased in various biological systems, encompassing both prokaryotic expression systems, like *Escherichia coli* [Rutschmann et al., 2014], and eukaryotic expression systems, like yeasts [Liu et al., 2006; Pozzolini et al., 2015], plants [Stein et al., 2009], insects [Tomita et al., 2003], and in mammalian cell cultures [Geddis et al., 1993; Woodley et al., 2004]. However, creating high protein throughput models has posed a significant challenge, especially regarding the incorporation of post-translational modifications in the mature protein [Davison-Kotler et al., 2019].

1.3.4 Employments of marine collagen

As mentioned earlier, there is an increasing demand for collagen sources that could represent a valid alternative to mammalian and avian ones. Marine collagen is one of the most promising options, as it possesses the majority of the positive chemical-physical and biological properties of the terrestrial feedstock (i.e., it is abundant, biocompatible, and biodegradable) but without carrying the same risk as the latter. However, the main disadvantage regards it's an overall reduced thermal/mechanical stability, due to a lower number of cross-links and imino acid (such as proline and hydroxyproline); in fact, except for some cases, marine animals generally live at a lower temperature respect to terrestrial ones, therefore they don't have the same physiological requirements. For this reason, it is advisable to choose species which live in a warmer environment [Tassara et al., 2023]; moreover, this inconvenience can be quite easily overcame applying some modifications like artificially introducing inter-chain crosslinks [Silva et al., 2014; Coppola et al., 2020; Pozzolini et al., 2020]. With the necessary precautions, marine collagen (both collagen itself as a biomaterial and its

derived molecules) can find employment in the same numerous fields as the "classical" one, especially in tissue engineering, wound healing, drug delivery, cosmetics, and nutraceuticals [Felician et al., 2018]. The most used among marine collagen is fish collagen. For instance, fish collagen 2D-membranes can exhibit significant potential for clinical use because of their highly porous structure, dense cell seeding capability, ability to allow nutrient and oxygen passage, excellent biostability, and minimal immune response compared to other naturally derived biomaterials [Chandika et al., 2015]; in association with bioactive glass, fish-derived collagen can be electrospinned in nanofibers and used as wound dressing, demonstrating the ability to protect against infection and to promote wound healing and skin regrowth in both *in vivo and in vitro* studies [Zhou et al., 2017]. Three-dimensional sponge-like structures have also been tested comparing growth and proliferation of fibroblast and keratinocytes with a rat model, showing a faster wound healing process in the fish collagen treated samples [Pal et al., 2016]. Even if not yet too widely explored, also fish collagen-based drug delivery systems are showing good potential for biomedical applications; Cao and co-authors, for example, produced a scaffold-controlled release system for skin tissue engineering that exhibited a tuneable protein release rate depending on the fish collagen concentration, good biocompatibility, and ability to promote fibroblasts proliferation and skin tissue regeneration [Cao et al., 2015]. Moreover, fish collagen has been successfully tested as a substrate for cell cultures for human mesenchymal stem cells (hMSCs) [Matsumoto et al., 2015] and murine preosteoblasts cells (MC3T3-E1) [Capati et al., 2016]. Finally, fish-collagen hydrolysates are extensively exploited in the nutraceutical and cosmetical fields [Subhan et al., 2020]. Besides fish, jellyfish-derived collagen is another economically valuable source with great biotechnological potential; in fact, collagen solubilized from various jellyfish species exhibited properties and biological effects on human cells comparable to mammalian type I collagen [Addad et al., 2011], and can play a role in a multitude of fields. For instance, cross-linked jellyfish collagen-based sponges exhibit haemostatic properties, making them suitable for haemostatic materials and wound-dressing applications [Cheng et al., 2017]; porous collagen scaffolds derived from jellyfish collagen show biocompatibility with primary human fibroblasts and endothelial cells, making them suitable for tissue engineering applications [Song et al., 2006; In Jeong et al., 2007]. This material seem especially suitable for cartilage tissue engineering, as supports pre-osteoblasts and nasal septal chondrocytes growth [Bermueller et al., 2013; Rastian et al., 2018]; recently, a *Rhyzostoma pulmo* type II collagen-based scaffold was used in combination with nanoreservoirs of the growth factor TGF-β3 and human stem cells, creating a device for articular cartilage repair [Keller et al., 2017]. The collagen of the same species was also employed in an unusual application: it has been cross-linked to a designed amine thrombin aptamer, thus creating a hybrid apta-sensor for the clinical detection of thrombin blood [Derkus et al., 2016]. In addition, also jellyfish collagen-derived peptides have a myriad of useful applications, as they have been demonstrated to have ACE-inhibitory effects, antioxidant properties, UV-protective properties, and can stimulate wound-healing by increasing the production of chemotactic factors in fibroblasts [Liu et al., 2016; Sionkowska et al., 2020; De Domenico et al., 2019]. Among other marine collagen sources besides fish and jellyfish, are commercially attractive only the ones who are already exploited for human consumption, like some echinoderms, as their alimentary waste can be transformed into a valuable resource of so-called "blue biomaterials"; this is the case, for instance, of sea urchins and sea cucumbers, which possess a unique mutable collagenous tissue (see 1.2) from which is possible to create membranes for guided tissue regeneration applications, even with better results if compared to a bovine model [Ferrario et al., 2017], and dermal template which have been recently tested in vivo for wound healing purposes [Carolo et al., 2023]. The same goes for cephalopods: for instance, Antarctic squid's mantle-derived collagen was used for the creation of biofilms [Uriarte-Montoya et al., 2010], while Peru squid cartilage derived-type II collagen have been efficiently tested in vivo, performing intra-articular injections that were able to relief osteoarthritis in rat models [Dai et al.,

2018]. However, the marine environment is a huge pool of molecular biodiversity that is still largely unexplored; as presented in 1.2, this concept is reflected in the multitude of different collagens observed in marine invertebrates, each one of them could potentially represent a valuable resource for biotechnological applications. The industrial exploit of many interesting marine species wouldn't be neither economically nor environmentally sustainable; however, some marine collagens, yet characterized and tested *in vitro* but still not available for industrial utilization, could at least be inspirational for biomimicry purposes. This is particularly the case of poriferan collagen: to date, it has not yet been possible to extract collagen molecules isolated using standard procedures, but only intact fibril extracts, likely due to the numerous cross-links between molecules that render the fibers extremely robust. Nevertheless, sponge-derived collagen offers many potential biotechnological usage options like those previously described. Sponge fibrillar extracts (alone or combined with other biopolymers or other types of collagens) can be used as a basis for preparing membranes primarily serving as scaffolds for various types of cell cultures [Pallela et al., 2011; Pozzolini et al., 2021]. The so-called spongins obtainable from marine sponges (see 1.2) can be employed as scaffolds for tissue engineering and in the development of bioinspired materials; in addition to biomedical applications, spongin-based scaffolds have been successfully used as adsorbents for various dyes and as supports for enzymatic immobilization. It has been demonstrated that some spongins (in their native structure) are thermostable up to 260°C, which opens the door to the possible development of new advanced composite materials [Ehrlich et al., 2018]. Among poriferans, *Chondrosia reniformis* is probably the most studied species for what concerns collagen; its potential applications are better listed in paragraph 2.4.

1.4 Extraction, purification, and characterization of collagen

1.4.1 Main methods of extraction and purification

The properties of collagen-based materials are influenced by the source of collagen as well as by the method of its purification and further treatment [Sionkowska et al., 2020]. Given the extreme diversity of tissues and collagen types, on one hand it is necessary to individuate the appropriate procedure to optimize the yield and the quality of the extraction, also depending on the further intended use of the material, while on the other hand it is important to develop standard extraction methods, to guarantee homogeneity in the features of the final product for all types of collagen and its different sources. The aim of the extraction is to remove all the non-collagenous matter and recover collagen as the final product [Lin et al., 2015]. Regardless of the source of collagen, the extraction procedure is usually carried out through pre-determined steps, listed below, each one of them can be adjusted in relation to the source and the further aim of the extracted material:

1. ***Pre-treatment***: this step, even if not always strictly necessary, is intended to prepare the starting material to facilitate and improving the yield of the actual extraction, as well as the quality of the material. The pre-treatment of the sample should weaken the covalent intermolecular crosslinks between collagen molecules, while leaving the collagen chains untouched; it should also defat and/or demineralise the tissue, when needed. Usually, the tissue is cut into smaller pieces of about 1 cm^2 or less; then, it can be subjected to one or more different pre-treatments accordingly to the feedstock characteristics:

- *Acidic pre-treatment*: for acidic pre-treatment, pieces of tissues are soaked in a diluted acid (usually acetic acid or hydrochloric acid) at a controlled temperature. This allows the tissue to expand up to three times its initial volume, while hydrolysing cross-links. acid permeates the

skin causing it to swell to between two- to three-times its initial volume and hydrolyzes the crosslinks. This kind of pre-treatment is suitable for tissue that have a low degree of fibre intertwinement, such as porcine and fish skins [Schmidt et al., 2016].

- *Alkaline pre-treatment*: alkali pre-treatment is performed using diluted sodium hydroxide (NaOH) or calcium hydroxide ($Ca(OH)_2$), for a time span that depends on the thickness of the sample and that can last up to several weeks. This method is usually preferred to the acidic one, as consent to eliminate non-collagenous proteins, lipids, pigments, and other unwanted material. The concentration of the alkali solution must be calibrated in order to be sufficient to remove non-collagenous material, but not too high, not to lose part of the acidic soluble collagen fraction within the tissue or degrade the native structure of it [Schmidt et al., 2016].
- *Other additional pre-treatments*: some samples, especially those rich in fats or mineral components, could require additional handling. Substances like butyl alcohol may be used to solubilize fat and pigments, while chelating agents such ethylene diamine tetra acetic acid (EDTA) can be used for demineralization [Matinong et al., 2022].

2. ***Extraction***: the step of actual extraction is crucial, as – in addition to the chosen pre-treatment – affects the physical-chemical collagen properties, like the average length of the polypeptide chains, solubility, viscosity, thermal and mechanical stability, emulsifying capacity, and water retention; therefore, it can be designed depending on the desired yield and the properties of the final product [Senadheera et al., 2020]. On the other hand, also the starting feature of the material (like the animal type and age) could influence the result of the extraction. This step must be carried out at a controlled low temperature (4 °C) to avoid collagen degradation. Many different techniques are successfully exploited:

- *Acidic hydrolysis*: one of the most common methods for collagen extraction is acidic hydrolysis, conducted under continuous stirring for 24-72 hours, which can be effectively carried out with both organic acids (like acetic, chloroacetic, citric and lactic acid) and inorganic acids (such as hydrochloric, sulfuric, and nitric acid). In an acidic environment, collagen molecules have a net positive charge, and the resulting repulsive force between them allows an easier molecular separation [Matinong et al., 2022].
- *Alkaline hydrolysis*: this method is less common that the previous one, because it can cause the rupture of collagen fibrils and the loss of amino acids such cysteine, histidine, serine, and threonine; in fact, it is prevalently used for the treatment of leather processing waste [Yang and Shu, 2014]. As acidic hydrolysis, also in this case the extraction is conducted under continuous stirring for a variable period of time, using sodium hydroxide, potassium hydroxide, calcium oxide, calcium hydroxide, or sodium carbonate as extractants [Suo-Lian et al., 2017].
- *Salt solubilization*: even if solutions of neutral salts (like citrates, phosphates, sodium chloride, and Tris-HCl) are applicable in collagen solubilization, this method is rarely used. This is due to the fact that collagen dissolves al low salt concentrations, but precipitates at salt concentrations higher than 1.0 kmol m^{-3}, creating an inconvenient limitation [Yang and Shu, 2014].
- *Enzymatic hydrolysis*: to overcome some of the limitations of traditional extraction techniques, enzymatic hydrolysis-based methods have been developed and optimized, to be used alone or in combination with some of the traditional chemical methods [Yang and Shu, 2014]. Enzymatic hydrolysis offers many advantages respect to chemical methods: first, it has a better selectivity towards non-collagenous proteins, thus causing less harm to collagen and

consequently increasing both the yield and purity of extraction. Moreover, although more expensive than chemicals, enzymes can be used under mild conditions: this guarantees a lower corrosion of the equipment on the long term, a lower consume of energy, produces less chemical waste, and produces a final product with lower salt content [Yang and Shu, 2014]. Collagen extraction can be pursued through various proteolytic enzymes, which may be of animal origin (for example, trypsin and pepsin), of vegetal origin (like bromelain, papain, and ficin), or even of microbial origin (such as collagenase and proteinase K). However, only pepsin, trypsin and papain can cleavage at the level of the non-helical portion of the collagen molecule, without impacting on the triple helix; in particular, the extracted pepsin-soluble collagen tends to have a higher purity, because this enzyme is able to hydrolyze non-collagenous proteins very efficiently. This is the reason why the hybrid extracting method using acetic acid together with pepsin is one of the most used procedures [Matinong et al., 2022].

- *Ultrasound-Assisted Extraction*: during collagen extraction, ultrasound application with a frequency of at least 20 kHz creates strong agitation in the liquid, leading to increased mass transfer and chemical reaction rates; this results in an improved yield and a reduced required time of extraction, especially in association with enzymatic treatment. In particular, the amplitude level is inversely proportional to the duration of the extraction time and directly related to the purity and the yield of the extract [Ran et al., 2014; Senadheera et al., 2020]. Ultrasound treatment is simple, cheap and reduces the need for corrosive chemicals, but too much amplitude can potentially damage the collagen structure, especially in combination with alkaline treatment; therefore, this kind of treatment must be optimized for the specific collagen source in terms of ultrasound amplitude, duration, concentration of enzyme or other chemicals, and temperature [Schmidt et al., 2016].
- *Other Extraction Methods*: collagen extraction can be assisted by other techniques, like microwaves – which has been found to deeply penetrate into tissues, aiding the disruption of cells and thus the release of collagen, especially together with enzymatic hydrolysis – and mechanical agitation – which improves all the methods of extraction listed before [Lin et al., 2010; Khong et al., 2018].

3. ***Purification***: once extracted, collagen must be isolated from its extraction mixture, containing unwanted substances, to obtain a higher degree of purity and quality. Many purification methods have been developed, including filtration, centrifugation, chromatography, and precipitation techniques [Mazzola et al., 2008].

- *Filtration/ultrafiltration*: This method foresees the removal of contaminants by letting the sample pass through a permeable medium which is able to retain contaminants or solid particles; in the case of collagen, filtration can be carried out with polyacrylonitrile/ hydrophilically modified polyacrylonitrile membranes, or polyethersulfone membranes. The filtration technique is overall energy saving and environmentally friendly; however, at industrial levels, it requires high costs of maintenance of the membranes to avoid biofilm creation on the filtration devices [Awang et al., 2020].
- *Chromatography*: this method can occur through affinity or ion-exchange. Affinity chromatography is based on reversible interaction between the chromatographic matrix and the protein that is needed to be recovered; this interaction must be specific for the desired protein, so that the other components can leave the system. The target molecule is then collected in a concentrated form by altering the binding conditions with the chromatographic matrix. This method seems suitable only for recovering collagens derived from recombinant systems, for

which ensures a really high purity in the extracts; however, the yield of production is not sustainable for large scale and long-term production [Awang et al., 2020]. Ion-exchange chromatography works through electrostatic interactions between charged protein groups and a solid matrix – a resin – which can be based on cation exchange or anion exchange. Sato and co-authors purified a porcine type V collagen from a pepsin digest using a filter paper-based diethyl aminoethyl cellulose (FBP DEAE), obtaining high yield and purity; however, this method is not recommended for industrial purification of collagen, as it is not easy, it requires expensive chemicals and equipment, and is exposed to a high risk of contamination [Sato et al., 2003].

- *Precipitation*: precipitation stands out as a traditional purification method and is extensively used in the processing of biomolecules. After extraction, a crude collagen extract typically contains neutral salts and non-collagenous proteins: to isolate the target protein (or to eliminate the unwanted ones), it is possible to employ a technique known as "salting out", which involves the addition of a high concentration of salt to the crude extract. The efficacy of salting out collagen, or other proteins, from a crude extract is influenced by several factors, including temperature, pH, ionic strength, the type of salt used, and the characteristics of the proteins involved. Typically, collagen exhibits lower solubility compared to contaminating proteins, making it more prone to precipitation through salting out. However, to ensure selective precipitation and minimize co-precipitation of other proteins, salting out conditions must be carefully chosen, also considering that salting out ability varies among different anions and cations present in salts. Anions generally demonstrate higher salting out efficiency than cations. Among anions, the salting out effectiveness decreases in the following order: $F^- \approx SO_4^{2-} > HPO_4^{2-} > CHCOO^- > Cl^- > NO^{3-} > Br^- > ClO^{3-} > I^- > ClO^{4-} > SCN^-$. Similarly, among cations, the salting out effectiveness diminishes in the following sequence: $NH^{4+} > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+}$. Neutral salts produced by combining a strong anion and a cation from the aforementioned series have the ability to precipitate collagen, which is then recovered through centrifugation and cleaned from the residue salt content with a cycle of dialysis. [Awang et al., 2020; Matinong et al., 2022].

The choice of the purification method mainly relies on the source of the collagen itself [Awang et al., 2020]. Regardless of the most standard procedures listed above, it is important to underline that every source of collagen has a specific method of extraction that fits it the most in terms of yield and purity, and that can be opportunely modified based on requirements. For instance, in the case of less mainstream collagens – such as poriferan collagens –, some precautions or modifications may be needed, like the employment of chelating agents (like EDTA), protein denaturants (like urea), or strong reducing agents (like or 2-mercaptoethanol) during the extraction phase, to facilitate the extrusion of collagen fibers from the tissues [Pozzolini et al., 2018].

1.4.2 Analyses for in vitro characterization

Once purified, collagen must be characterized to assess the efficiency and the quality of the extraction, as well as the physical-chemical and biological properties of the final product, to better tailor it for adequate applications. The most common *in vitro* tests to which collagen can be subjected are reported below.

- ***Quantification***
  - *Hydroxyproline Assay*: hydroxyproline is a collagen-specific amino acid; therefore, hydroxyproline content can be considered a proxy for assessing the presence and quantifying

the collagen content into a sample. However, a precise quantification is possible only knowing the typical proline hydroxylation level of the considered collagen source. Reddy and colleagues developed an effective method for quantifying the Hyp content into a collagen sample: collagen is first hydrolysed, releasing hydroxyproline in the solution; free Hyp can then react with a chloramine-T and *p*-Dimethyl Amino Benzaldehyde (PDAB) to form a colored compound. The intensity of the colour is directly proportional to the hydroxyproline concentration, allowing for quantification through spectrophotometric measurement at 550 nm [Reddy and Enwemeka, 1996].

- *Sirius Red-based Methods*: the Sirius Red (or Picrosirius Red) staining is able to selectively bind collagen fibers [Constantine and Mowry, 1968]; this principle is at the basis of several collagen quantification assays, like the Sircol Assay (Biocolor Ltd, Carrickfergus, County Antrim, UK), in which the samples are stained with the Sirius Red dye, forming a stable complex. The bound dye is then solubilized, and the absorbance of the resulting solution is measured spectrophotometrically at a 540 nm wavelength.

- ***Structural information***
  - *Mass spectrometry*: mass spectrometry (or mass spectroscopy) follows the principle by which ions are deflected by magnetic or electric fields according to their mass-to-charge ratio, allowing for the identification and characterization of compounds based on their mass *spectra*. Mass spectroscopy can be therefore employed to provide molecular mass information about collagen proteins (with reported molecular masses of 283-300 kDa), complete of data about the dimensions of telopeptides in the mature protein, post-translational modifications, cross-linking patterns and possible contaminants [Abraham et al., 2008].
  - *Circular dichroism*: the circular dichroism (CD) technique utilizes the differential absorption of circular polarized light in an asymmetrical environment to assess the structure of the analyzed sample. In the case of collagen, this method allows not only to identify the secondary structure elements present within the protein, discerning features such as alpha-helices, beta-sheets, or random coils, but also to determine the relative proportions of these structural components [Bhatnagar and Gough, 1996].
  - *X-ray Photoelectron Diffraction (XPD)*: this method analyses the diffraction of photoelectrons emitted from a surface, thus providing information about the crystal structure and atomic arrangement of the surface layers; it consent indeed to evaluate the collagen fibril structure and organization through the analysis of X-ray scattering patterns [Abraham et al., 2008].
  - *Fourier Transform Infrared Spectroscopy (FTIR)*: FTIR can provide insights into the types of chemical bonds present within the collagen molecule, such as amide bonds (which are characteristic of proteins like collagen), and can also identify functional groups like amine, carbonyl, and hydroxyl groups, aiding in the characterization of collagen's secondary structure [Belbachir et al., 2009].
  - *Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)*: SDS-PAGE is one of the easiest and most common techniques used for obtaining information regarding the molecular weight and the integrity of a protein sample [Abraham et al., 2008]. Analysing collagen with this technique consent to determine the molecular weight of the protein and to assess if a collagen is homotrimeric or heterotrimeric, separating eventual different alpha-chains [Illidge et al., 2001].

- ***Chemical information***
  - *X-ray Photoelectron Spectroscopy (XPS)*: this technique is exploited to characterize the atoms present on the first 10 nm of a material's surface, consenting to assess the superficial atomic composition, hence identifying possible contaminants or modifications which may affect collagen's properties [Abraham et al., 2008].
  - *Contact angle*: using a contact angle goniometer, the angle of contact of a small drop (sessile drop method) of fluid placed on a surface of interest can be measured. Contact angle analysis on collagen surfaces provides information about their wettability and surface energy. This data is crucial when characterizing a collagen-based biomaterial because the hydrophobicity of a surface has a direct effect on cell adhesion and spreading; the contact angle method can therefore help in predicting and explain cell attachment on a collagen-based biomaterial. Furthermore, it can be used to check the efficacy of surface chemical modification reactions which should manifest as alterations in surface hydrophobicity or hydrophilicity [Abraham et al., 2008].
- ***Morphological information – imaging***
  - *Atomic Force Microscopy (AFM)*: AFM can reach a nanometre-scale resolution, thus providing data related to surface receptors and protein interactions. This technique can be used to characterize the surface roughness of collagen-based biomaterials [Abraham et al., 2008].
  - *Scanning Electron Microscopy (SEM)*: SEM analysis provides high-resolution images of collagen morphology and structure at the microscale level; is therefore often employed to visualize cells growth on a collagen surface, in terms of adhesion and spreading [Holmes et al., 2001; Abraham et al., 2008].
  - *Environmental Scanning Electron Microscopy (ESEM)*: through ESEM it is possible to visualize both native and denatured collagen and it can be employed, for instance, to characterize collagen coated constructs or self-assembled fibrils of collagen composites for bone-tissue engineering [Holmes et al., 2001; Abraham et al., 2008].
  - Transmission Electron Microscopy (TEM): TEM is a microscopy technique that utilizes a beam of electrons transmitted through a thin specimen to produce high-resolution images at a nanoscale level; it can offer detailed visualization of collagen fibrils ultrastructure and organization at nanoscale resolution [Holmes et al., 2001; Abraham et al., 2008].
  - *Immunohistochemistry*: immunohistochemistry is a common technique for sample visualization. It foresees the use of antibodies to target and bind the substance of interest present in a tissue sample; these antibodies are typically conjugated with a detectable marker, such as a fluorescent dye or an enzyme. When the antibody binds to its target collagen protein, the marker allows visualization of the substance distribution within the tissue sample. Analysing collagen sample with histochemical methods allows to identify and localize collagen fibers in tissues, to define their abundance, organization, and modification under various conditions. Moreover, immunohistochemistry can provide spatial information about the distribution and expression levels of collagen proteins within tissue samples [Kabiraj et al., 2015; Bielajew et al., 2020].
- ***Mechanical properties***

- *Tensile testing*: tensile testing for collagen involves subjecting a collagen sample to controlled stretching until it breaks, measuring the force applied and the corresponding elongation. This test provides information on the mechanical properties of collagen, including its tensile strength, elasticity, stiffness, and ultimate strain. It is therefore used in assessing the structural integrity and quality of collagen-based materials [Miyazaki and Hayashi, 1999; Quigley e al., 2018].
- *Rheological testing*: rheological testing for collagen regards assessing its flow and deformation behaviour under applied stress in different conditions, such as temperature, pH, or concentration, measuring its viscosity, elasticity, shear stress, and shear rate. This technique provides crucial information about collagen's structural organization, stability, and suitability for various biotechnological applications [Zhang et al., 2010; Skočilas, et al., 2016].

- ***Degradation and stability***
  - *Enzymatic degradation*: it's common to test collagen samples, such as 2D-membranes, by exposing them to specific enzymes that naturally break down collagen, such as collagenases or matrix metalloproteinases (MMPs). This is useful to assesses the resistance of collagen samples of interest towards enzymatic breakdown [Helling et al., 2017].
  - *Differential Scanning Calorimetry (DSC)*: in differential scanning calorimetry, a sample is subjected to a progressive heating until reaching a pre-set temperature, together with a reference sample. The difference in electrical energy required to raise the temperature of the sample and the temperature of the reference solvent is normalized by the heating rate to calculate the difference in heat capacity. In other words, the heat flow associated with changes in the sample structure is measured as a function of temperature. In the case of collagen, this technique provides information about collagen's thermal properties, including denaturation temperature, enthalpy of denaturation, and thermal stability [Fessas et al., 2006; Abraham et al., 2008].
  - *Thermogravimetric Analysis (TGA)*: this method is often performed in association with differential scanning calorimetry; in fact, similarly to DSC, here samples are subjected to controlled heating while measuring changes in their mass as a function of temperature. This technique provides information about collagen's thermal stability, decomposition temperature, and degradation kinetics [Fessas et al., 2006].

- **Cross-linking Analysis**:
  - *Cross-linking Density Assay*: it is possible to measure the degree of cross-linking density within collagen molecules, generally using techniques like the quantification of di-hydroxylysinonorleucine (DHLNL), which is a crosslinking amino acid derived from collagen. The method foresees collagen hydrolysis, from which DHLNL is released, and then quantifying it, typically using techniques like liquid chromatography with mass spectrometry. DHLNL levels reflect collagen's cross-linking degree, impacting its mechanical properties and biological functions [Paschalis et al., 2001].
  - *Pentosidine Assay*: this assay is used to quantify advanced glycation end products (AGEs) such as pentosidine, which are markers of collagen cross-linking and aging. In particular, pentosidine assay involves quantifying the levels of pentosidine, a fluorescent crosslinking compound formed by the reaction of collagen with reducing sugars. This assay provides information about collagen's degree of glycation and age-related changes and can serve as a marker for tissue aging and diabetes-related complications. [Sivan et al., 2006].

- ***Biocompatibility***
  - *Cell adhesion and proliferation*: when testing a biomaterial intended for biomedical purposes, it is crucial that the biomaterial in analysis is biocompatible. Two of the most common methods used to evaluate in vitro the biocompatibility of a sample are cell adhesion assay and cell proliferation assay: the former evaluates the ability of collagen to promote cell attachment and spreading, while the latter measures cells' ability to proliferate using the tested sample as a support for their own growth [see for instance Pozzolini et al., 2018].

Chapter II – THE MARINE DEMOSPONGE *Chondrosia reniformis* (Nardo, 1847)

« *The peculiar being on which I address your courteous attention, it is among those over which nature seems to have wanted to spread its thickest veil, to further test the curiosity of those who most yearn to penetrate the depths of its secrets. It is not rare in our Adriatic Sea, and it is known by the fishermen by the name of "sea kidney" […]* »

"Anatomic observations on the marine animal commonly known as Sea Kidney", G. Nardo, 1847

Chondrosia reniformis (Chondrosiida, Chondrosiidae) is a marine demosponge known to sailors and fishermen since ancient times. One of its earliest official references appeared in the 16th century in Pietro Andrea Gregorio Mattioli's treatise *"Commentarii in sex libros Pedacii Dioscoridis Anazarbei de Medica materia*"; subsequently, Carl Linnaeus described it as *Spongia reniformis* in his "*Systema Naturae*" (1758). In 1847, Giandomenico Nardo assigned it its current scientific name.

***Figure 6**: Chondrosia reniformis.*

2.1 Ecology

2.1.1 Habitat

Chondrosia reniformis is a common species throughout the entire Mediterranean area, although its presence has also been documented in several locations in the Eastern Central Atlantic [World Porifera Database - Species - Chondrosia reniformis Nardo, 1847 (marinespecies.org), last accessed January 23, 2024]; in the Western Atlantic, other species from the same genus could be erroneously identified and recorded as *C. reniformis*, but these observations are not accurate, as genetic differences between Western Atlantic and Eastern Atlantic/Mediterranean populations have been conclusively demonstrated [Lazoski *et al.*, 2001]. This sponge can typically be found from shallow waters to relatively mesophotic depths (~100 m) [Orel *et al.*, 2021] and is commonly observed on rocky substrates, often within coralligenous habitats. Occasionally, although infrequently, it can also be

found on sandy bottoms. It tends to prefer less illuminated sites, such as dark zones beneath rock overhangs, within cracks, or at the entrances of underwater caves [Göthel, H. 1992]; its growth is, in fact, inhibited by light [Wilkinson and Vacelet, 1979].

2.1.2 Reproduction strategies

C. reniformis is gonochoric and oviparous [Liaci and Sciscioli, 1975], but it can reproduce asexually too [Bavestrello *et al*., 1995; Bavestrello *et al*., 1998], and the occurrence of one or another strategy depends on environmental conditions. In sexual reproduction, gamete dispersal is expected to be quite limited, as they are unlikely to travel far from the parent sponges, and the spermatozoa are thought to remain suspended in the water column for a maximum of only a few hours [Lévi and Lévi, 1976]; the zygote undergoes development into a lecithotrophic free-swimming larva known as a *parenchymella* before eventually settling on a substrate, where it matures into a juvenile sponge. It has been observed that sexual reproduction is characterized by a maximum number of oocytes during summer [Riesgo and Maldonado, 2008; Di Camillo *et al*., 2011]. Asexual reproduction occurs primarily during late spring and summer via "drop-like" propagules: this sponge can generate very long outgrowths which extend downwards from the parental body, favoured by the force of gravity [Fassini *et al*., 2012], for up to 3 meters, then detach and form a new sponge.

2.1.3 Microbiome

Marine sponges host within their body exceptionally diverse microbial communities, such that is common to refer to them as "holobionts". This term encompasses not only the host sponge but also its associated microbiota, emphasizing that both are interdependent entities, rather than autonomous organisms. The microbial companions play a vital role in supporting the well-being of the sponge by contributing to its nutrition, defence mechanisms, immunity, and overall development [Pita *et al*., 2018]. According to the abundance and density of microbes in their tissues, sponges can be classified into two groups: high microbial abundance (HMA) sponges, which host significantly greater microbe densities compared to their surrounding environment, and low microbial abundance (LMA) sponges, in which microbe density does not significantly differ from that of the surrounding water [Pita *et al*., 2018]. *C. reniformis* is classified as a HMA sponge [Erwin *et al*., 2015], and its microbial heritage is transmitted vertically from the parent sponge in sexual reproduction [Lévi and Lévi, 1976]. The composition of the prokaryotic community of *C. reniformis* was investigated by [Roveta *et al*., 2023]: this sponge is associated with fifteen prokaryotic *phyla*, thirteen of them belonging to the domain Bacteria and two to the Archaea. Notably, the composition of these prokaryotic communities remains consistent throughout various regions of the sponge's body and is co-dominated by three lineages of ammonium-oxidizing organisms (*Cenarchaeum symbiosum*, *Nitrosopumilus maritimus*, and *Nitrosococcus* sp.), suggesting that ammonium oxidation/nitrification processes probably play a pivotal role in certain metabolic pathways of the sponge [Roveta *et al*., 2023].

2.2 Morphology

C. reniformis occurs in solid masses, sometimes solitary and sometimes aggregated, with a shape that can vary significantly, including globose, elongated, lobed, or flattened forms, often conforming to the substrate to which the sponge is attached [Nardo, 1847]. Its coloration ranges among different specimens from more or less dark marbled shades of brown or grey, somewhat resembling a leopard print. The body surface is smooth and glossy, interrupted only by a few round *oscula* of moderate

size, typically measuring around less than 1 mm to 1 cm in diameter, as observed by the author. However, despite the significant intraspecific variability, on average the specimens occupy a substrate area ranging from 10 to 100 cm², and generally exhibit a thickness of 2-4 cm [Garrone *et al*., 1975].

Cutting a transversal section of tissue, it is possible to notice two clearly distinct zones: an external dark-coloured *cortex*, which is the ectosome, and an internal yellow-whitish *medulla*, which constitutes the choanosome of the sponge (that is, the region in which are concentrated the choanocytes responsible for generating water currents and capturing food and in which is hosted the abundant population of symbiotic bacteria of this animal [Garrone *et al*., 1975]). Apart from choanocytes, the *cortex* encompasses all the different cell types present in the sponge, including lophocytes (that is, specialized cells contributing to the production of molecules essential for structural support of the sponge). The morphology of the filtration system of *C. reniformis* has been investigated by [Bavestrello *et al*., 1988]: the cortical layer contains intricate incurrent canals resembling tree branches, with smaller branches connecting to clusters of inhalant pores (the *ostia*). These canal stems penetrate into the sponge's choanosome, where they branch out and interconnect to form a complex network. Larger cloacal ducts originating from the oscular openings give rise to the main excurrent canals, which intersect with the incurrent network. Both incurrent and excurrent canals contain sharp, dead-end capillaries. Morphometric data analysis reveals that canal branch diameter scaling follows a linear pattern.

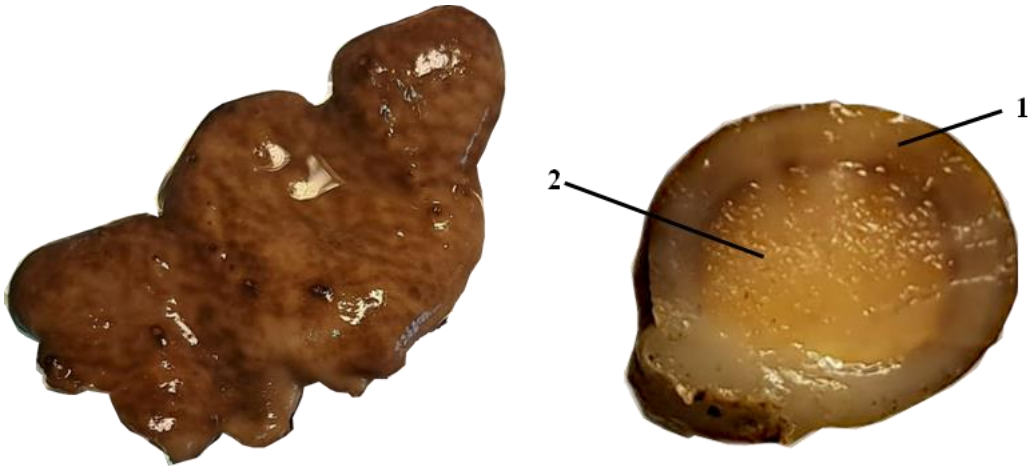


***Figure 7**: a whole specimen of Chondrosia reniformis on the left and a close-up of a transversal section of tissue on the right, where the ectosome (1) and the choanosome (2) are clearly visible.*

2.2.1 The collagen of *Chondrosia reniformis*

The structural integrity of sponges can be maintained through a combination of inorganic elements known as spicules, which can be either siliceous or made of calcium carbonate, and organic elements especially of proteinaceous – fibrous nature (see *1.1.5.1*). However, like only a few other *genera* of demosponges (notably, *Oscarella* and *Halisarca*), *C. reniformis* is devoid both of spicules and of reinforcing spongin fibers (where, for "spongin", the author means those peculiar molecules of collagenous nature typical of many horny sponges, such as those from the *genera Ircinia*, *Sarcotragus* or *Spongia*; see again *1.1.5.1*) [Garrone *et al*., 1975; Fassini *et al*., 2014]; instead, the greatest fraction of its body mass is constituted of an intercellular, fibrillar collagen which was found to be very similar to type I collagen from calf skin [Heinemann *et al*., 2007]. The presence of another type of collagen, belonging to the short-chain spongin-like subfamily and similar to type IV collagen, was also

identified in this species, where could be involved in the formation of the ectosome, in the process of attachment to the substrata and in the selective sediment incorporation typical of these species. [Pozzolini *et al*., 2012]; however, the author will default to using the term "collagen" to refer to the fibrillar type, unless otherwise specified, as it is way more abundant than the non-fibrillar one. A collagenous body allows this sponge to display great plasticity in terms of both shape and mechanical properties: for instance, after only a few days after an excision, fragments of *C. reniformis* experience a rounding off of cut surfaces and noticeable bending [Nickel and Brümmer, 2003]. The collagen of *C. reniformis* can be extracted from the sponge following different methods, with slightly different outcomes, as described in [Pozzolini *et al*., 2018a], and it appears as a viscous brownish suspension, in which the colour is probably due to melanin impurities [author's observations]. Indeed, the whole extracellular matrix of this species has been – and still is – subjected to many studies regarding its chemical-physical characterization.

2.2.1.1 Organization and distribution of the collagen fibres

Although omnipresent in the body of *Chondrosia reniformis*, collagen is organized differently between the two major zones recognizable in this sponge (see 2.2): the collagenous framework within the choanosome exhibits a notably less compact structure compared to the ectosome; the latter, instead, seems to be primarily comprised of intricate interwoven fibrillar bundles, where collagen fibrils are densely packed together. This arrangement consists of fibrils with a great consistency in their diameter, averaging around 200 Å [Garrone *et al*., 1975]. Within these bundles, fibrils are also interconnected by tinier filaments that play a role in linking individual fibrils, assisting in the creation of more complex structures. This suggests the presence of robust intrafibrillar cross-links, similar to the covalent cross-links stabilizing fibrillar aggregates in Vertebrates; nevertheless, the specific chemistry of these cross-links in this sponge collagen remains unknown [Heinemann *et al*., 2007].

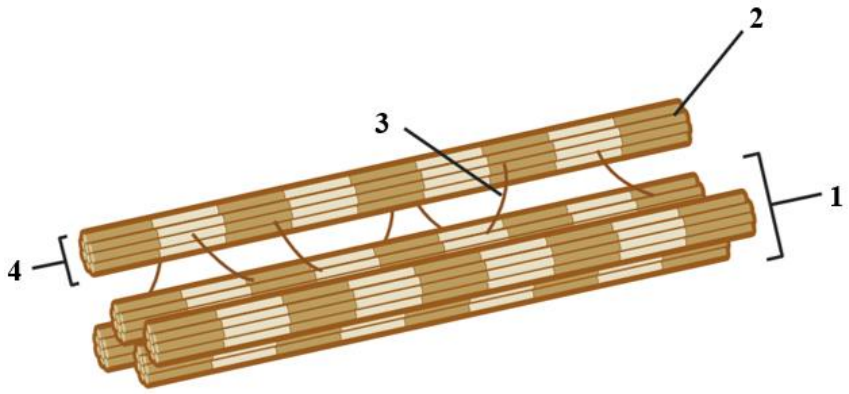


Figure 8*: Schematic representation of the bundle of fibers (***1***) observed in the collagenous matrix of* C. reniformis*, where (***2***) is a single fibril and (***3***) are some of the filaments which interlink the fibrils; the diameter of each single fiber (***4***) is about 200 Å.*

2.2.1.2 Composition and physical-chemical properties

The collagenous extract derived from the extracellular matrix (ECM) of *Chondrosia reniformis* exhibits distinctive biochemical behaviour compared to more common collagens found in fish, jellyfish, or mammals. Notably, it demonstrates high insolubility in acidic environments [Heinemann et al., 2007], making it resistant to extraction using traditional methods (refer to section 1.4.1). Instead, an alternative methodology is required for its extraction [Pozzolini et al., 2018]. A similar

demeanour of insolubility in acidic environment is observed in both aged type I human collagen and in the collagen of marine glass sponges [Heinemann et al., 2007]. When a droplet of *C. reniformis*' collagen is introduced into an acidic solution, it constricts, forming a rounded, gummy structure. This phenomenon is reversible, as long as the acidic solution is not overly concentrated: when the "solidified" droplet is reintroduced into a solution with a higher pH, it disperses, returning to its original state of suspension. Another peculiarity of this collagen is its resistance to collagenases; as reported by several authors, however, this seems to be a common feature in many sponge collagens [Garrone et al., 1975]. The reason behind these features has to be ascribed to the biochemical characteristics of the collagen, in particular the aminoacidic sequence and the ECM molecules which happen to be more or less intimately associated with the collagen fibers. Various authors have conducted amino acid composition analyses of *C. reniformis* collagen [Garrone et al., 1975; Swatschek et al., 2002; Tassara et al., 2023]. Despite the challenge of comparing results due to the different methods used, a consistent observation across the cited works is the higher presence of phenylalanine and aspartic acid compared to a control of bovine type I collagen. Nevertheless, the overall amino acid profile of this collagen appears similar to that of bovine type I collagen, showing a clear correlation also in their infrared spectra [Heinemann et al., 2007] and, even though less evidently, in their mechanical properties [Garrone et al., 1975]. As explained in the previous paragraph, the collagen bundles in the *C. reniformis*' cortex show great cohesiveness: this suggests the presence of a ground substance that effectively helps the bundles to bind together, behaving like a sort of "cement", and that is represented by sugary molecules (such as proteoglycans, glycosaminoglycans, and more or less complex carbohydrates) distributed all along the fibrils [Garrone et al., 1975] or dispersed in the extracellular matrix. In fact, sponges display one of the most glycosylated collagen of the entire animal kingdom [Garrone et al., 1975]. For what concerns *Chondrosia reniformis*, very little is known about the exact nature and composition of the sugars associated to its collagen fibers; however, seen the great thermal stability of this collagen [Tassara et al., 2023], which is higher than those of fish and of calf skin [Garrone et al., 1975], it is probable that they play a role in this context, as this feature seem not to be ascribed to high levels of hydroxyproline [Tassara et al., 2023]. Finally, it has to be highlighted that the collagen of *C. reniformis* is very plastic: apart from displaying considerable intraspecific variability (author's observations), it can exhibit biochemical variations considering the different environmental conditions in which the animal is able to live, being this species eurybate and eurythermal [Orel et al., 2021; Tassara et al., 2023]; consequently, the chemical-physical properties of this extract can undergo changes. More specifically, higher levels of collagen glycosylation, and hence increased thermal stability, are observed during the summer and in shallow sponge populations [Orel et al., 2021; Tassara et al., 2023].

2.2.1.3 Contraction ability

Despite lack of a nervous system, sponges exhibit various behavioural responses, such as the contraction of the entire body or specific parts, along with changes in contraction rhythm and amplitude. These responses can be influenced by the application of externally administered neuroactive compounds [Nickel, 2010]; moreover, larvae and adults of various species both demonstrate evident coordination phenomena in response to external and internal *stimuli* [Fassini *et al*., 2014]. In the case of *C. reniformis*, it has been observed that the body of the animal exhibits a noticeable stiffening when handled, *i. e.* in response to mechanical stimulation [Wilkie *et al*., 2006; Fassini *et al*., 2014]. A similar situation also occurs in other *phyla*, for instance in the so-called mutable collagenous tissue (MCT) of echinoderms [Wilkie, 2005], which is another example of anatomical structure with a conspicuous collagenous composition, exhibiting rapid adjustments in its passive mechanical properties. Sea cucumbers can rapidly alter the mechanical properties of their

ECM under nervous control through the releasing of a stiffening factor, normally sequestered inside cells, which is able to reinforce interfibrillar cohesion. In *C. reniformis* there is evidence for the presence of potentially contractile elements, as its mesohyl contains a sparse and loose network of cell processes, probably belonging to endopinacocytes and equipped with actin microfilaments [Wilkie *et al*., 2006]. However, morphological studies have revealed that the quantity and the arrangement of these potentially contractile cells are not sufficient to influence significantly the passive mechanical properties of the internal tissue. Therefore, [Wilkie *et al*., 2006] proposed that this mechanical behaviour of the *C. reniformis* mesohyl could be regulated physiologically. The cited authors hypothesized and demonstrated that cells within the mesohyl store a stiffening factor (or more than just one) and that the observed stiffening in intact animals upon being touched is attributed to the release of this factor through direct interaction with the extracellular matrix, enhancing the cohesive forces and therefore prevent slippage between fibrils or fibers. This phenomenon is Ca^{2+}-dependent and is irreversible if cells are permanently damaged, leading to the conclusion that both stiffening and destiffening of the mesohyl (or at least one of them) seem to be an active, cell-mediated process. Subsequently, [Fassini *et al*., 2014] deepened the evidence of the existence of a chemical factor that influences the interaction among collagen fibrils, consequently altering the mechanical properties of the mesohyl. In particular, these authors demonstrated not only a notable correlation between the ectosome stiffness and the extent of mechanical stimulation, but also that the manipulation of entire sponge led to a contraction along all three axes, with the most significant dimensional changes occurring along the axis parallel the line of action of the applied pressure. Moreover, the stiffening process is accompanied by shrinkage of both canals and mesohyl and it extends to the entire region close to the stimulated area: this suggested the presence of a functional conduction system. The biological significance of the stiffening phenomenon remains uncertain; however, it has been hypothesized that the physiological regulation of mesohyl stiffness might be connected, at least in part, to the control of the release of reproductive propagules. [Fassini *et al*., 2014].

2.2.1.4 The "creeping" phenomenon

It has already been highlighted in this manuscript that *Chondrosia reniformis* displays remarkable tissue plasticity and body deformability. As briefly mentioned in before, *C. reniformis* can undergo a unique process called "creeping", forming long and slender outgrowths from the parent body which may extend up to 3 meters in length (see Fig. 8); this is noteworthy, especially considering that an average adult sponge of this species typically measures around 15-20 cm in diameter [Fassini *et al*., 2012]. Usually, these outgrowths extend downwards, as if under the force of gravity, but they can be observed also developing horizontally [Bonasoro *et al*., 2001]. Similar behaviour is observed in other few sponges, including *Oscarella lobularis* and *Chondrilla nucula*; these species, like *C. reniformis*, lack an organized endoskeleton of spicules or spongin fibers and are primarily found inhabiting exposed cliffs. [Bonasoro *et al*., 2001].

The phenomenon of "creeping", like the stiffening ability widely described in the previous paragraph, has been attributed to an alteration of the interactions between the collagen fibrils of the mesohyl [Fassini *et al*., 2014], and it has been studied both from a morphological and a physiological perspective. During the dynamic creep condition, it is possible to identify four different anatomical regions [Bonasoro *et al*., 2001] (see Fig. 8): 1) the *parent region*, that represents the part of the sponge body from which the outgrow develops, 2) the *elongation region*, that is the part that stretches, 3) the *transition region*, which represents the anticipated rupture site, where the weakened filament breaks, leading to the division of the sponge into two parts, and 4) the *propagule region*, considered as the "new" sponge body.

Figure 9*: On the left,* Chondrosia reniformis *in its habitat during active creeping phenomenon (photography credits: Jordi Regas); on the right, schematic representation of the different anatomical parts individuated during creeping, where (**1**) is the parent region, (**2**) is the elongation region, (**3**) is the transition region and (**4**) is the propagule region.*

From a microscopic anatomical point of view, an exhaustive description of the above-mentioned regions observable during the creeping behaviour of *C. reniformis* have been provided by [Bonasoro *et al*., 2001]. The parent region appears similar to the sponge during the standard resting condition, with the same texture and microarchitecture of the collagen fibres; however, a few ultrastructural features appear different during this functional state, in particular novel and highly prevalent intercellular connections between the elongated processes of endopinacocytes and lophocytes and the spherulous cells. In these areas, the cell membranes appear closely associated, forming structures resembling junctions. On the contrary, significant differences in the ultrastructural organisation of both the cellular and the extracellular components of the mesohyl can be observed in the elongation region: while the exopinacoderm appeared still to form a more or less continuous layer, with only a few gaps and probable ruptures, the overall development of the aquiferous system was clearly reduced (even if not completely absent). Associations between spherulous cells (which content looks much more heterogeneous than in the parent region) and endopinacocytes or lophocytes are evident. In the extracellular stroma of the mesohyl, collagen bundles are loosely arranged, often showing disarray and disaggregation. The orientation of collagen fibers is different from the one observed in the parent region, forming a longitudinally stretched meshwork of parallel bundles. However, the size and structure of collagen fibrils appear unchanged. Wide areas between collagen bundles are irregularly filled with a loose network of microfilaments. Notably, the production and the rearrangement of ECM components not only continues but became even more evident during this stage. The transition region further differs from the previous two: cell types, subjected to different degrees of vacuolisation or even degeneration, are difficult to tell apart; more specifically, the outer layer of cells constituting the exopinacoderm is almost completely unstructured, therefore causing a massive invasion and spreading of bacteria all throughout the mesohyl. A general disorganization can be seen also in the ECM component, where the bundles of fibers are disorganized and scattered randomly. Finally, the propagule region appears again very similar to the intact parent sponge.

The creeping phenomenon can be observed not only in nature, but it can be also artificially induced by subjecting the sponge, or even fragments of it, to prolonged compression or tension, for instance applying weights to sponge tissue samples [Fassini *et al*., 2012]. Moreover, it sometimes occurs even horizontally, without the detachment from the substrate, especially if the sponge had fronted a prolonged state of stress and/or starvation (personal observation). In any case, given that this dynamic condition can happen without rupture or lengthening of the collagen fibrils [Bonasoro *et al*., 2001], it is assumed that the collagen fibrils making up the ECM of *C. reniformis* are not continuous, implying that individual fibrils do not entirely encircle the sponge's body. Additionally, the molecular interactions binding adjacent fibrils together must be unstable enough to allow fibrils or bundles of fibrils to slide past each other.

The ecological significance of this phenomenon is still not completely clarified. Different authors gave different interpretation of it, not necessarily mutually exclusive, but rather coexisting. In most cases, creeping is correlated with reproductive processes and can be seen as an opportunistic form of asexual reproduction [Fassini *et al*., 2012]. This is supported by the fact that this kind of event is apparently seasonal, with a strong direct correlation with sea water temperature and therefore showing its maximum peaks during the warmer months of the year (from May to September) [Fassini *et al*., 2012]. However, creeping can also be viewed as a passive reaction to mechanical pressures resulting from alterations in the environment or rearrangements of the substrate, thereby enhancing the survival rate of the sponge in instances where a portion of the substrate collapses. Finally, it has been suggested that these dynamic creep phenomena may represent an unconventional type of localized movement, again potentially occurring before asexual reproduction.

2.3 A unique skill: selective engulfment and dissolution of quartz particles

Many marine organisms can actively interact with the substratum for various ecological reasons, usually regarding the maintenance of body support or the creation of a niche to live in. Among sponges, one of the most remarkable examples is the Clionidae family, which are able to bore calcareous surfaces, therefore playing a crucial role as bioeroders [Dieudonne and Carroll, 2022]. The erosion of siliceous substrates, on the other hand, is highly uncommon. Among terrestrial organisms, only a few lichens are recognized for penetrating siliceous rocks, and this capability is basically unobserved in the animal kingdom [Bavestrello *et al*., 1995]. *Chondrosia reniformis*, however, is an exception to this rule. As other demosponges, it incorporates a wide range of external material in its ectosome (such as sand grains or foreign spicules), specifically to improve body support, lacking its own skeleton. The peculiarity lies in the distinct reactivity of this species' exopinacoderm towards crystalline and amorphous silica: on one hand, the engulfing of amorphous silicon (such the one found in opaline spicules) elicit a motile response in exopinacocytes, that end up forming a cellular coat covering the opaline material and leavening it unaltered; on the other hand, crystalline silicon particles (i. e., quartz particles), are gradually incorporated by exopinacocytes and actively etched till made uniform in size [Bavestrello *et al*., 1995; Bavestrello *et al*., 1998b]. Additionally, it has been observed the presence of a polarity in the mineral selectivity of the sponge: the upper side of the ectosome collects quartz and amorphous silicates, in a dynamic process comprising incorporation of the particles and resizing of the quartz grains, with their elimination via the aquiferous system [Bavestrello *et al*., 1995]. The lower side, instead, specifically engulfs the calcareous substrata, thus facilitating the anchorage of the sponge to the bottom [Bavestrello *et al*., 1998c]. Giving the presence of a thick collagenous ectosome, and considering that ascorbic acid is the reducing agent in proline

hydroxylation (a key process in collagen biosynthesis, see 1.1.2), the latter has been identified as the molecular mediator of this process [Bavestrello *et al*., 1995]. Ascorbic acid, in fact, can react with quartz, and not with amorphous silica [Fenoglio *et al*., 2000]. [Pozzolini *et al*., 2012] demonstrated the direct role of dissolved silica in the positive regulation of a collagen gene in *C. reniformis*; these results have led to the hypothesis that quartz erosion through ascorbic acid locally releases soluble silicate, thereby promoting collagen gene expression [Giovine *et al*., 2013]. The biological importance of this behavior appears to be further supported by the presence, particularly in the ectosome, of aquaporin molecules that may have the potential function of facilitating the permeation of H_2O_2 released during quartz erosion by ascorbic acid [Pozzolini *et al*., 2016].

2.4 Biotechnological appeal

Due to its distinctive features, *Chondrosia reniformis* has been the subject of extensive studies in recent years, simultaneously gaining attention for its biotechnological potential. So far, both the entire alive organism and its collagen have been explored for possible usages, especially in the fields of bioremediation and biomedicine. As a whole organism, *C. reniformis* can be employed as bioindicator for heavy metal pollution: sponges, in fact, filter notable volumes of water during their feeding process, capturing food particles and various compounds, and therefore frequently accumulating within their mesohyl pollutants found in the water column, such as hydrocarbons, organochlorinated compounds and heavy metals. In particular, *C. reniformis* has been proven to bioaccumulate in its choanosome trace elements like Fe, As, Cu, Cr and Hg, showing all the fundamental features requested for being designed as an ideal biomonitor [Roveta *et al*., 2022]. This species is also known for the production of biomolecules of high biotechnological interest, both bioactive compounds and biomaterials. For instance, when it experiences significant stress or undergoes injury, *C. reniformis* secretes into the surrounding environment a cytotoxic protein named chondrosin; [Scarfì *et al*., 2020] investigated the anti-tumoral activity of this protein, which toxic effect is mediated by an extracellular Ca^{2+} intake followed by a ROS overproduction in the cytoplasm, and proved a high homology with the best known defensins, present in other animal groups like Arthropods and Reptiles. However, a greater interest is reserved for the collagen of this sponge, which has a wide list of demonstrated potential applications as a biomaterial from marine origin:

- ***Moisturiser in cosmetic preparations*** as a substitute for more conventional collagens. [Swatschek *et al*., 2002] added a collagenous extract of *C. reniformis* to two different cosmetic formulations and tested its effect on volunteers; compared with their extract-free analogues and with a commercially available bovine collagen-containing product, the sponge collagen formulations determined a slight increase of skin hydration and significant boost of lipids, without any influence on the physiological skin surface pH.

- ***Injectable hydrogel medium*** for biomedical applications. [Fassini *et al*., 2017] described a method of extraction to obtain from *C. reniformis* a sticky and highly hydrated collagenous hydrogel, enriched with other associated proteins and GAGs, which can be used collagen to create scaffolds for tissue engineering and regenerative medicine.

- ***Collagen-derived bioactive peptides***, [Pozzolini *et al*., 2018]. *C. reniformis*' collagen can be hydrolysed to obtain bioactive peptides able to stimulate cell growth, protect cells from UV-induced death and with significant antioxidant activity, suggesting a possible profitable employment in drugs or cosmetic formulations for damaged or photoaged skin repair.

- ***Production of 2D membranes*** for tissue engineering and regenerative medicine. This collagen can be poured into moulds and let dry, to obtain thin, biocompatible 2D membranes exploitable for instance for skin repair purposes [Pozzolini *et al*., 2018]. Notably, the chemical-physical features of these membranes show some differences based on the season (and therefore, on the temperature of the water) in which the sponge was collected before the collagen extraction: more specifically, membranes produced with collagen obtained from sponges harvested in summer result more thermally stable and richer in glycosylation than their winter counterparts, that on the contrary have a greater mechanical stability [Tassara *et al*., 2023].
- ***Marine gelatine***. [Silva *et al*., 2016] developed a simple and environmentally sustainable methodology to obtain a *C. reniformis*-derived collagen/gelatine mixture, suggesting the possibility to scale up this process to an industrial scale for the production of a marine biomaterial suitable for biomedical, cosmetic, and pharmaceutical applications.

2.5 Sustainable production of *Chondrosia reniformis*

Notwithstanding its great biotechnological potential, widely described in the previous paragraphs, the exploitation of *Chondrosia reniformis* faces a significant limitation in achieving sustainable harvesting. This sponge exhibits a relatively slow growth rate, necessitating large-scale production to meet market demands at least partially. This raises concerns about the potential risk of overexploitation, a scenario that has already occurred with other species of sponges rich in secondary metabolites [Tinghino, 2020]. To date, also given the good recovery capacity and regeneration of this species after cutting [Nickel and Brümmer, 2003], many attempts have been made with the aim of obtaining an environmentally conscious production of this animal, put in place especially through the *in vitro* culture of sponge fragments and through mariculture. For instance, [Nickel and Brümmer, 2003] chose *C. reniformis* to develop strategies for biotechnological upscaling of sponge cultures using fragments of about 50–80 mm^{-3}, that were cultured in 6-well plates with filtered supplemented sea water; the authors were able to maintain viable fragments for more than a year, but their metabolic state progressively lowered with time, with the consequent depletion of the aquiferous system and cell density. On the other hand, while the current yield rate is still insufficient to meet industrial demands, the mariculture efforts for this species appear highly promising. [Gökalp *et al*., 2022] investigated a cost-effective, easily implementable, and sustainable production method for cultivating *Chondrosia reniformis* based on the employment of "lantern-shaped" structures held vertically in the water column by buoys and anchors, in which the estimated minimum requirements for economically viable demosponge culture of 90% survival and growth of 100% per annum were satisfied, suggesting that this strategy could be included, for example, in a marine collagen production pipeline, possibly in synergy with feeding fish mariculture. By now, however, it has been demonstrated that *Chondrosia reniformis* can thrive almost equally well in tanks, as long as the latter have enough volume capacity, maintaining good health with minimal treatment requirements [Tinghino, 2020]. It must be taken in consideration that tanks have the advantage of being a controlled system in which environmental parameters can be managed, but some precautions are requested to maintain vital and healthy sponges. In particular, hydrodynamics is essential for *C. reniformis* viability, as the current facilitates the removal of stress factors such as bacterial biofilm, algae, sediments, and copepods, which can represent a health risk for the sponges; moreover, insufficient current levels inside the tanks could result in water stagnation, leading to bacterial proliferation and alterations in chemical parameters, such as nitrate levels (that for this species must be ≤ 25 mg/L) or tank pH (which value shouldn't go below 8). Salinity should be around 35 psu, while an optimal diet formulation consists in commercial

phytoplankton added with ascorbic acid and sediments [Tinghino, 2020]. Finally, a crucial factor is water temperature: the best results in terms of growth rate are in fact obtained at 25°C [Orel *et al*., 2021]. Currently, however, despite the encouraging results, additional and extended experiments are required to guarantee the sustainable production of biomass over the long term.

Section B – Research Activity

Chapter III – MOLECULAR CHARACTERIZATION OF THE FIBRILLAR COLLAGENS IN *Chondrosia reniformis* (Nardo, 1847)

In the previous chapters, the distinctive traits of the collagen of *Chondrosia reniformis* have been extensively discussed, together with its biotechnological potential in many fields. To date, however, an in-depth molecular description of this collagen, from the gene sequences to the deduced final proteins, is still missing; a previous attempt was made by [Pozzolini *et al.*, 2016], who reported a single partial collagen sequence extrapolated from a transcriptome analysis of this sponge. This part of the doctoral project aims to identify and describe all the fibrillar collagen gene sequences occurring in the genome of *C. reniformis*, their level of expression in the adult animal and their final products as mature proteins.

3.1 Materials and methods

3.1.1 Acquisition of Fibrillar Collagen Gene Sequences

To obtain the fibrillar collagen gene sequences of *C. reniformis*, a genomic analysis approach associated to a transcriptomic analysis approach were necessary. Simultaneously with the beginning of the present doctoral activity, the Wellcome Sanger Institute (Hinxton, Cambridgeshire, UK) set up the "*Moore Aquatic Symbiosis Genomics Project*", with the aim to generate high-quality gene sequences for organisms associated with microbial symbionts, and *C. reniformis* was included in the list of target organisms; hence, to optimize efforts and costs, the "wet lab" activity of this part of research was exclusively dedicated to the transcriptomic analysis. Except for sample collection and preparation, described in *3.1.1.1*, this part of activity was performed in the molecular biology laboratories of the National Museum of Natural Sciences of Madrid (Spain), as part of a traineeship lasting three months.

3.1.1.1 Sample collection and preparation

Several medium sized specimens of *Chondrosia reniformis* were harvested by scuba diving in the waters surrounding the Gallinara Island (Savona, Liguria, Italy), in the Ligurian Sea. The animals were kept in fresh seawater, collected in the same location of sampling, until their arrival in laboratory. There, each specimen was promptly cut into < 1 cm^3 pieces with a sterile scalpel and immersed into an adequate volume of RNA*later*®, and finally sent to the National Museum of Natural Sciences of Madrid, where they were further processed.

3.1.1.2 Total RNA extraction

Given the highly fibrous tissue of *C. reniformis*, which appears and feel like a piece of rubber, total RNA was extracted using the TRIzol® Reagent in association with the PureLink® RNA Mini Kit, as the conjunction between the two methods was found to be the best option in case of samples that are difficult to lyse. Firstly, the pieces of *C. reniformis* tissue were rinsed with sterile RNAse free water and furtherly cutted into tinier portions with a sterile scalpel, then homogenized directly into the TRIzol® Reagent (following a ratio of 50-100 mg tissue:1 mL reagent) and incubated at room temperature for five minutes, to allow complete dissociation of nucleoprotein complexes.

Subsequently, an amount of chloroform was added, equal to 1/5 to the total volume of starting reagent, shaking the tube by hand for 15 seconds and avoiding vortexing. After 3 minutes at room temperature, samples were centrifuged at 12000 × *g* for 15 minutes at 4°C. The transparent aqueous upper phase was transferred to a clean RNAse-free tube and an equal volume of 70% ethanol was added, then the extraction procedure was brought on with the PureLink® RNA Mini Kit, according to manufacturer's instructions. Quantity and quality of total RNA thus obtained were measured with a NanoDrop 2000 (Thermo Scientific).

3.1.1.3 Purification of mRNA and preparation of a cDNA library for sequencing

The mRNA purification and the cDNA library preparation were performed starting from 25-1000 ng of total RNA for each sample, using the Illumina Stranded mRNA Prep Kit (Illumina, Inc., San Diego, CA, USA), following manufacturer's guidelines and using Illumina adapters. The final concentration of the cDNA libraries was assessed with the Qubit dsDNA High Sensitivity (HS) Assay Kit using the Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA), while their quality was checked using the Agilent 2200 TapeStation system, with the High Sensitivity D1000 ScreenTape. Successful cDNA libraries (i.e., with a final concentration of >1 ng/µl and with an optimum quality profile) were finally sent to Novogene (UK) Company Limited for whole paired-end transcriptome sequencing.

3.1.1.4 Transcriptome assembly

The quality assessment of the raw reads provided by sequencing was performed with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, last accessed February 22, 2024). Considering that a *Phred* score equal to 30 corresponds to a 99.9% base accuracy and to a probability of 10^{-3} of incorrect base calling, the terminal bases resulted from the FastQC analysis with a *Phred* quality score <30 were discarded with Trimmomatic (http://www.usadellab.org/cms/index.php?page=trimmomatic, last accessed February 22, 2024), thus producing sequences of unequal size. Reads were then screened using FastQC again, checking for the presence of residual adapters or primers sequences; if present, adapters or primers were trimmed with Trimmomatic. *De novo* assembling with clean data sets was performed with Trinity (https://github.com/trinityrnaseq/trinityrnaseq, last accessed February 22, 2024). The final assembled transcriptome derives from the merging of the sequenced reads obtained from 27 specimens of *C. reniformis*. Subsequently, the quality of the assembly was checked: firstly, the read composition of the assembly was assessed with Bowtie2 (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml, last accessed February 22, 2024), mapping the original unassembled reads to the assembled transcripts and hence evaluating how well the assembled transcripts represented the original reads. Finally, the BUSCO tool (https://busco.ezlab.org/, last accessed February 22, 2024) was run for assessing the completeness of the assembly, together with a gVolante analysis (https://gvolante.riken.jp/, last accessed February 22, 2024).

3.1.1.5 Sequence annotation

To identify predicted protein-coding sequences from transcript sequences, TransDecoder (https://bio.tools/TransDecoder, last accessed February 22, 2024) was run, selecting a minimum protein size of 100 amino acids as default and retaining only the single best open reading frame (ORF) per transcript. The functional annotation of the translated transcriptome was performed with Diamond (https://bio.tools/diamond, last accessed February 22, 2024), using the latest release of the Swissprot database (https://www.uniprot.org/help/downloads, last accessed February 22, 2024) as a reference, using an e-value cutoff of $1e^{-5}$.

3.1.1.6 Research of fibrillar collagen sequences

Fibrillar collagen sequences were searched both in the annotated genome and the annotated transcriptome, obtained as described in the previous paragraphs. *C. reniformis*' genome (GCA_947172415.1) was submitted on NCBI (https://www.ncbi.nlm.nih.gov/, last accessed February 23, 2024) by the Wellcome Sanger Institute in October 2022. The gene prediction within the genome was provided by the Riesgo Lab and was obtained using Augustus (https://github.com/Gaius-Augustus/Augustus, last accessed February 23, 2024) and BLAST (https://www.ncbi.nlm.nih.gov/books/NBK279670/, last accessed February 23, 2024). The information derived from this analysis was collected into three text files:

- One .gtf file (GenePrediction.gtf), containing the list of the predicted genes, the chromosome in which are located, and the predicted starting and ending position of the open reading frame (ORF) of the genes within their chromosome in terms of number of nucleotides.
- One .txt file (BlastGeneAnnotation.txt), containing the annotation of the predicted genes.
- One .fasta file, containing the nucleotide sequences of the predicted open reading frames (ORF) of the annotated genes (NT.fasta).

The name assigned from the bioinformatic tools to the genes is maintained in all files.

In order to individuate fibrillar collagen sequences within the genome, the keyword "COLL" was used as a query to manually search for predicted collagen genes in the file BlastGeneAnnotation.txt; only fibrillar collagens-related genes were retained. The name assigned to fibrillar collagens-related predicted genes was then used as a query in the NT.fasta and AA.fasta files, to obtain their predicted CDS and protein sequence. Finally, knowing the predicted starting and ending position of the genes of interest thanks to the GenePrediction.gtf file, it was possible to obtain the whole gene sequences running the following command line in an Ubuntu terminal within a conda environment:

```
~$seqkit faidx Chromosome.fasta chrname:startingposition-endingposition
```

where "Chromosome.fasta" is the fasta file of the chromosome in which the gene of interest is located; "chrname" is the name of the chromosome indicated at the beginning of the Chromosome.fasta file); "startingposition-endingposition**"** is the chosen range containing the sequence of interest (it has to be indicated only with numbers). Considering that the GenePrediction.gtf file reports only the starting and the ending position of the gene's ORF, in order to obtain also the 5'-UTR and 3'-UTR regions the same command line was run but considering a default of 3000 bp before and after the starting and ending point of the ORF.

The transcriptome analysis provided information collected into two main text files:

- One .fasta file (AssembledTranscripts.fasta) containing the nucleotide sequences of the assembled transcripts; each transcript is identified by a specific name.
- One .txt file (AnnotatedTranscripts.txt) containing the predicted Diamond annotation relative to each transcript.

To obtain the fibrillar collagen-related transcript sequences, the keyword "COLL" was again used as a query to manually search for the transcripts of interest within the AnnotatedTranscripts.txt, retaining only fibrillar collagens-related transcripts; then, the name of the thus found transcripts was used as a query in the AssembledTranscripts.fasta file to obtain the corresponding nucleotide sequences of the transcripts.

3.1.2 Analysis of Fibrillar Collagen Gene Sequences

Considering that the majority of the bioinformatic tools designated for gene sequence analysis are calibrated on human, mouse, rat or other model organisms' genomes, most of the operations aimed at the analysis of the *C. reniformis*' sequences obtained as described in 3.1.1 were carried out manually. To obtain their translation products, the transcripts were translated into protein sequences using ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/, last accessed February 23, 2024) or the ExPASy translation tool (https://web.expasy.org/translate/, last accessed February 23, 2024). The intron/exon assessment was carried out using Augustus and LALIGN Pairwise Sequence Alignment (https://www.ebi.ac.uk/jdispatcher/psa/lalign, last accessed February 23, 2024), and then checked manually. The sequence structure of both gene and protein sequences was also analyzed manually, as well as the percentage of collagen key amino acids in each protein sequence. The prediction of potential N-glycosylation sites and O-glycosylation sites was performed with the standard predictor of GlycoEP (https://webs.iiitd.edu.in/raghava/glycoep/, last accessed March 1, 2024) based on the composition profile of patterns and setting 0.8 as a threshold value for the prediction job.

3.1.2.1 Research of potential regulatory sequences in the genes' upstream regions

In 2019, Pozzolini and co-workers demonstrated the involvement of a TGF-related mechanism in the growth and regeneration of *Chondrosia reniformis* [Pozzolini et al., 2019]. In particular, it has been individuated in a transcriptome of this animal the presence of seven TGF superfamily ligands, three TGF superfamily receptors and several Smad proteins, which sequences appeared to be extremely conserved among animal *phyla*. In the same work, the administration of a strong TGF-inhibitor altered the restoration of the exopinacoderm layer of the sponge, suggesting that this molecular pathway could be involved in the regulation of the expression of collagen genes in *Chondrosia reniformis*. For this reasons, the region preceding the coding sequence of the fibrillar collagen genes (notably, -724 bp for G1 and G2 and -2000 bp for G3, G4, and G5) was considered for the research of potential Smad-binding motifs, especially the 5'-CAGAC-3' (considered as the main binding element for Smad2/3/4) motif and GC-rich motifs (preferred by Smad1/5/8), which are the main recognized binding motifs for Smad transcription factors [Martin-Malpartida et al., 2017].

3.1.3 Gene expression profile of fibrillar collagen genes in *Chondrosia reniformis*

The expression profile of the obtained fibrillar collagen genes was checked within adult specimens of *C. reniformis*, comparing the relative expression of the afore-mentioned genes in the cortex and in the choanosome, respectively.

3.1.3.1 Sample collection

Three specimens of *C. reniformis* were harvested by scuba-diving in the Portofino MPA (Marine Protected Area) in shallow waters during spring of 2023. The animals were kept into refrigerated bags containing the same sea water of the collection site until their arrival in laboratory; then, they were carefully put into an aquarium with the same temperature and salinity parameters of their original environment and let acclimatize for 24 hours before processing.

3.1.3.2 Gene expression evaluation

For each specimen, two cylindrical sections (Ø =1 cm; height= about 2 cm) were excised, making sure to encompass both the cortex and the choanosome. The tissue samples thus obtained were then cut with a sterile scalpel in order to separate the cortex from the choanosome. Each portion was promptly chopped into tinier fragments and put into 3 mL of Isol-RNA Lysis Reagent (5 Prime GmbH, Hilden, Germany) for total RNA extraction. Then, the poly-A fraction was obtained with the

FastTrack® MAG mRNA isolation kit (Life Technologies, Milan, Italy) respecting the manufacturer's instructions. Complementary DNA (cDNA) was produced with the Revert Aid Reverse Transcriptase Kit (Thermo Fisher Scientific, Milan, Italy), using 1 μg of purified mRNA from each sample. The qPCR reactions were carried out in triplicate for each sample in a final volume of 15 μL containing: 1× iQ SYBR®Green master mix (Bio-Rad, Milan, Italy), 0.2 μM of each primer, and 3 μL of a 1:5 dilution of cDNA. The thermal profile involved the following steps: initial denaturation for 3 mins, followed by 45 cycles with denaturation at 95°C for 15 seconds, annealing and elongation at 57.7°C for 60 seconds. At the end of each elongation step, fluorescence was measured. The final step involved a slow heating (1°C/s) of the amplified product from 55°C to 92 °C, to generate a melting temperature curve. GAPDH was used as the reference gene and the values were normalized to its expression. The PCR primers (see Table) were designed using the OligoArchitect™ software (Sigma-Aldrich) and were provided by Eurofins Genomics (Milan, Italy). Data analyses were performed with the DNA Engine Opticon® 3 Real-Time Detection System Software program (3.03 version). In order to calculate the relative gene expression of G1, G2, G3, G4, and G5, the comparative threshold Ct method [Aarskog and Vedeler, 2000] was used within the Gene Expression Analysis for iCycler iQ Real Time Detection System software (Bio-Rad) [Vandesompele et al., 2002]. For the tissue distribution gene expression experiment, the transcript level was relative to the sample with the lowest Ct value. Data are the mean ± standard deviation of three independent experiments performed in triplicate. Statistical analyses were performed using a one-way analysis of variance followed by Tukey's posthoc test (GraphPad Software, Inc., San Diego, CA, USA). A p-value < 0.05 was considered significant.

	Primer sequences	
Gene	*Forward*	*Reverse*
GAPDH	5'- AAGCCACCATCAAGAAGG-3'	5'-CCACCAGTTTCACAAAGC-3'
G1	5'-GCTTCTTGCGTATGCTCAG-3'	5'-TTCCTCAGTTGGCTCTATCATT-3'
G2	5'- GTCATCCAGGTCGTCAAG-3'	5'- TAGGTCAATCTCTAATTCAGC-3'
G3	5'- TTCCTGGTCCTCCTGGTTCTC-3'	5'- TGGCACCTGGATTACCTCTTCT -3'
G4	5'- TGGATATTGTGATGGCTGTA -3'	5'-CTGTCAGTGTTGAGGAAGA-3'
G5	5'- ACACTTCATCAGCAGACTTG-3'	5'- CTCTCCAGCAGAACACTTG-3'

***Table 1**: primer sequences used for evaluating the expression profile of G1, G2, G3, G4, and G5, using GAPDH as housekeeping gene.*

3.2 Results

3.2.1 Transcriptomic analysis

The results of the transcriptomic analysis are reported in Tables 2 and 3. Table 2 shows the length statistics and composition of the assembled sequences, while Table 3 contains the completeness assessment of the transcriptome.

Length statistics and composition of assembled sequences	
Number of sequences	612503
Total length (nt)	493944409
Longest sequence (nt)	52543
Shortest sequence (nt)	176

Mean sequence length (nt)	806
Median sequence length (nt)	355
N50 sequence length (nt)	1825
L50 sequence count	57305
Number of sequences > 1K (nt)	97164 (15.9% of total number)
Number of sequences > 10K (nt)	2869 (0.5% of total number)
Number of sequences > 100K (nt)	0 (0.0% of total number)
Number of sequences > 1M (nt)	0 (0.0% of total number)
Number of sequences > 10M (nt)	0 (0.0% of total number)
Base composition (%)	A: 27.97; T: 28.39; G: 21.66; C: 21.98
GC-content (%)	43.64
Number of non-ACGTN (nt)	0

***Table 2**: length statistics and composition of the assembled sequences.*

Completeness assessment of the transcriptome	
Total number of core genes queried	954
Number of core genes detected	
Complete	852 (89.31%)
Complete + partial	891 (93.40%)
Number of missing core genes	63 (6.60%)
Average number of orthologs per core genes	8.58
% of detected core genes that have more than 1 ortholog	93.19

***Table 3**: completeness assessment of the transcriptome.*

3.2.2 Gene sequences and their translation products

We found five different genes encoding for fibrillar collagens within the genome of *Chondrosia reniformis* (which encompasses 14 chromosomes), that were arbitrarily named gene 1 (G1), gene 2 (G2) and gene 3 (G3) located on Chromosome-13, and gene 4 (G4) and gene 5 (G5) located on Chromosome-2. Here, we especially considered the genes from their 5'-UTR regions to their 3'-UTR regions, which flank the open reading frame (ORF), i.e., the continuous stretch of DNA codons that begins with a start codon and ends at a stop codon. The exact dimensions of each gene are still unknown, because with the methods that we applied it was not possible to determine the precise length of the 5'-UTR regions; therefore, their size is indicated approximatively.

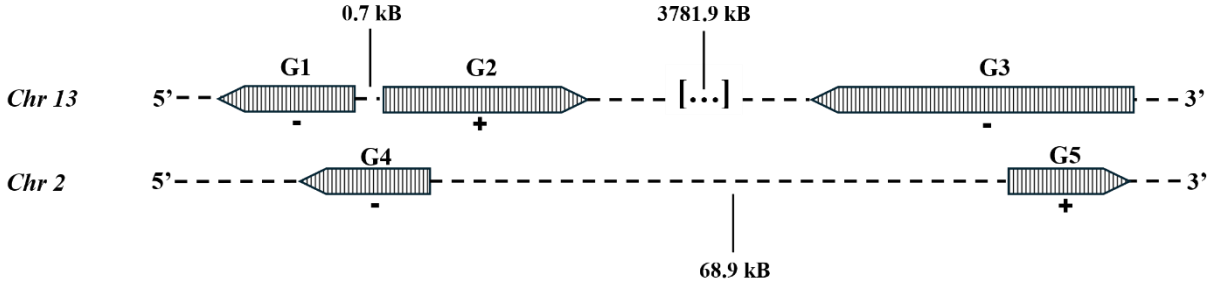


***Figure 10**: position of* C. reniformis' *fibrillar collagen genes on chromosomes and their relative dimensions and distance.*

3.2.2.1 Gene 1

G1 is located on chromosome 13 (genomic coordinates of the ORF: Chr13:1259974-1268442, with *minus* orientation) and is approximately 9 kb long. The ORF has 31 exons and 30 introns; the size of exons ranges from only 5 nucleotides (exon 1) to 449 (exon 3). Most of the exons begin with a complete codon, except for exons 2, 3, 5, 18, 20, 23, and 27. Contrarily to what occurs in the greatest part of Metazoans, not all exons encoding the major triple helical domain start with a complete codon for glycine (namely, codon 18, 20, 23, and 27). The biggest intron measure 759 bp, while the smallest one is only 56 bp long. The 3'-UTR region is 702 bp long and is coded by two exons, measuring 310 and 116 bp, respectively, separated by one 276 bp long intron. It's noteworthy to underline that G1 is in a head-to-head position with G2 and that they are separated by only 774 bp. G1 structure is represented in Figure 11.

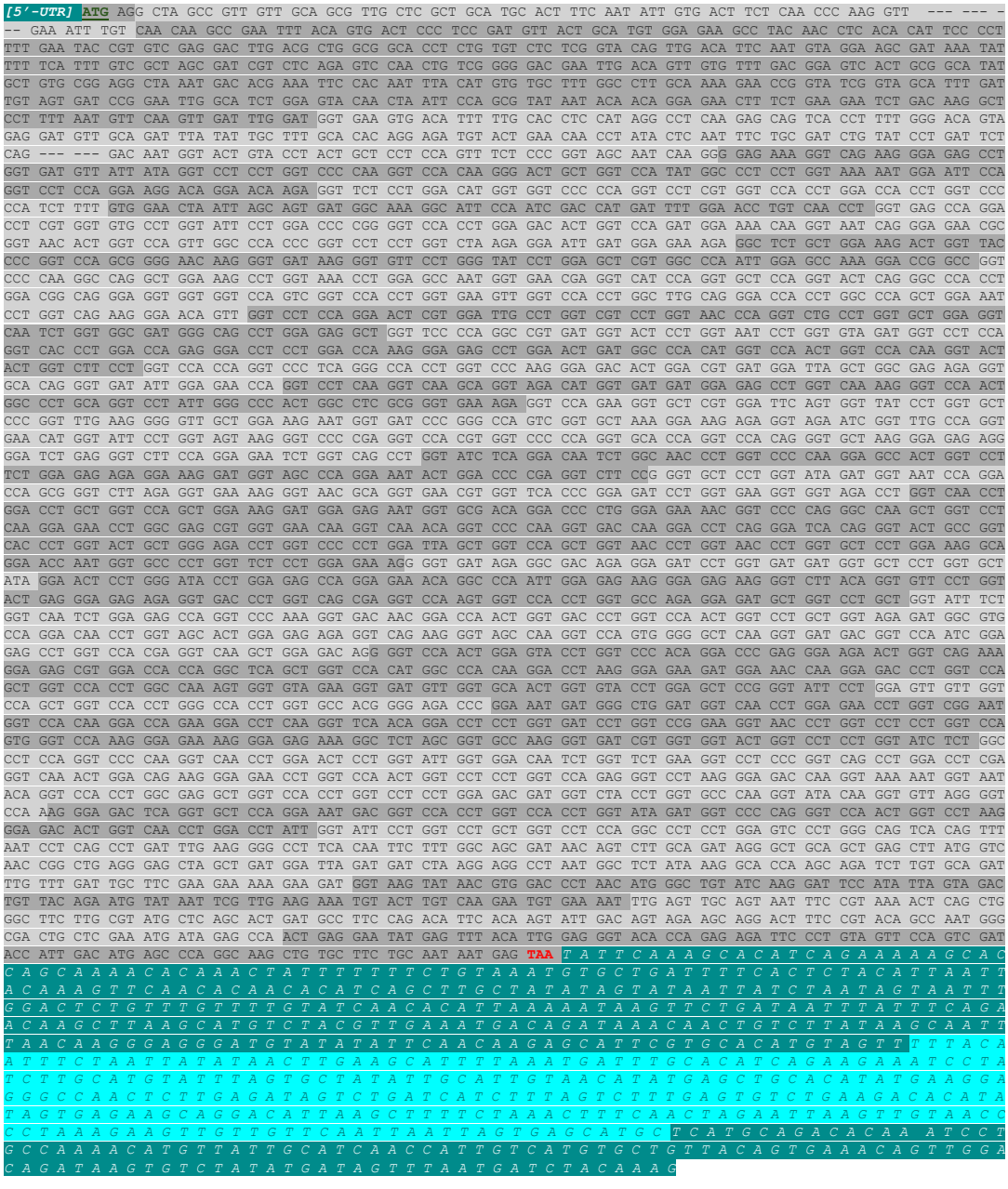


***Figure 11**: G1 introns-exons assessment. The introns and the exons of the coding sequence are highlighted in light grey and dark grey, respectively. The introns and the exons of the 3'-UTR are highlighted in light blue and dark blue, respectively. The starting codon is in bold underlined dark green, while the stop codon is in bold red.*

3.2.2.1.1 Putative protein from G1 – structure and features

The coding sequence of G1 produces a pre-proprotein (p-pP1) of 1549 amino acids if predictively translated (that is, eliminating the intron sequences from the obtained G1 ORF), or of 1544 amino acid if translated from the transcript obtained from transcriptomic analysis, but with an overall identity of 98.9%. The reason behind this little difference in size and AA residues still must be clarified. P-pP1 has a molecular weight of 149.9 kDa and an isoelectric point of 5.00; the amino acid composition of p-pP1 is shown in Figure 31. The structure of p-pP1 (considering the amino acid sequence obtained from the transcriptome analysis) is represented in figure and encompasses a 20 AA signal peptide, a N-terminal 56 AA minor triple helix, a 1017 AA major triple helix, and a 165 AA C-terminal COLFI domain (e-value: 3.85e-9); the prediction system found 4 potential N-glycosylation sequons and 23 potential O-glycosylation sites with the set threshold value of 0.8. Figure 12, Figure 13, and Figure 14 report the structure of p-pP1 referred to its corresponding gene (G1), the amino acid sequence of p-pP1, and the predicted domains of p-pP1, respectively.

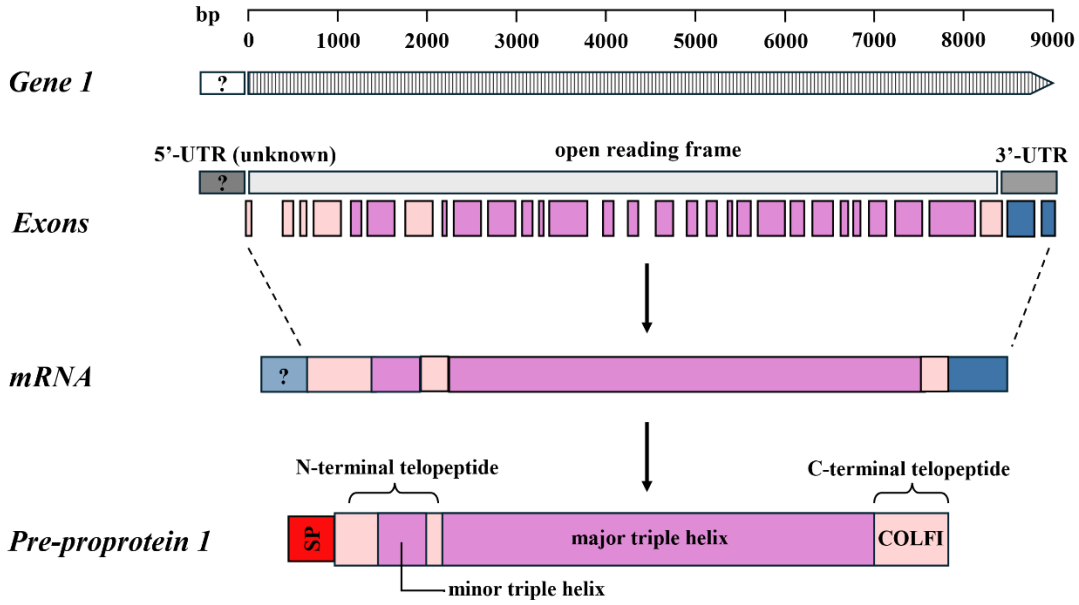


***Figure 12**: the path from gene 1 to p-pP1. The exons coding for triple helices domains are coloured in dark violet, as well as the triple helices domains in the pre-proprotein; the UTRs are coloured in blue, while the signal peptide (SP) is indicated in red.*

MRLAVVAALLAACTFNIVTSQPKVEICQQAEFTVTPSDVTACGEAYNLTHSPFEYRVEDLTLAAPLCLSVQLTFNVGSDKYFSFVASDRLRVQLSGDELT
VVFDGVTAAYAVRRLNDTKFHNLHVCFGLAKEPVSVAFDCSDPELASGVQLIPAYNTTGELSEESDKAPFNVQVDLDGEVTFLHLHRPQEQSPFGTVEDV
ADLYCFAHRRCTEQPILNFCDLYPDSQDNGTVPTAPPVSPGSNQGEKGQKGEPGDVIIGPPGPQGPQGTAGPYGPPGKNGIPGPPGRTGTRGSPGHGGPP
GPRGPPGPPGPPSFVELISSDGKGIPIDHDFGTCQP*GEPGPRGVPGIPGPRGPPGDTGPDGKQGNQGERGNTGPVGPPGPPGLRGIDGERGSAGKTGYPG*
PAGNKGDKGVPGYPGARGPIGAKGPAGPQGQAGKPGKPGANGERGHPGAPGTQGPPGRQGGGGPVGPPGEVGPPGLQGPPGPAGNPGQKGTVGPPGTRGL
PGRPGNPGLPGAGGQSGGDGQPGEAGSPGRDGTPGNPGVDGPPGHPGPEGPPGPKGEPGTDGPHGPTGPQGTTGLPGPPGPSGPPGPKGDTGRDGLAGER
GAQGDIGEPGPQGQAGRHGDDGEPGQKGPTGPAGPIGPTGLAGERGPEGARGFSGYPGAPGLKGVAGKNGDPGPVGAKGKRGRIGLPGEHGIPGSKGPRG
PRGPPGAPGPQGAKGERGSEGLPGESGQPGISGQSGNPGPQGATGPSGERGKDGSPGNTGPRGLPGAPGIDGNPGPAGLRGEKGNAGERGSPGDPGEGGR
PGQPGPAGPAGKDGENGATGPLGENGPQGQAGPQGEPGERGEQGQTGPQGDQGPQGSQGTAGHPGTAGRPGPPGLAGPAGNPGNPGAPGKAGTNGAPGSP
GERGDRGDRGDPGDDGAPGAIGTPGIPGEPGETGPIGEKGEKGLTGVPGTEGERGDPGQRGPSGPPGARGDAGPAGISGQSGEPGPKGDNGPTGDPGPTG
PAGRDGVPGQPGSTGERGQKGSQGPVGAQGDDGPIGEPGPRGQAGDRGPTGVPGPTGPEGRTGQKGERGPPGSAGPHGPQGPKGEDGNQGDPGPAGPPGQ
SGVEGDVGATGVPGAPGIPGVVGPAGPPGPPGATGRPGNDGLDGQPGEPGRNGPQGPEGPQGSTGPPGDPGPEGNPGPPGPVGPKGEKGEKGSSGAKGDR
GGTGPPGISGPPGPQGQPGTPGIGGQSGSEGPPGQPGPRGQTGQKGEPGPTGPPGPEGPKGDQGKNGNTGPPGEAGPPGPPGDDGLPGAQGIQGVRGPKG
*DSGAPGNDGPPGPPGIDGPQGPTGPKGDTGQPGPIGIPGPAGPPGPPGVPGQS*QFNPQPDLKGPSQFFGSDNSLAD**RAAAELMVNRLRELADGLDDLRRP**
NGSIKAPSRSCADLFDCFEEKEDGKYNVDPNMGCIKDSILVDCYRMYNSLKKCTCQECENLSCSNFRKTQLGFLRMLSTDAFQTFTSIDSRSRTFRTANG
RLLEMIEPTEEYEF**T**LEG**T**PERFPVVPVDTIDMSPGKLCFCNNE

***Figure 13**: amino acid sequence of p-pP1. the signal peptide is highlighted in yellow, the minor triple helix is bold light blue, the major triple helix is in bold black italics and the C-terminal fibrillar domain is in bold green; potential N-glycosylation sequons are in bold red and underlined, while potential O-glycosylation sites are in bold dark blue and underlined as well. The RGD motifs are in bold pink and underlined.*

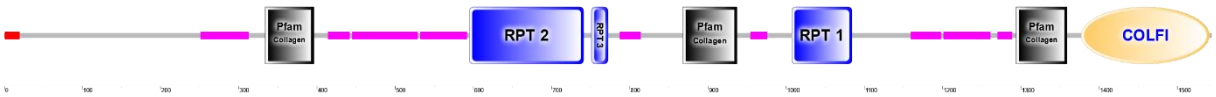


***Figure 14**: predicted domains of p-pP1 (image obtained with http://smart.embl-heidelberg.de/).*

3.2.2.2 Gene 2

Like G1, G2 is located on chromosome 13 (genomic coordinates of the ORF: Chr13:1269216-1281913, with *plus* orientation) and is about 13 kb long. It has 52 exons and 50 introns: the shortest exon (exon 28) measures 43 nucleotides, while the longest one (exon 2) is 499 nucleotides long. While most exons begin with a complete codon, many of them, even in the triple helix domain, start with an incomplete codon. Quite similarly to G1, the biggest intron is 862 bp long, while the smallest one measures 59 bp. The 3'-UTR region is coded by one exon measuring 220 bp. As underlined in 3.2.1, G2 is in a head-head position with G1. G2 structure is represented in Figure 15.

[5'-UTR] ATG GCT CCC CGG AGC CTG CTT TTC GCT TTT GTC GTC GTT GTG GTA GTC CGA TTT AGC TCA GCT GAC ACA GAC TAT TAC TCG
GAG AAG TGT TAC AAC CTT CCT GAA GGT CAT ATT GAC TTA GTG CAG AAA TCA GGT ATC ACC TAT CTG CCT GGC AAT GTG GAA TTA ACT
GGA GAT GAT AAC TTT CTA CCA GCC TAT AAA TTG CAA GCT CAA GAT GGC ATT GCT GTC AAA TGC GAA GAA GTT TTA GAA AAT GGT CTT
TTG GAT GAG TTT TCA TTG ATT GCT AAG GTC CGC TTA GAT AAA TCC TTG AAA AAA GCA AAC ATT TTG ACC ATC ATG AAC AAG GAT GCT
GTT GAT TTG AGT GTC AAC ATG TCG CTA ACA GAG ACA GGA ATA ACA GTT CAT TTT GAG TGT TCT AAT ATA CTT GCT GAC TTT GCA GTA
AAT TTG AAG GAT CTG GGA GAG TGG CAT AAA CTG GGG ATA GTG ATC ACA ACC AAA GCC ATT CAG CTA ACA TAT GAC GGT CAA CTA GTG
GCA GAA GGC TTT GTA CCC GAA AGT GAA GCC TGC ATG TTT ATG TGC AGA GAC AGG GAA ATT CAT ATA GCT CGC AAT AAC GAT GGT CCA
ATC ACT GTT CAG CAA CTA GTT ATG CTT CCT GAT CTT GAG TAT CAG AAT CAT GAT AAA TAC AGC AGC TGC GCA GCC AAA AGT CGT CAA
GCT CGT GGT AGT CCC ATA ACA GGA AGT GGA AGT GGA ATG ATT GAG AAA GGT ACT GCT GGT GAT CCT GGT GAT AAG GGT GCC TCT ATT
CCT GGA CGT CCT GGA CCA CCT GGA AAT CCA GGT GTT CCT GGT GAT GCT GGT ACC CCT GGT ATA CCT GGT ACT TCT GGC TCA AGA GGC
CCG CCT GGA GCT GTT GGA CCT GAT GGT CAT GAT GGG TTG GCA GGT CAA TTT GGG CCA CAA GGA CCA CCT GGT CCA ATT ATG TTT AGA
CCT GGG GTG TAT GCA ATC TAT TCA TCT GGC GGA AAG GGT CCA GTT GGA ACA CAA AAA ACA GCT CTT CCT CTG CCT GCC GGA AAT CCA
GGT AAC GAG GGA CCC CCA GGG GCT CAG GGC TCA GCT GGT GCT CAA GGG CAC CCT GGT TTA CTT GGT GAA ACT GGA ATC CAA GGT CAA
CAG GGA GAA CCT GGT GTT GCA GGG TCG AAA GGC CCT CCT GGT GCT TCA GGA TCA GAT GGT GGA GTG GGT AGT ACT GGT CCC CGT GGT
AAG AAT GGA CAT CCA GGT CCA AAT GGA GAT CCA GGT GAT CCT GGA AAT GGC GGA CCA ATT GGA CAA CAG GGT TTT CCT GGT CAC CAG
GGC CGA CCT GGC CCA CCA GGA ACT GAT GGG ACG AAT GGA GCT GAT GGA TTA ACT GGA GAC CCA GGT CCT GAA GGT CCT CCA GGT TTA
CCA GGA ACA GAT GGT TCT CCT GGA GAC CAA GGC AAA CGT GGA TCT GTT GGA CCA CCT GGT AAT CCT GGA GTA CAG GGT GTA TCT GGT
CCG CCT GGA ACA CCA GGA GAG ATT GGC CCT CCT GGA CTG ATT GGC CCC CCT GGT AAT GCA GGA AAT CGT GGG GCA AGT GGT AGC CCT
GGT CTA CAA GGC CAA GCT GGA ACT CCT GGT GAA AGA GGA GCT TCT GGT GAG CCC GGA CGT GCT GGT GCT CCA GGA AAA GAC GGT AGA
CAA GGC TTC CAA GGT ATT CCT GGT GCA GTA GGA AGC AGA GGG CCT CCA GGT TCT GAT GGT AGG CAA GGT TTT CAG GGA ACT CAG GGT
CCC CAA GGA CAA CCT GGA CCA GCT GGT AGT GAA GGT AGA ATT GGA CAA AAG GGA CCA ACT GGT AAG GAG GGA ATT GAC GGT ATA CCT
GGA CCA AAA GGA GAA AAA GGA ATC ACT GGA TTA CCC GGA CGT CCC GGT GAG CCT GGA CGT GAT GGA TAT GAT GGT GAG CGG GGC CAT
CCC GGA ATA CCT GGA CCA CCT GGT CCA TCT GGT CCA CCA GGT CAA AGA GGT GAA CGT GGA ATT CCT GGG CCT CAA GGT TTT ACA GGA
GCA ATT GGT GTA ACA GGT TCT CCA GGT GAA ATT GGT GCT CCC GGA AGA AGA GGG CCA CAA GGC CCC CCT GGT CCC CCT GGC CCA CGG
GGA GCC CCA GGT GCA GAT GGC TTC CCT GGC GAG ACT GGT AAC GAA GGT GCC ACT GGA GAG CCA GGA GTT ATC GGA GAC CCT GGT GGA
CCT GGT GAT CCT GGT CCT GTA GGA AGA CCT GGA AGT TCT GGT ACA ATT GGT CAA GTT GGT GTT GAA GGT GAC CCT GGA CAA CCT GGC
TCA CAG GGT CCA AGA GGA ACT AAT GGT GAT GAT GGA TTC GAA GGA CAG AGA GGA GAC CCT GGT ATT CCT GGT ATA AAC GGA AGA CCG
GGG CCC TTG GGA CCT CAA GGC AGA TCT GGA GAG TCT GGA GAG CGC GGC GCT GTT GGC TTT CCA GGA AGA AGA GGC AGT GAT GGC AAT
CCA GGT CCT CAA GGT GCT GCT GGT GAA CCC GGT GAA CCT GGT GCT GTG GGA GCA CCC GGT AGC TCT GGA CCA ATT GGT GAA CCT GGT
CCA ACT GGG TCC AAG GGA GAA AGA GGA CAA CTT GGA TTT CCT GGT GCT AAG GGG CCA CGA GGT CCT GAT GGT GAT GTC GGT GAA AAG
GGT GCA GAT GGA CAG AAT GGA GCC AGA GGT AAC CCT GGT GAC GCT GGC ACT CCC GGA CAT CCT GGA ACA CCC GGT CAA CCG GGT GAA
GAT GGT CCC GAT GGA GCT CAG GGC TTA AGA GGA CCA AAG GGA AGC AAG GGT GAA CAG GGA GAG AAG GGA AGA GCT GGG TCT GAA GGA
CCA AAG GGT TTG AAT GGG GCA ATA GGT AGA CCT GGT CCG GTT GGC AGA CGA GGT CCA AAG GGA TCC CGA GGT GAC CCT GGT GAT GGT
GGA GCT GCT GGT CCG AAG GGA CCA GAG GGG GTT GAT GGT CTT ATT GGA GAG CCT GGT CAA CCT GGT CCA ATT GGT GCC GAG GGT TCT
TCT GGT TTA GAA GGA TTC CTT GGG GAC AAA GGT TCC AAG GGT GCA AGG GGT GGT CCT GGA AAC CGT GGG CGA CCT GGT CAA GAT GGT
GTT CCA GGA CAA GAT GGT AGA GCG GGA GAG AAG GGT GAA GGT GGG GAA ACA GGT GAC AGG GGA CAA CAA GGT TTG AGA GGC AAG GTT
GGA GAC CCT GGA CTC GTG GGA GAT CTT GGA GCA CAG GGC CCA CAA GGC AGC CAA GGC CTT GTA GGT CCT CCT GGA ATC CCT GGA GAG
CCT GGA AGT GGT GGA GAA CCA GGA GAT CAA GGA CCA AGA GGG CCA GAG GGA CCT CAG GGT TCT CCT GGA GTT AGA GGA GGC CGT GGA
GAA CGT GGT ACA CCT GGT GCA GTT GGA CCC AAA GGA CCT CCT GGT AAA AAC GGA GCC GAT GGA CCC AGA GGT CTT CCT GGA GCA TCT
GGT CCA CCT GGT TCC CCT GGC AAC CAA GGT CCA GAA GGA TCT AGG GGA GCA GAT GGA GAA TAT AAT GGA TTC CCA GGT GAT GAT GGT GAA
AAC GGC CTT GTT GGT ATT CCG GGT GAA CCT GGT CCC AAG GGA GCG AGA GGG ACA CGA GGA GAA CTG GGA AAA ACT GGT CGT GGA GGT
CCA GCT GGT CTA CAA GGT GCG GCA GGT AAT CCT GGA GAC CCT GGA GAT AGA GGT CAG GCC GGT GAG ATT GGT CTA CCT GGT ACT GAG
GGC CAG AGA GGT CAA GGA GGT TCC AGA GGA GAT GAT GGT ATC GGT GGT CAA TCT GGT ACC GAT GGA GAC CCT GGT AAT GAT GGT GTG
GCA GGT ATC CGA GGA GCT AGA GGA GAG CCA GGT GCT ACT GGT CCT GAA GGT GCT GCT GGT CAA AAG GGA GAC CGG GGC AGA TTC GGA
GAA CAG GGT AGG CCA GGA AAT GAT GGT CCC CCT GGC CGA CGA GGA AGG GTT GGA AAC TTG GGA GAG ACT GGT GCA GAA GGT GAT GAG
GGG ACC AGA GGT TAC ACT GGA GAT AGG GGT CCT GAG GGT GCT ATT GGA ATT TCT GGA GTT ACT GGT AAC CCA GGT CCA CAA GGA ATT
AAG GGA CCA CCT GGT GAC ACT GGT CAT CCA GGT CGT CAA GGT CCA TCT GGT CCA CAA GGT CCA CCT GGT ATT CCT GGA ACA GAT GGT
CTG ACA ATC CAT AAT CTC ATC AAG CCA CCT TCA CAA TTC TTT GAT GCA ACA TCC TCA AGT GAC CCC CTT ACA GAT GCT GTC GTG GAG
TCA ATC CTC AAG TCT TTC CAA TAT GCT GAA TTA GAG ATT GAC CTA ACC AAA AAG CCT GAT GGT ACA ATG AAG TAT CCG GCA ATC TCT
TGT GAT GAT CTA CAT AAA GAT TAC CCT CAG CTT CCT AGC GGT AAT TAC ACC CTC GAC CCT AAT GGT GGC TGT AAA AAT GAT GCC TTT
GAG ACA TAC TGT GAA TTT AAC AAC TCT GTG AAA ATG TGC TTA ACT CCA AAA ATA CCC ACA CTT CTA CCA ATG GGT ACT TAC AAG TAC
TAT GTA AAC AGT GAG GGA TAC TAT TCA CCA AAT GAT TTT GGA CTA AAT TTG AGG TTC TTT GAA TAC TAT GGG TCA GTG ACA CAA CTT
AAG TTT TTA CAA ACT AAA GCA ACA AGA GTA ACA CAG ACC ATT AGA GTA TTA TGC AAG AAT TAC GAT CCT CTC CAT AAA CAA CCA GTA
TTT ATT GGA ATG AAT GAT GAA ACC GTT ATG GAC GAA CCA CGT ATG GAA GAA AAC CAG TGT CAG TAC TTC AAT GGT CTA AGT GCT CAT
GTA GAA CTT GAG CTA AGC AGC AAC GAT CCA TCC TAT TTA CCC ATC TAT GAA ATG CGT CTC TAC TTG GGA AGA AAG ACG AAT GAA GAG
TTG GGC ATT GAA CTT GGG GAT CTC TGC TTT GAA TAC **TAA** *G T C A T G T A A A C T T A G C A C T A T C T T G T A T T G T A*
C C C C T A A A T T T G G A G G T C A C T G A C A C A A G T T C A T T G T T T T T C T A G T G T T C A C G C A T C T
C A T C T G C T T T G A A T A A C T C C A A T A C A T T T A A C T G T A T C T T T A C A C T G T T A T A A T A A A T
A T A T T G C A T A T G T C T C A A C T A A T A T C A T A G C T A T A T G T C T T T C A C C A T T A A T T A A T T A
C A T G C A A T A A T G A A

***Figure 15**: G2 introns-exons assessment. The introns and the exons of the coding sequence are highlighted in light grey and dark grey, respectively. The introns and the exons of the 3'-UTR are highlighted in light blue and dark blue, respectively. The starting codon is in bold underlined dark green, while the stop codon is in bold red.*

3.2.2.2.1 Putative protein from G2 – structure and features

The pre-proprotein (p-pP2) deriving from the CDS of G2 is 1605 amino acids long, has a molecular weight of 158.6 kDa and a predicted isoelectric point of 5.08. For the amino acid composition of p-pP2, see Figure 31. P-pP2 comprehend a 21 AA signal peptide, a N-terminal 84 AA minor triple helix, a 1020 AA major triple helix, and a 219 AA C-terminal COLFI domain (e-value: 7.22e-13); the prediction system highlighted zero potential N-glycosylation sites and 23 potential O-glycosylation sites with the set threshold value of 0.8. Figure 16, Figure 17, and Figure 18 report the structure of p-pP2 referred to its corresponding gene (G2), the amino acid sequence of p-pP2, and the predicted domains of p-pP2, respectively.

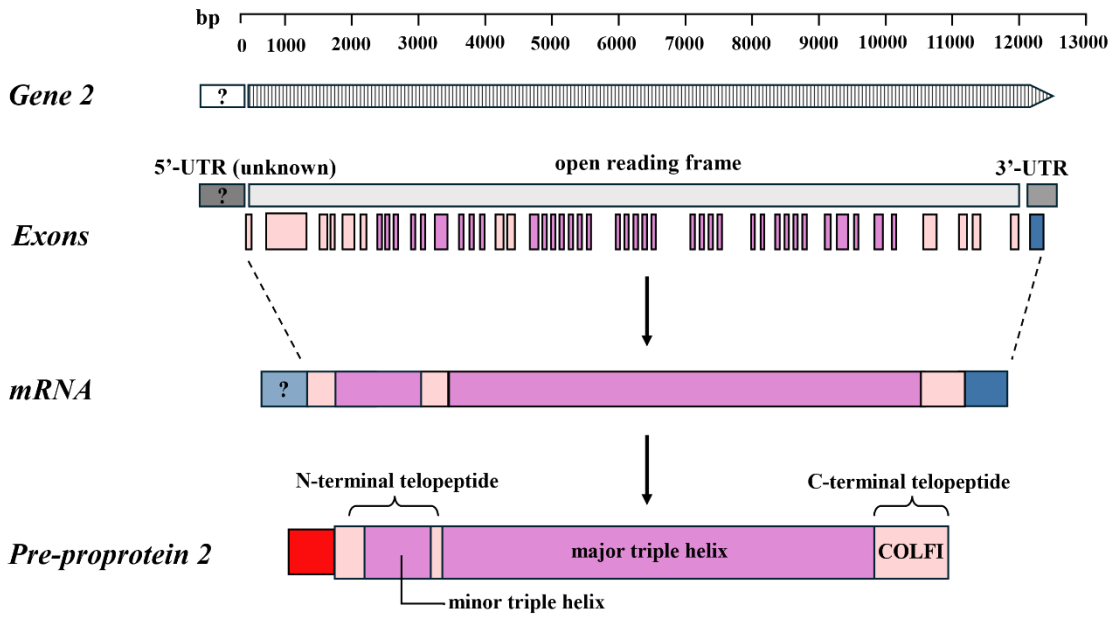


***Figure 16**: the path from gene 2 to p-pP2. The exons coding for triple helices domains are coloured in dark violet, as well as the triple helices domains in the pre-proprotein; the UTRs are coloured in blue, while the signal peptide (SP) is indicated in red.*

MAPRSLLFAFVVVVVVRFSSADTDYYSEKCYNLPEGHIDLVQKSGITYLPGNVELTGDDNFLPAYKLQAQDGIAVKCEEVLENGLLDEFSLIAKVRLDKS
LKKANILTIMNKDAVDLSVNMSLTETGITVHFECSNILADFAVNLKDLGEWHKLGIVITTKAIQLTYDGQLVAEGFVPESEACMFMCRDREIHIARNNDG
PITVQQLVMLPDLEYQNHDKYSSCAAKSRQARGSPITGSGSGMIEKGTAGDPGDKGASIPGRPGPPGNPGVPGDAGTPGIPGTSGSRGPPGAVGPDGHDG
LAGQFGPQGPPGPIMFRPGVYAIYSSGGKGPVGTQKTALPLPA*GNPGNEGPPGAQGSAGAQGHPGLLGETGIQGQQGEPGVAGSKGPPGASGSDGGVGST*
GPRGKNGHPGPNGDPGDPGNGGPIGQQGFPGHQGRPGPPGTDGTNGADGLTGDPGPEGPPGLPGTDGSPGDQGKRGSVGPPGNPGVQGVSGPPGTPGEIG
PPGLIGPPGNAGNRGASGSPGLQGQAGTPGERGASGEPGRAGAPGKDGRQGFQGIPGAVGSRGPPGSDGRQGFQGTQGPQGQPGPAGSEGRIGQKGPTGK
EGIDGIPGPKGEKGITGLPGRPGEPGRDGYDGERGHPGIPGPPGPSGPPGQRGERGIPGPQGFTGAIGVTGSPGEIGAPGRRGPQGPPGPPGPRGAPGAD
GFPGETGNEGATGEPGVIGDPGGPGDPGPVGRPGSSGTIGQVGVEGDPGQPGSQGPRGTNGDDGFEGQRGDPGIPGINGRPGPLGPQGRSGESGERGAVG
FPGRRGSDGNPGPQGAAGEPGEPGAVGAPGSSGPIGEPGPTGSKGERGQLGFPGAKGPRGPDGDVGEKGADGQNGARGNPGDAGTPGHPGTPGQPGEDGP
DGAQGLRGPKGSKGEQGEKGRAGSEGPKGLNGAIGRPGPVGRRGPKGSRGDPGDGGAAGPKGPEGVDGLIGEPGQPGPIGAEGSSGLEGFLGDKGSKGAR
GGPGNRGRPGQDGVPGQDGRAGEKGEGGETGDRGQQGLRGKVGDPGLVGDLGAQGPQGSQGLVGPPGIPGEPGSGGEPGDQGPRGPEGPQGSPGVRGGRG
ERGTPGAVGPKGPPGKNGADGPRGLPGASGPPGSPGNQGPEGSRGADGNNGFPGDDGENGLVGIPGEPGPKGARGTRGELGKTGRGGPAGLQGAAGNPGD
PGDRGQAGEIGLPGTEGQRGQGGSRGDDGIGGQSGTDGDPGNDGVAGIRGARGEPGATGPEGAAGQKGDRGRFGEQGRPGNDGPPGRRGRVGNLGETGAE
*GDEGTRGYTGDRGPEGAIGISGVTGNPGPQGIKGPPGDTGHPGRQGPSGPQGPPGIPGTDGLT*IHNLIKPPSQFFDATSSSDPLTDAVVESILKSFQYAE
LEIDLTKKPDGTMKYPAISCDDLHKDYPQLPSGNYTLDPNGGCKNDAFETYCEFNNSVKMCLTPKIPTLLPMGTYKYYVNSEGYYSPNDFGLNLRFFEYY
GSVTQLKFLQTKATRVTQTIRVLCKNYDPLHKQPVFIGMNDETVMDEPRMEENQCQYFNGLSAHVELELSSNDPSYLPIYEMRLYLGRKTNEELGIELGD
LCFEY-

***Figure 17**: amino acid sequence of p-pP2. the signal peptide is highlighted in yellow, the minor triple helix is bold light blue, the major triple helix is in bold black italics and the C-terminal fibrillar domain is in bold green; potential N-glycosylation sequons are in bold red and underlined, while potential O-glycosylation sites are in bold dark blue and underlined as well. The RGD motifs are in bold pink and underlined.*

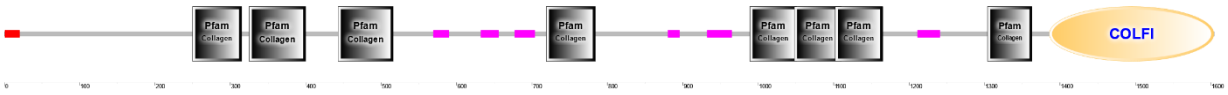


***Figure 18**: predicted domains of p-pP2 (image obtained with http://smart.embl-heidelberg.de/).*

3.2.2.3 Gene 3

G3, like G1 and G2, is sited on chromosome 13 (genomic coordinates of the ORF: Chr13: 5063887-5083881, with *minus* orientation) and is the longest one, measuring about 20 kb. It is constituted by 57 exons, of which the biggest one (exon 2) is 569 bp long, while the smallest (exon 51) is only 18 bp long. A total of 16 exons, distributed quite evenly along the ORF, begin with an incomplete codon. The biggest and the smallest introns are long 718 bp and 55 bp, respectively. The 3'-UTR region is coded by a single 78 bp long exon. G3 structure is represented in Figure 19.

[5'-UTR] **ATG** ATT CTA CGC GGA GTG ATC TCC GTC CAT ATA CTG TTG TCG TCG TGG TTC ATC TTC CAA CTC GGC CAC GAT GTC
CAA GCT GCT CCT CGC GTT TGG AAA CGC TCC TCT AAC GAT GCA TCC TCG CAC AGC AAC GTC CCA GTA GTC GAC CTG CTA GGT AGG
CAG AGC TTT TCT TCA CCA TCT CCC ATC CCA CCT TAC CAA TCC AAT ATT TCA GAC GAA GCT ATT CCA TCA TTG CCA AAT AAC TTC
AGC TTA TTT ATT GTA TTC AGA GAA ATC CCT ATT ACA AAC TTT CGA TTA GTA TCT ATT TAT AAA AGA GAA TTT TCT ACA AGT GAC
GAT TTT AGC TAC TCC TTG CAA AAC TCT ACA AAT TTT CAT GAG GTA CCA TCT ACC AAC ACA TCT AAT TTC ACC TTT AAT GGA CAA
AGT GGC TTT ACC TTT GAG AGC TCA CTA ACT ATT GGA AGT GAT GTC AAT GTC ATC TAT AAC TCA GAT TAT GAG TCA GTG TGG CAA
GAC ATT GCA TTG CAT GAA GTT TTT CTG CCT CCA ACT GTC ACT TCT ATC AAC CAA TCA AAA TTA ATA TTA ACT GTG GAG AAT GGT
GTT CTA ACA TCT TGT TTG AAT GGG GAG TAT CTT TCA AAG GAG ATT ACG TTG TGG AAT GCT TCA GAC ACA GAT CAA GTC CTG TTG
ATA TTT TCT GAG GAA TTT TAT TTA CCA TTC ATT CAG ACA TCT TTG TGT GAA GCA ATA GTT ATT CTG GAA CCC ATA ACT TCA ACA
GAC TAT AAA GCG TGA ACC AGA AAT TCC TAT TTT TTG CAA GCC AGA GCA ACA AAG AAG CCC CTG TAG AGG TCC TCC AGG CCC TCC
TGG TCC ACC TGG GCC ACC TGG TCC TCC AGG CCC CAA GGG ACC CAT AGG TGT CAG AGG TTA CCG GGG TTT GAA AGG TCC CAG GGG
TCC TGT TGG ATT TAG AGG TCC CAA GGG ACT ACC TGG CCC TCC TGG TCC AAA AGG GGA GAG AGG TCA TAC AGG AAG GCA TGG TAA
AAC AGG AGA AGT AGG CAG CAA GGG AGA TCC TGG GCC AAG AGG TCA TAC TGG TCC AAC AGG GGA TCC TGG AGG TGA AGG ACC ACC
CGG AGA GAG AGG TAG CCC GGG ACC GCC AGG AAG TGA TGG AAA AAA TGG AGA TAT GGG TCT TAC TGG TAA TGA TGC CCC AGG GGA
CCA GAA GGC AGA AAG GGT GTT AGA GGA GCA GTG GGT CCT AAG GGA GAA CCA GGG CAT CGT GGG GAA CCA GGT CTT CAA GGT CCA
CCA GGG CCT GAA GGG AAG AAT GCT GAA CGT GGG CCA ACC GGA GTA CAA GGC ATT CAA GGT CCA CCT GGG CCG CCT GGT CCC AAG
GGA TTG AAA GGA GAG AAG GGC ATG GAT GGT CCA CCT GGG GAG AAG GGA CAT CCC GGA CAA CAA GGA AAA ACT GGT ATA CCT GGA
CCA GAT GGA CCG CCC GGC CGG AGA GGA CCA GTG GGG CCC CAT GGA TCA AGG GGA TTA CCC GGT CCT AGA GGA AGA AGG GGT AGA
ATG GGT CTA CCT GGC CCT CCT GGA CAG AAA GGA AAT ATT GGC TAT TCA GGA GCC AGA GGA CGT CCA GGA TTA CCA GGT AAT CAA
GGG CCT CGT GGT CCA GAG GGT CTT CCA GGA ATA AAA GGA GAA AAG GGA GAA GAA GGG ATA AGT GGT ATT CCA GGT CTT GAT GGC
CTT TTC GGT CCT CCT GGA GAG AAG GGT GAC ATA GGA GTA GTT GGT GAG AGG GGA ATA CTT GGA CCT GAA GGA GAG AAA GGT GAA
CGA GGT CTA GCA GGA GCT AAT GGA ACA GAT GGT ACA AAG GGT ACC AAA GGA GAG AAG GGA CTC CCA GGT TAT CCG GGA AGA CCT
GGC CCC AAG GCA GGT ATC TCT GGT TTA CCT GGT CCC CGT GGT GAC CCT GCA ACA CCT GGA AAG GTT GGT CAT CCT GGA CCT CCT
GGT ACA CCT GGG ATA GCA GGA GAA AGG GGA TTA GTT GGT GAG ACA GGA CCC CGA GGG CCT CCT GGA CCA CAA GGC CCC CCT GGT
CTC CCT GGA CTG CCG GGG TTT GTG GGC CTC CCT GGA TTT AGG GGC CCA CCT GGT CCC CGA GGA CTT CCC GGA CAT CAT GGT AAA
AGA GGA CCA CCG GGT TAT GCT GGA AAT CCT GGT CAC GAA GGT CCC AAA GGA CAA CAA GGA AAG AAG GGT GAA AAA GGT GAA AAT
GGG CAT ACT GGA TCA ACA GGA TTA CCA GGT GAT GAT GGT TTA GAG GGA CCA CCA GGT CCT CAA GGG CCT CAA GGA CGC AAG GGT
GCA CAA GGT CCT AAA GGT ATT CCT GGA CTA GCT GGT GCC ATT GGA AAA CCA GGA AGA AAG GGG GAT CAG GGT GAG CCT GGT GAA
GTG GGT AGT AGA GGT GAG AGA GGT GAT AAA GGA TCG GAA GGT GAT GCT GGA GAA AAT GGA CGT CAC GGT TCA CCT GGC CTG CCT
GGA CAT AAT GGT CCA CCC GGA GTC CCA GGT GGA ACT GGA CCA CCT GGT CCA GTT GGA AAC AAA GGA GGA GTG GGA CCA CGT GGC
GAC AAT GGA CTC TAT GGT GAG AAG GGA GAT GAA GGA GAA AAG GGT CAT AAG GGC TTA CAG GGG AGT ACG GGT GTC AAG GGA GAC
GTT GGT GAT ATT GGA AGC GCT GGT GAA AAG GGA GAA AAG GGT CAA ATA GGT AGA AAT GGA CGT TCA GGG GAA AAG GGG GAA ATG
GGT GAT GAC GGA GAG GCT GGC CCA GCT GGA ATT AAA GGT TTT AAA GGT GTT GCC GGT ACC CCT GGT AAA CCA GGT AAA GAT GGC
CAC AAA GGG ATA ATG GGT GCT AAG GGC AAC GAG GGT CTG CAA GGA CCA CCG GGA GAG CTT GGA AAG GCT GGA CCT AAA GGA GTG
CAA GGA AGG AAA GGA GAC AAA GGA TCA GAT GGA TTT CAG GGA CCA CCA GGC ATT GAT GGA GAC AGA GGA CCA ATT GGA CCA CCA
GGT CCT CAG GGA CAT AAG GGA AAA CGT GGG AAG CCC GGT CCT CCT GGG GAA GAA GGT CCT GCA GGT CCG ACT GCA GAC CCT GGG
AAA TGG GGG CCT CCC GGG GAA GCT GGA GAG CCT GGG CCC AAA GGT CCT CAT GGT CCA CTT GGT GTA CGA GGT GAT GTT GGA GAT
CCT GGA TTA CCT GGT CCA GCA GGA CCA GAT GGT GCC CAA GGT CGC CCT GGA CAT CCT GGA CCA GAC GGA GAT GTA GGT CCT GCA
GGT GCC CCA GGT AGG CCT GGT TCA GTC GGA AAT CCA GGA CCA ATG GGA GAT CCT GGA CCA CCT GGA GTG AAG GGG AGA CTG GGA
GGT TAT GGG CCA AGG GGT CCT GTA GGC CGT GCT GGT TTA AAA GGA GAT CCT GGA GAG AAG GGT CCT GTT GGT CAC CCT GGG CCA
CCT TCT CCA AAA GGA GTT AAG GGA GTA AAG GGA TTG AAA GGT GTA AAA GGA ATC AGG GGG GCA AGG GGA GTG ATG GGG GAA AAA
GGT GAA ATT GGC TTT CCT GGT CCT CCT GGT TCT CCT GGT TTC AAT GGA GCT CGA GGA TCT GAC GGC AGA AGA GGT AAT CCA GGT
GCC AGT GGG GTC AAA GGA GAC AAA GGA TAT AGA GGA CCC AAA GGA CGC ATG GGG CTC TTT GGT ACC AAG GGA GAC CAG GGT TTT
CCT GGA CCG ATA GGC CCC AAG GGC TTT GCT GGG AGA CCA GGA CTA CCG GGT GAA CCT GGT AGT CCT GGT GAT ATG GGG TTA ATT
GGG AGG AAA GGA GAA AAG GGA TTA AGA GGT GAA AAG GGG AAC CCA GGA GTA GTG GGG TTA CCA GGG ATA CCT GGA AAC AGA GGA
TTG CAT GGC TTC AAG GGA CCA CCT GGT CCC ATT GGT GCT TTT GGA CCA CCG GGG CAA AAA GGT CCA CCA GGG CAT CCA GGA GCT
GCA GGC AGA AGT ACA ATA TTT TAT GAC TCC GCT AAC CCA GAA TCA CAA AAG GAT TTA TTT AAA GTT CTT AAT TTA CTG GTC CAG
AGA TAT CTA GGC AAC CAA ACA AGA CCT CTG CCT TCA ACT TTG TGT GTT CCT CCC TTA AAT GGA GTG TTA GGA GAA AGG GAC ACT
CCG GCC TTG TCG TGC AAT TCC ATC AAA CAT GTT TAT CCA CAA GCC CAA AAT GAT TGG TAC TGG ATT GAC CCT AAC GAA GGA TGC
TTA AAG GAT GCT GTA GAG GTT TTT TGT GAT TTT ACT TTA AAT CAA ACT TGC ATC AAT CCA AGA AGA AGC AAA ACA ACA GAA TAT
CAT TGG TTA CCT ACC TCA AAT GAT ACT CTC TCT ATC TCT CAT ATG GAT CCC TTA TTT AAG CCA AAC TAT GCT ATT GGA AGT GTG
CAA CTC AAT TTC ATA AAA TTA TAC AGC AAA AGA GTT GAG CAG GGT TTG AAT TTA CAT TGT ACA AGG GTC ACC AAT ACA AAA TGC
TCA GAT TGC TAT CTT CAA ATT CAA CCA CTC CTT ACA GAA AAG CAA GAA TTA GCA CAG TTA TTA CAA CTC AAT ATA TCT TTA CAA
AAT TGT TTA AGT CCA TTT GGA CAC ACA CTA AAG CTT ACA TCT GAA AGA ATT TCA CTG CTC AAT AAA TTT CCT ATC CAA GAT GTT
TCA TTA CAA TTA TAT GCA GAA AAT GTT CCA CCT AAC TTC AAT CTA AAT CTG GAA TGG GAC TTT GGA CCT CTT TGC TTT TCA CAG
GGT **TGA** *T T A T A T T C A T T A T A A T G T C T G A A G T T T T T A T T T C T T C A T A A C A A T G T A A T T A*
A T T T T A T A A A A A T T T T G T G C A T C T A A

***Figure 19**: G3 introns-exons assessment. The introns and the exons of the coding sequence are highlighted in light grey and dark grey, respectively. The introns and the exons of the 3'-UTR are highlighted in light blue and dark blue, respectively. The starting codon is in bold underlined dark green, while the stop codon is in bold red.*

3.2.2.3.1 Putative protein from G3 – structure and features

The pre-proprotein (p-pP3) predicted from the CDS of G3 is 1540 amino acids long, has a molecular weight of 156.3 kDa and a predicted isoelectric point of 9.24. The amino acid profile of p-pP3 is shown in Figure 31. P-pP3 comprehend a 27 AA signal peptide, a 1018 AA major triple helix, and a 236 AA C-terminal COLFI domain (e-value: 3.64e-8); no N-terminal minor triple helix was found. The prediction system highlighted 10 potential N-glycosylation sites and 11 potential O-glycosylation sites with the set threshold value of 0.8. Differently from the other pre-proproteins, p-pP3 displays some imperfections in the major triple helix, represented by the insertion of some alanine or serine residues (see Figure 21). Figure 20, Figure 21, and Figure 22 report the structure of p-pP3 referred to its corresponding gene (G3), the amino acid sequence of p-pP3, and the predicted domains of p-pP3, respectively.

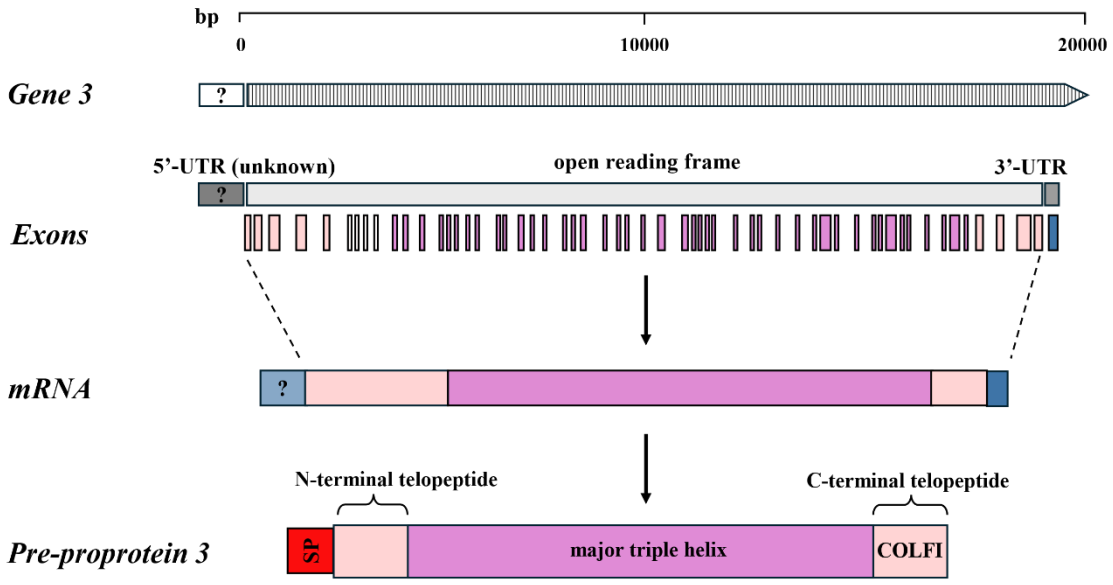


***Figure 20**: the path from gene 3 to p-pP3. The exons coding for triple helices domains are coloured in dark violet, as well as the triple helices domains in the pre-proprotein; the UTRs are coloured in blue, while the signal peptide (SP) is indicated in red.*

MILRGVISVHILLSSWFIFQLGHDVQAAPRVWKRSSNDASSHSNVPVVDLLGRQSFSSPSPIPPYQSNISDEAIPSLPNNFSLFIVFREIPITNFRLVSI
YKREFSTSDDFSYSLQNSTNFHEVPSTNTSNFTFNGQSGFTFESSLTIGSDVNVIYNSDYESVWQDIALHEVFLPPTVTSINQSKLILTVENGVLTSCLN
GEYLSKEITLWNASDTDQVLLIFSEEFYLPFIQTSLCEAIVILEPITSTDYNELCEPEIPIFCKPEQQRSPCR***GPPGPPGPPGPPGPPGPKGPIGVRGYR***
GLKGPRGPVGFRGPKGLPGPPGPKGERGHTGRHGKTGEVGSKGDPGPRGHTGPTGDPGGEGPPGERGSPGPPGSDGKNGDMGLTGNDGPRGPEGRKGVRG
AVGPKGEPGHRGEPGLQGPPGPEGKNAERGPTGVQGIQGPPGPPGPKGLKGEKGMDGPPGEKGHPGQQGKTGIPGPDGPPGRRGPVGPHGSRGLPGPRGR
RGRMGLPGPPGQKGNIGYSGARGRPGLPGNQGPRGPEGLPGIKGEKGEEGISGIPGLDGLFGPPGEKGDIGVVGERGILGPEGEKGERGLAGANGTDGTK
GTKGEKGLPGYPGRPGPKAGISGLPGPRGDPATPGKVGHPGPPGTPGIAGERGLVGETGPRGPPGPQGPPGLPGLPGFVGLPGFRGPPGPRGLPGHHGKR
GPPGYAGNPGHEGPKGQQGKKGEKGENGHTGSTGLPGDDGLEGPPGPQGPQGRKGAQGPKGIPGLAGAIGKPGRKGDQGEPGEVGSRGERGDKGSEGDAG
ENGRHGSPGLPGHNGPPGVPGGTGPPGPVGNKGGVGPRGDNGLYGEKGDEGEKGHKGLQGSTGVKGDVGDIGSAGEKGEKGQIGRNGRSGEKGEMGDDGE
AGPAGIKGFKGVAGTPGKPGKDGHKGIMGAKGNEGLQGPPGELGKAGPKGVQGRKGDKGSDGFQGPPGIDGDRGPIGPPGPQGHKGKRGKPGPPGEEGPA
GPTADPGKWGPPGEAGEPGPKGPHGPLGVRGDVGDPGLPGPAGPDGAQGRPGHPGPDGDVGPAGAPGRPGSVGNPGPMGDPGPPGVKGRLGGYGPRGPVG
RAGLKGDPGEKGPVGHPGPPSPKGVKGVKGLKGVKGIRGARGVMGEKGEIGFPGPPGSPGFNGARGSDGRRGNPGASGVKGDKGYRGPKGRMGLFGTKGD
***QGFPGPIGPKGFAGRPGLPGEPGSPGDMGLIGRKGEKGLRGEKGNPGVVGLPGIPGNRGLHGFKGPPGPIGAFGPPGQKGPPGHPGAAGRS**TIFYDSANP
ESQKDLFKVLNLLVQRYLGNQTRPLPSTLCVPPLNGVLGERDTPALSCNSIKHVYPQAQNDWYWIDPNEGCLKDAVEVFCDFTLNQTCINPRRSKTTEYH
WLPTSNDTLSISHMDPLFKPNYAIGSVQLNFIKLYSKRVEQGLNLHCTRVTNTKCSDCYLQIQPLLTEKQELAQLLQLNISLQNCLSPFGHTLKLTSERI
SLLNKFPIQDVSLQLYAENVPPNFNLNLEWDFGPLCFSQG

***Figure 21**: amino acid sequence of p-pP3. The signal peptide is highlighted in yellow, the major triple helix is in bold black italics and the C-terminal fibrillar domain is in bold green; potential N-glycosylation sequons are in bold red and underlined, while potential O-glycosylation sites are in bold dark blue and underlined as well. The RGD motifs are in bold pink and underlined. The imperfections in the major triple helix are highlighted in light blue.*

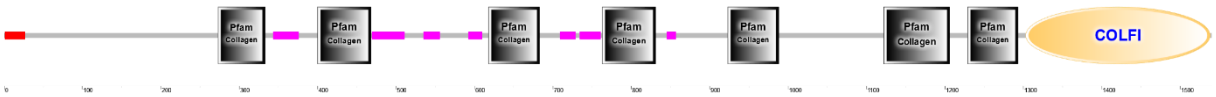


***Figure 22**: predicted domains of p-pP3 (image obtained with http://smart.embl-heidelberg.de/).*

3.2.2.4 Gene 4

G4 is located on chromosome 2 (genomic coordinates of the ORF: Chr2: 926545-934883, with *minus* orientation) and is about 8.5 kb long. The transcript of this gene is incomplete in the 5' region, therefore the intron-exon assessment is here based on the Augustus tool prediction., which reports the presence of 23 exons and 22 introns; the longest exons are exon 6 and exon 8, both measuring 324 bp, while the shortest one is exon 1, which is 19 bp long. The longest intron is 1073 bp long, while the smallest one is 56 bp long. The 3'-UTR region is coded by a single long exon of 495 bp. The structure of G4 is showed in Figure 23.

[5'-UTR] ATG AGA CTT GTA ATC TTA GGT CTA CTT GTA GGC CTC TGT GCT GTG ACT GTA TTA GGT CAA CCA GGT GAA AAG GGT GAC ACG
GGA CCT GTA GGA CCT CCT GGT CCA CCT GGA GTT TGT ACA ACT TGC CCA GCT GCG GGA GGC GGC GGA TCA GAC GGC AAG GCC GAC GGT
TCC GGT TCA ATA CAA GGT CCC CCC GGT CCC CCT GGT GCT CCC GGA GAC CCT GGT CCT GCT GGA CAG CAA GGA CCT CCT GGT CCC CGT
GGA ACT ACT GGT AGG CAA GGA CCT GCA GGA GAT AGA GGA GCT CAA GGA CCG CTA GGT CCA AGG GGT CCC CAA GGT GTT CCT GGA ACT
CCA GGT GTA CCT GGC CGT GAT GGT AGC GAT GGA CCT TCT GGA TTA CAA GGA CCA CCA GGA CCA ATT GGT CCA GCT GGA TTA CAA GGT
CCA CCA GGA CCA ATA GGT CCA AAA GGT CCA CAA GGC GAG CCT GGA CGA GAT GGA TTA CCT GGA GCT GAT GGA GAT ACT GGA AAC CGA
GGT TCA ACT GGT GCT GCT GGT CCA CCT GGA CCA CGT GGA TAT CCA GGT AAT GAT GGC TCA GAT GGT GCT CGA GGA CAG CCT GGT TCA
CCT GGG CCA GCT GGT AAG AAT GGT ATT CCC GGT GGT ATT GGA AGA CCT GGT CAA AAG GGA GAA CCA GGA CCA GGT GGT GCT AAT GGT
AAT GAT GGA CCT ACG GGA CCT CAA GGT GAT ACT GGA GCA TCA GGA AGC GCG GGA CCA CCA GGA GAA CGT GGC CTT CCT GGT GCA ACT
GGT GGT TCT GGT CAG AAA GGA GAA GCG GGA GCT CGT GGA GGA AGT GGT GCA CCT GGT TTA CCT GGT GCC ATT GGC AAC CAG GGA TCT
CCT GGT CCA TCT GGT CCA TCT GGG CCT GCG GGT GAT CAA GGT CCT GAG GGT GCT CCT GGA CCA CCA GGT GCT GAT GGA TCT CCG GGC
AAC GCT GGT ACT CCT GGC ACA CCT GGT AAC ACT GGT CCA CCT GGG CCT CCT GGA AAG AAA GGT AGA CGT GGA GAG AGA GGT CCA GCT
GGT GAT CCT GGC CCT ACT GGT CCA AGA GGT TTT ACT GGC CAA CGT GGT TTC AAT GGT GCA AAT GGC GCA CCA GGA AAG CAA GGA GAG
CCC GGA CTT CCA GGT GAG ACT GGA CCC AGA GGT CTC TCT GGA GCA AAG GGA GAG AGA GGT GAA GAA GGT AGA CCA GGT CCT CCA GGA
CAG CGT GGA GCC CCT GGT AAC CCT GGG CAG AAT GGA GAC AAT GGA ATA GAT GGA CAG CGT GGT GCT TCT GGT ATC GAT GGA GCT CCC
GGC CAG CCT GGG CCT CCT GGT GAC AGA GGA GCA CCT GGT TCA GCT GGC CCA GCT GGT AAC ACT GGA GAT AGG GGA GAT ACT GGT GAA
AAG GGA GCA ATG GGA GAG AAG GGA GAG AAA GGA GAT GAA GGT CCC CAA GGT CCA TCT GGT GGT GAC GGT GCG AAT GGA CAA AAT GGT
GGA GAT GGA CCG CCT GGT GAC ACT GGT GAT GGT GGT CCA CAA GGA GCA CCT GGA GGA CCT GGT CCT ACT GGC CCA TCT GGT CCT CCT
GGT GCA ACT GGT CCA CCC GGT CCA AAG GGA ACC GAT GGA CCT CCA GGT GCA ACG GGT CCA CAA GGA TCT GTT GGA AAT CCC GGT CCT
CCA GGC CCT CCT GGA GCT ACA GGT CAA CCC GGA CAG AAG GGT GAA CGT GGT GCT ACT GGT GAT GGT GGT GAA GCT GGT GCA CCT GGT
CCT GCT GGT GAC CCA GGT GCT CAG GGA GAT CGT GGT CTT GAT GGC CCA TCT GGT CTT ACT GGT CCA CAA GGA CCC CAG GGT CCC CAA
GGC TCC CGT GGT GCC CCT GGT CCC GAG GGC AGA CCT GGT AAT GCT GGT CCA CAA GGT GAA ATT GGA CCC GCA GGT CCT CAA GGC ATA
CCT GGA ATG GAT GGT ACC AAT GGT GTG GTT GGC CCT GAT GGT GAA AAG GGT GAA AAG GGA GCA GAT GGT CCT GCT GGT CCA CCT GGA
CTA CAG GGT CCT CCT GGT CCA AAT GGT ATT CCT GGT CGT ATT GGA CCT GCT GGT CCA CAA GGA CAA CCT GGG CCA CCT GGA CCA GGT
GGT CCA ATG GGC CCA CCT GGT TCT TTC GGT GGA CAA GGT TCA GGC GGA GCC GAT GGT GCT AAG GGA GAT GAT GGT CCA CCA GGA CCC
ACT GGT CCT CCT GGT GAC CGT GGT GTT GAT GGA AAT CCT GGA ATC AAG GGA GCC ACT GGT TCT CAA GGT GGA CCT GGT CAA CCT GGT
CCA CCT GGT CCA GTT GGT CGT CAA GGT CCT CCT GGT AGT CCT GGT TCT CCA GGT GAT TCC GGT GAT GAT GGT CCT AAA GGT GAT ACT
GGT CCA CCT GGT CAA GCG GGA CAG CCT GGT CCA ATT GGC CCG CAA GGC CCT CCT GGT GCT AAT GGA GCA AAG GGA TCC AAG GGA GAT
CGT GGT AGT GAT GGA AAC ACT GGT CCA GCT GGC GCA ACA GGA GAC AGT GGT GTT ATT GGG CCA CCT GGT CCT CCA GGT CCT CAG GGT
CCT CGT GGT GCT GAC GGA GGT CAA GGT GCT GAC GGA AAG GCA GGA TCA ATG GGA CAA ATG GGT TTG AAG GGA GAA AGA GGA TCT CCA
GGT CCA CCT GGT GTT CCT GGT CCG TCC GGC CCA AGA GGT GAC CCT GGT TCT GAT GGT ATT CCT GGT CAG GAT GGT GCT AAT GGA CCC
CCT GGA GAT GCT GGT CCC ATT GGT GCC CCT GGC GCC CCT GGT ATC CCT GGT ACG TCT GGA GCC CCC GGT CTT CCT GGA CCA CCT GGT
CCT CGT GGA AGC CCT GGA GCA TCG GGT AGA GCT GGT GCT CAA GGT CAA AGA GGA GGA GAC GGT GAG AAG GGT GAG AGG GGA GTC AAT
GGT GCA ACA GGA CAG CAA GGT GCC AAG GGA GAG ATT GGC GCT ACT GGT GTT AAG GGA GAT AGA GGT GGA TTT GGG TTG CCG GGT AGA
CAA GGT CAA TCT GGA TCT GTT GGA CCA CCC GGA CCA CCC GGT CCA CCT GGT CCT TCT GGA GAT CAA GGA ACA AGG GGT TCT CCT GGT
CCA GCT GGT CCT CGT GGT GAT CCT GGC TCC CAA GGT CCT CAG GGT ACA TCT GGT CCA GAA GGT GGA CGT GGT AGA GAT GGT TTC GAT
GGA GAC AAA GGA CAC CCT GGT CCT CGT GGT TCA CCT GGC CCA CCA GGT CCA CCT GGA TAT TGT GAT GGC TGT ACA ACT TTG GCT GAC
CCA AAG AAG CAC TTC AAC TCA GCC CAC GGA GGA GGA GAG GAA GAC GGC CCT GAA TGG TAC GAC GCC TCA AGA AGG TAT TGT GAT GGT
ACA GAG GAT AAG CCT TGC CCG ACA TGT CGA GAT GTC TTC CTC AAC ACT GAC AGT GCA AAA ACA GGC AAA TAT TTC ATT GAT CCC AAC
TTT GGA TCG CCA AGT GAC AAG GTA GAG GTA GTT TGC AAT TTC AAC GTC CAA ATG GTT AAT AAT GTT CCA GAA CAG CAC TAC AGA ACC
TGT GTT AAT GTC ACA GAC GTT TCA CAA GTC AGT GAA GTT TGG AAT GTC AAA TAC AAG ATG CAG GAT GAT CAA ATG TGG TCA CTG ATG
AGA TTA AGT TCA AAA GTG GAA CAG AAG ATG TCA GTT GCA TGC AAT GAT ATG CCA GAG ACT GTG GGA TGT TCA ATT GAA GAC AAG AGT
ACC ATA CAA CCT TCC TGT GTA TCA GGT CAA GAG ATC TGA *GTCTGTTAACTGCATATTAACTTTAGAAGACT*
TAATTTTTTGTGCAACAAACTTCAGAGTTATGCACTTGCAATCACGCTCACAGTATCT
TCTCTCTCGTGTAAACTAATAAGTTTTTCATAATAATTTCCAGCTAAATCTCTATTAT
TAATTAATCTATGCAAATATCTTAGACTGTTATGAGCGATGGCCAAGCCCTTAATTAA
CACTTGCAATCACACTCACATTATCTTCTCTCTCATGTGAACTAAAGCTAAATCTCTA
TTATTAATCTATACAAATTGGCCAAGCCCTTAACTAGATTGACATTCATTAACACTGC
TTGTATATTGAGAAGTATATTTCACCTTGAAATGAATTACCGTATATTGGGAATTGAA
TGCGTGTGTCAAAAGAATGTGAATTGCGAGTTTTGTAGCCATAAAATAAACATAACAT
GCAGGGATAACATTGTCAATTGAGGACACTGCACAAAATGAAATTCTACAATAGT

***Figure 23**: G4 introns-exons assessment. The introns and the exons of the coding sequence are highlighted in light grey and dark grey, respectively. The introns and the exons of the 3'-UTR are highlighted in light blue and dark blue, respectively. The starting codon is in bold underlined dark green, while the stop codon is in bold red.*

3.2.2.4.1 Putative protein from G4 – structure and features

The CDS of G4 produces 1228 amino acids long putative pre-proprotein (p-pP4), has a molecular weight of 113.83 kDa and a predicted isoelectric point of 4.90. Figure 31 shows the complete amino

acid composition of p-pP4. P-pP4 contains a 19 AA signal peptide, an 18 AA minor triple helix in the N-terminal region, a 1002 AA major triple helix, and a 140 AA C-terminal COLFI domain (e-value: 0.0338); the prediction system found 1 potential N-glycosylation sites and 25 potential O-glycosylation sites with the set threshold value of 0.8. Figure 24, Figure 25, and Figure 26 report the structure of p-pP4 referred to its corresponding gene (G4), the amino acid sequence of p-pP4, and the predicted domains of p-pP4, respectively.

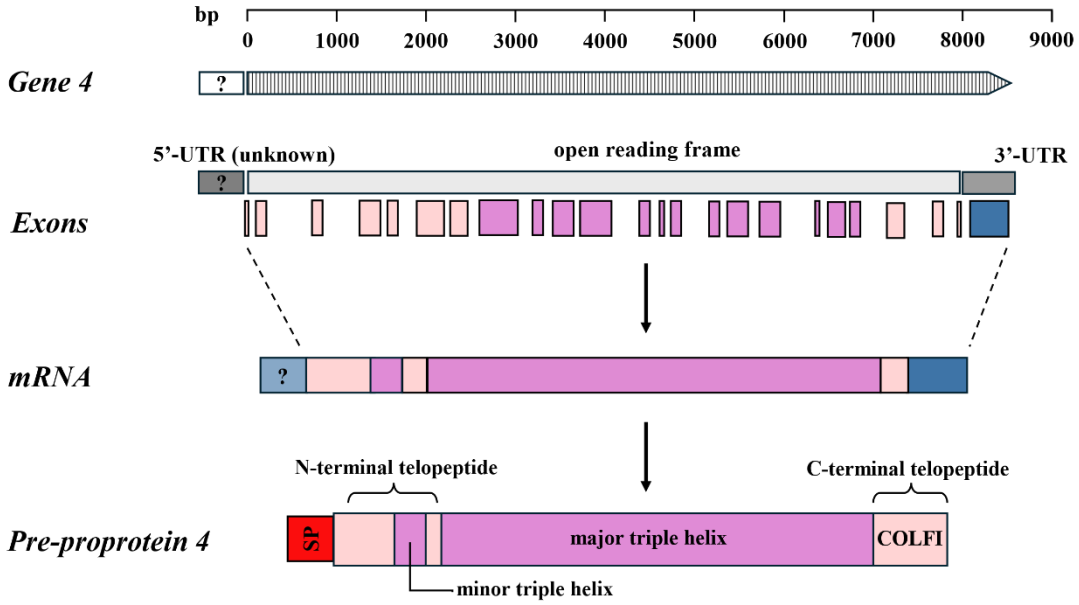


***Figure 24**: the path from gene 4 to p-pP4. The exons coding for triple helices domains are coloured in dark violet, as well as the triple helices domains in the pre-proprotein; the UTRs are coloured in blue, while the signal peptide (SP) is indicated in red.*

MRLVILGLLVGLCAVTVLGQPGEKGDTGPVGPPGPPGVCTTCPAAGGGGSDGKADGSGSIQ*GPPGPPGAPGDPGPAGQQGPPGPRGTTGRQGPAGDRGAQ*
GPLGPRGPQGVPGTPGVPGRDGSDGPSGLQGPPGPIGPAGLQGPPGPIGPKGPQGEPGRDGLPGADGDTGNRGSTGAAGPPGPRGYPGNDGSDGARGQPG
SPGPAGKNGIPGGIGRPGQKGEPGPGGANGNDGPTGPQGDTGASGSAGPPGERGLPGATGGSGQKGEAGARGGSGAPGLPGAIGNQGSPGPSGPSGPAGD
QGPEGAPGPPGADGSPGNAGTPGTPGNTGPPGPPGKKGRRGERGPAGDPGPTGPRGFTGQRGFNGANGAPGKQGEPGLPGETGPRGLSGAKGERGEEGRP
GPPGQRGAPGNPGQNGDNGIDGQRGASGIDGAPGQPGPPGDRGAPGSAGPAGNTGDRGDTGEKGAMGEKGEKGDEGPQGPSGGDGANGQNGGDGPPGDTG
DGGPQGAPGGPGPTGPSGPPGATGPPGPKGTDGPPGATGPQGSVGNPGPPGPPGATGQPGQKGERGATGDGGEAGAPGPAGDPGAQGDRGLDGPSGLTGP
QGPQGPQGSRGAPGPEGRPGNAGPQGEIGPAGPQGIPGMDGTNGVVGPDGEKGEKGADGPAGPPGLQGPPGPNGIPGRIGPAGPQGQPGPPGPGGPMGPP
GSFGGQGSGGADGAKGDDGPPGPTGPPGDRGVDGNPGIKGATGSQGGPGQPGPPGPVGRQGPPGSPGSPGDSGDDGPKGDTGPPGQAGQPGPIGPQGPPG
ANGAKGSKGDRGSDGNTGPAGATGDSGVIGPPGPPGPQGPRGADGGQGADGKAGSMGQMGLKGERGSPGPPGVPGPSGPRGDPGSDGIPGQDGANGPPGD
AGPIGAPGAPGIPGTSGAPGLPGPPGPRGSPGASGRAGAQGQRGGDGEKGERGVNGATGQQGAKGEIGATGVKGDRGGFGLPGRQGQSGSVGPPGPPGPP
*GPSGDQGTRGSPGPAGPRGDPGSQGPQGTSGPEGGRGRDGFDGDKGHPGPRGSPGPPGPPGYC*DGCTTLADPKKHFNSAHGGGEEDGPEWYDASRRYCDG
TEDKPCPTCRDVFLNTDSAKTGKYFIDPNFGSPSDKVEVVCNFNVQMVNNVPEQHYRTCVNVTDVSQVSEVWNVKYKMQDDQMWSLMRLSSKVEQKMSVA
CNDMPETVGCSIEDKSTIQPSCVSGQEI

***Figure 25**: amino acid sequence of p-pP4. The signal peptide is highlighted in yellow, the minor triple helix is in bold light blue, the major triple helix is in bold black italics and the C-terminal fibrillar domain is in bold green; potential N-glycosylation sequons are in bold red and underlined, while potential O-glycosylation sites are in bold dark blue and underlined as well. The RGD motifs are in bold pink and underlined.*

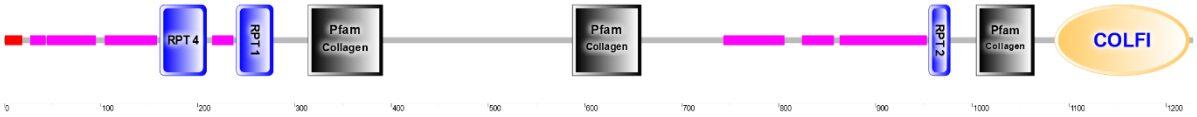


***Figure 26**: predicted domains of p-pP4 (image obtained with http://smart.embl-heidelberg.de/).*

3.2.2.5 Gene 5

G5 is located on chromosome 2 in a head-head position with G4 (genomic coordinates of the ORF: Chr2:1003795-1011325, with *plus* orientation) and measures approximately 8 kb. As for G4, the transcript deriving from G5 is incomplete in the 5' region in the results of our transcriptomic analysis; gene prediction also gave an incomplete sequence in which wasn't possible to find the start codon. The intron-exon assessment was made following the Augustus tool prediction; considering that the first exons are lacking, the incomplete ORF results being composed by 22 exons and 21 introns. The longest exon has 324 bp, and the shortest one 45 bp. The largest and the smallest introns measure 1049 and 52 bp, respectively. The 3'-UTR region derives from one 271 bp long exon. The structure of G5, however incomplete, is shown in Figure 27.

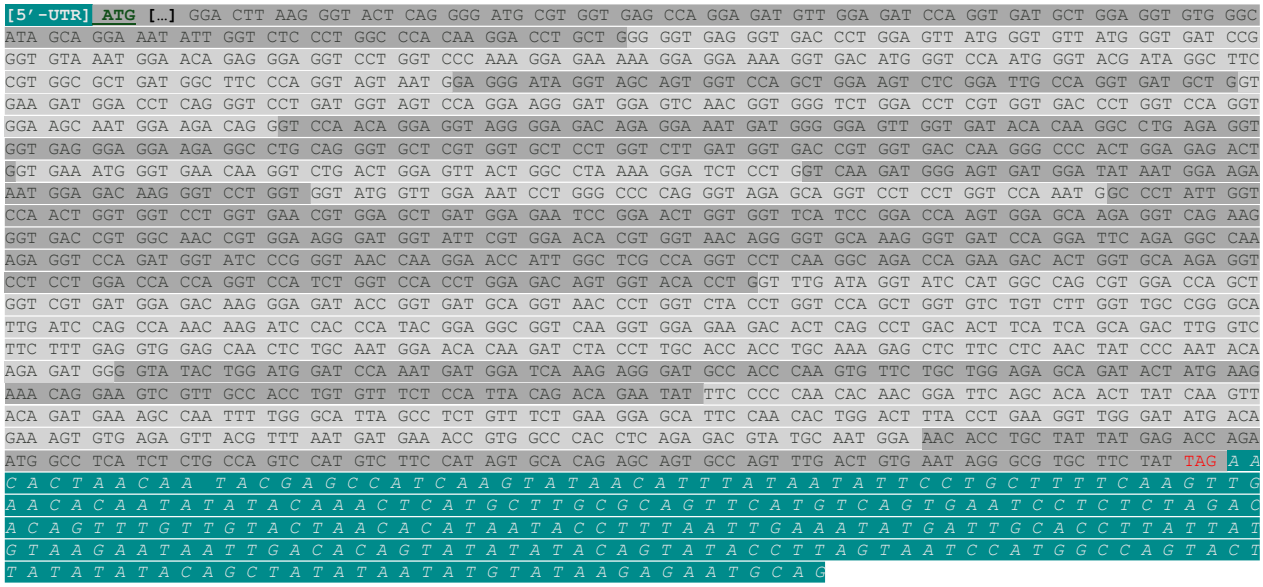


***Figure 27**: G5 introns-exons assessment. The introns and the exons of the coding sequence are highlighted in light grey and dark grey, respectively. The introns and the exons of the 3'-UTR are highlighted in light blue and dark blue, respectively. The starting codon is in bold underlined dark green, while the stop codon is in bold red.*

3.2.2.5.1 Putative protein from G5 – structure and features

In our analysis, the putative pre-proprotein derived from G5 (p-pP5) lacks the N-terminal domain; therefore, the results refer to an incomplete protein. Lacking the N-terminal region, it was not possible to precisely calculate the molecular weight and the isoelectric point of the total p-pP, and neither a signal peptide nor a minor N-terminal triple helix was individuated; however, a major triple helix of 843 AA and a COLFI C-terminal domain (e-value: 1.96e-7) of 184 AA were found. The amino acid composition of the major triple helix of p-pP5 is reported I Figure 31B. The prediction system individuated 3 potential N-glycosylation sequons and 14 potential O-glycosylation sites with the set threshold value of 0.8. Figure 28, Figure 29, and Figure 30 show the structure of p-pP5 referred to its corresponding gene (G5), the amino acid sequence of p-pP5, and the predicted domains of p-pP5, respectively.

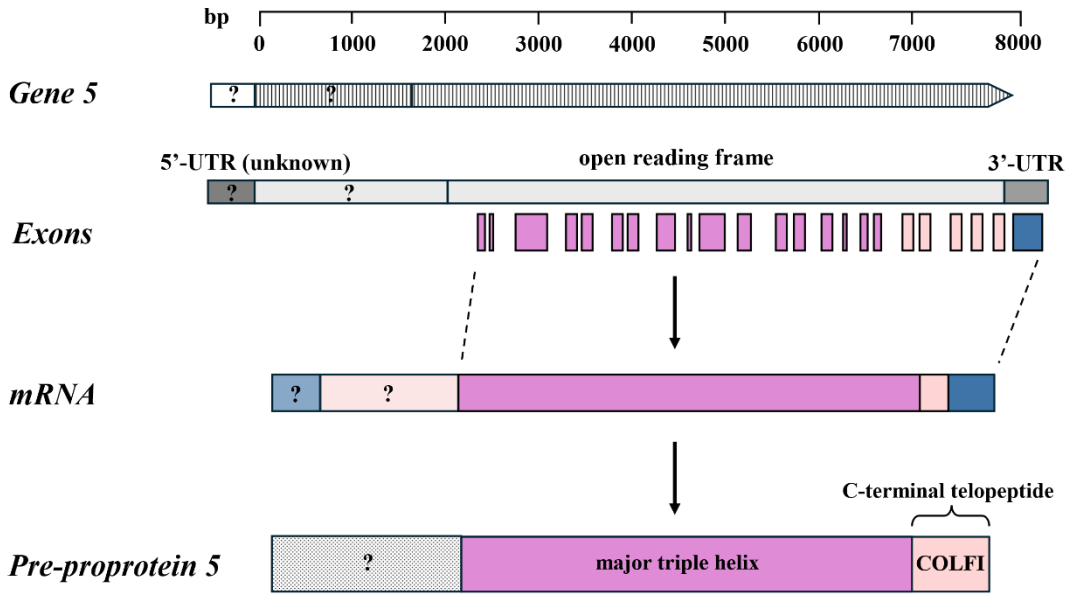


***Figure 28**: the path from gene 5 to p-pP5. The exons coding for triple helices domains are coloured in dark violet, as well as the triple helices domains in the pre-proprotein; the UTRs are coloured in blue.*

M[…]L*GERGDPGRQGNPGAPGSAGSSGVGGSGGVEGPQGEKGVAGASGVKGASGDVGAPGASGEPGLPGRQGTDGPPGPPGPPGNVGAEGKQGSPGRDGS PGNPGAFGENGVDGRPGRIGDQGKPGQKGEKGPIGSQGARGIRGLTGFNGAVGKPGVGGSQGDRGENGLRGAPGAIGGRGIRGFDGEPGNPGAVGDPGSN GEHGSPGKHGDSGPNGLSGIAGAAGVSGLTGATGAQGIKGVKGDIGSTGAQGDQGEKGMRGNAGAPGGDGGTGTAGTKGNKGAKGVAGSEGLQGAAGSRG LAGGPGITGSQGLPGTGGDRGPTGPAGPQGADGSTGLAGPQGITGPKGQTGEPGLDGKDGRPGDEGPEGEKGIRGDIGAAGLAGGQGPAGNPGESGPEGR IGLQGARGARGERGDPGPLGAQGPNGTDGLMGLKGTQGMRGEPGDVGDPGDAGGVGIAGNIGLPGPQGPAGGEGDPGVMGVMGDPGVNGTEGGPGPKGEK GGKGDMGPMGTIGFRGADGFPGSNGGIGSSGPAGSLGLPGDAGEDGPQGPDGSPGRDGVNGGSGPRGDPGPGGSNGRQGPTGGRGDRGNDGGVGDPGDDG VPGGPGEKGVPGSPGTQGLRGGEGGRGLQGARGAPGLDGDRGDQGPTGETGEMGEQGLTGVTGLKGSPGQDGSDGYNGRNGDKGPGGMVGNPGPQGRAGP PGPNGPIGPTGGPGERGADGESGTGGSSGPSGARGQKGDRGNRGRDGIRGTRGNRGAKGDPGFRGQRGPDGIPGNQGTIGSPGPQGRPGDTGARGPPGPP GPSGPPGDSGTPGLIGIHGQRGPAGRDGDKGDTGDAGNPGLPGPAGVCLGCRALIQPNKIHPYGGGQGGEDTQPDTSSADLVFFEVEQLCNGTQDLPCTT CKELFLNYPNTRDGVYWMDPNDGSKRDATQVFCWRADTMKKQEVVATCVSPLQTEYFPQHNGFSTTYQVTDESQFWALASVSEGAFQHWTLPEGWDMTES VRVTFNDETVAHLRDVCNGNTCYYETRMASSLPVHVFHSAQSSASLTVNRACFY*

***Figure 29**: amino acid sequence of p-pP5. The major triple helix is in bold black italics and the C-terminal fibrillar domain is in bold green; potential N-glycosylation sequons are in bold red and underlined, while potential O-glycosylation sites are in bold dark blue and underlined as well. The RGD motifs are in bold pink and underlined.*

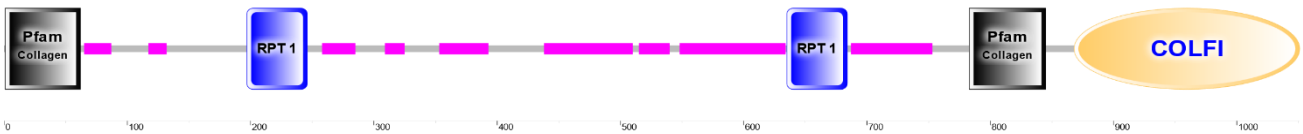


Figure 30: predicted domains of p-pP5 (image obtained with http://smart.embl-heidelberg.de/).

3.2.3 Amino acid composition of the pre-proproteins

Figure 31 reports the content of each amino acid (expressed as percentage of residues per amino acid sequence) within the translation products of G1, G2, G3, G4, and G5, both as pre-proproteins (Figure 31A) and as mature proteins, where the telopeptides will be cleavaged during post-translational processes (Figure 31B); in the latter case, only the major triple helix was considered. For p-pP5, only the triple helix was considered, as the N-terminal domain is completely missing. The

high amount of glycine confirms the collagenous nature of the p-pPs. Despite having the shortest minor triple helix, p-pP4 has the highest content of glycine and proline when comparing the amino acid sequences complete of the telopeptides (see Figure 31A). Considering only the triple helix (see Figure 31B), p-pP1 has the highest content of proline and threonine. P-pP2 and p-pP4 are the most similar between each other, even if p-pP2 has a lower content of alanine and proline and a higher amount of arginine, glutamic acid, and valine. P-pP3 is the one the most differs from the others, because it shows a considerably higher number of residues of lysine, leucine, histidine, and valine, while it has the lowest content of alanine, asparagine, aspartic acid, glutamine, serine, and threonine. Finally, p-pP5 has the highest amount of glycine and aspartic acid, and the lowest content of proline.

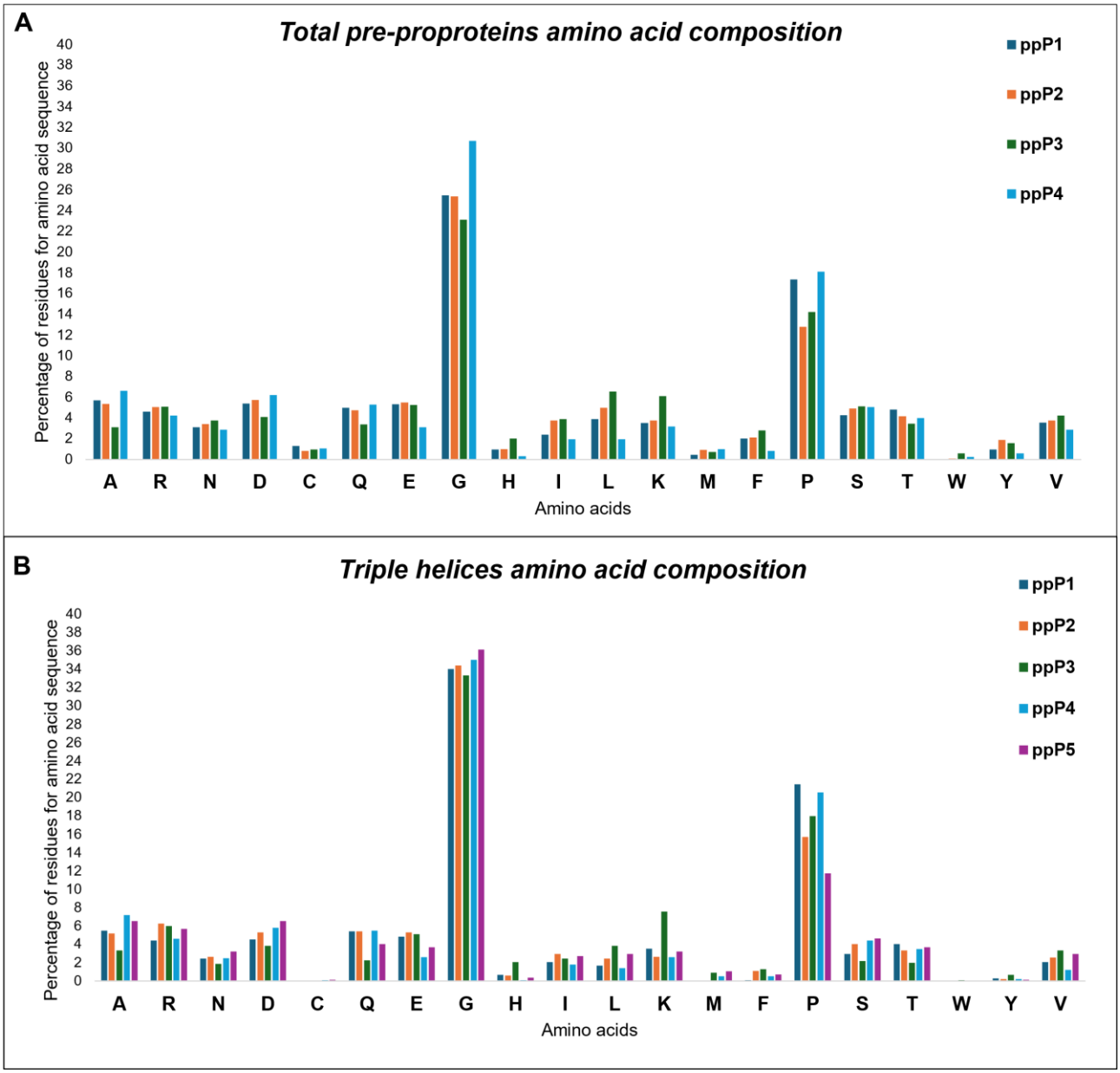


***Figure 31**: amino acid content expressed as percentage (%) of residues for each translation product, considering the whole pre-proprotein (where the telopeptides are still present) and only the triple helix (that is considered the main component of the mature protein).*

3.2.4 Comparison with the C-terminal domain of other fibrillar collagens

The putative sequence of several fibrillar collagens from both vertebrates and invertebrates were aligned with the five putative sequences found for *Chondrosia reniformis*. A special attention was reserved for the C-terminal domains, as it is representative of this molecule and one of its most conserved portions. The COLFI domain of the five p-pPs is, in fact, similar to that of other organisms, including *Homo* sapiens; in particular, considering human fibrillar collagens COLFI domains, p-pP1 has a 24% identity with the alpha-1 chain of the type I fibrillar collagen, p-pP2 and p-pP5 share the most identity with alpha-1 and alpha-3 chains of type V collagen (21.8% and 20% identity, respectively), while p-pP3 and p-pP4 are more similar to the alpha-1 chain of type XXIV collagen (19% of identity). The results of the complete COLFI domains alignment with also other invertebrates COLFI domains are reported as a supplementary file (S1).

3.2.5 Potential regulatory sequences in the genes' upstream regions

Considering the 2000 bp (or 724 bp for G1 and G2) upstream regions of the genes, the 5'-CAGAC-3' motif and several GC-rich motifs were searched as a further validation of the TGF-related pathway previously identified in the regeneration mechanism of *Chondrosia reniformis*. Notably, the 5'-CAGAC-3' motif was individuated in the upstream region of G1 (in which is present singularly), G3 (where is repeated three times) and G4 (in which is repeated two times). The motifs 5'-GTCT-3' and 5'-GGCTG-3' have been individuated in the upstream region of each gene. The motif 5'-GGCGC-3' has been found only for G1 and G2, while 5'-GGCGG-3' was found twice for G1, once for G4 and thrice for G5. The 5'- GGCCGC-3' has been identified only for G2 two times, while 5'-GGCCGCG-3' was not found in any of the upstream regions. The upstream regions with the highlighted potential Smad-binding sequences are reported in the Supplementary File S2.

3.2.6 Gene expression analysis

The fibrillar collagen genes expression distribution within the exopinacoderm and the mesohyl, as well as the relative gene expression of G1, G2, G3, G4, and G5, was assessed. Regarding tissue distribution, all experiments show that all genes are significantly more expressed in the internal part of the sponge, except for G3, which expression is considerably higher in the cortical region (see Figure 32; note that the different graphs should not be compared between each other); this assessment is constant in all replicates, even the level of expression may change due to variable lophocytes concentration in different animals. In particular, G1, G2, G4 and G5 are always respectively at least 4, 2.5, 4.4 and 9 times more expressed in the mesohyl, while G3 is always between 8 to 169 times more expressed in the cortex. Figure 32 shows a representative result of one experiment, where SD derive from the values of the triplicates of one single qPCR analysis.

The relative gene expression of G1, G2, G3, G4, and G5, considering G1 as calibrator sample, is reported in Figure 33. G1 was chosen as a calibrator arbitrarily. In the cortical part of the sponge, as well as in the choanosome, the most abundant transcripts derive from G1 and G4, while the transcripts deriving from G2 and G3 are the less represented (see Figure 33A). mRNA copies of G1, G2, G4, and G5 are distributed quite evenly among both cortex and mesohyle; instead, a higher number of mRNA copies of G3 can be found in the cortex, while the same transcript is basically absent within the choanosome (see Figure 33B).

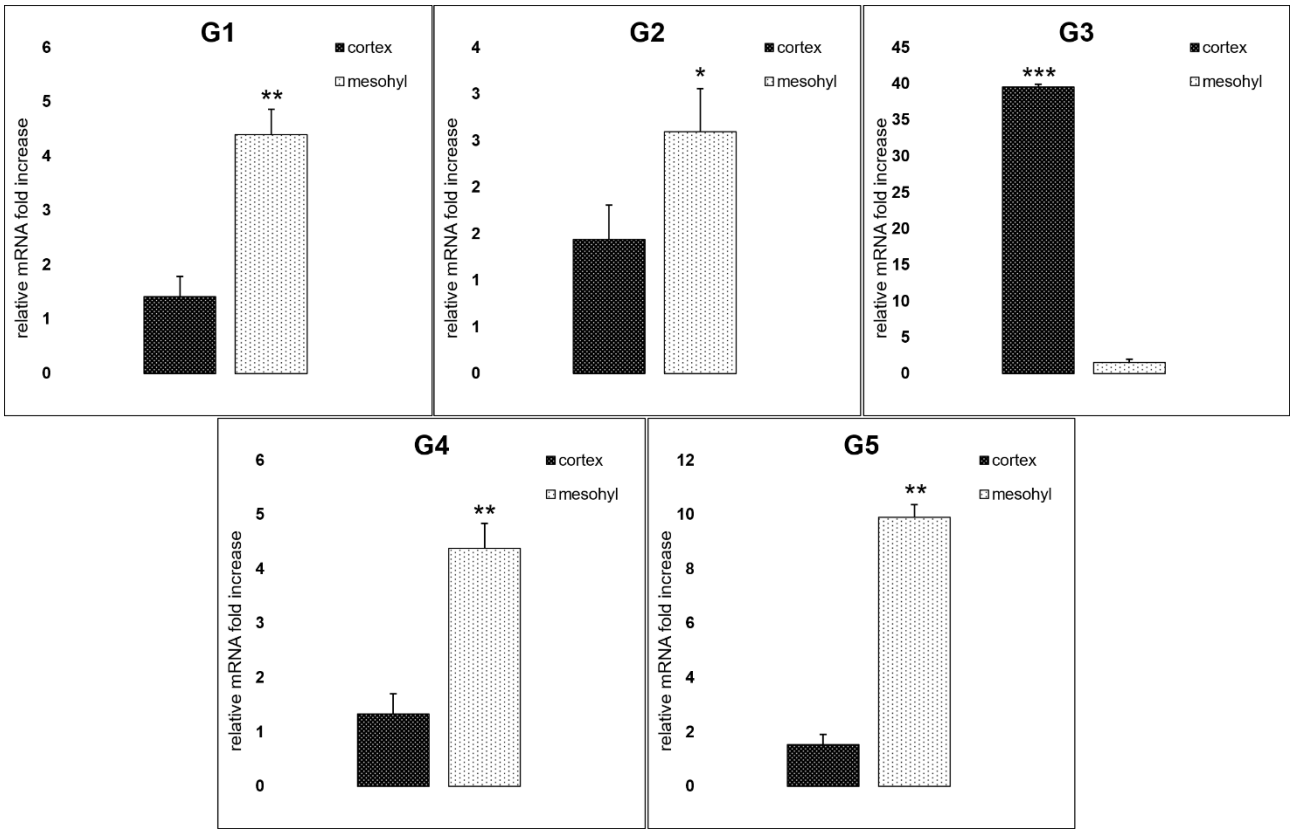


***Figure 32**: gene expression distribution of the different fibrillar collagen genes in the two main region of the body of* C. reniformis*, i.e. cortex (represented in black) and mesohyle (represented in light grey). The five different graphs should not be compared between each other. *= p-value<0.5; **= p-value<0.01; ***= p-value<0.001*

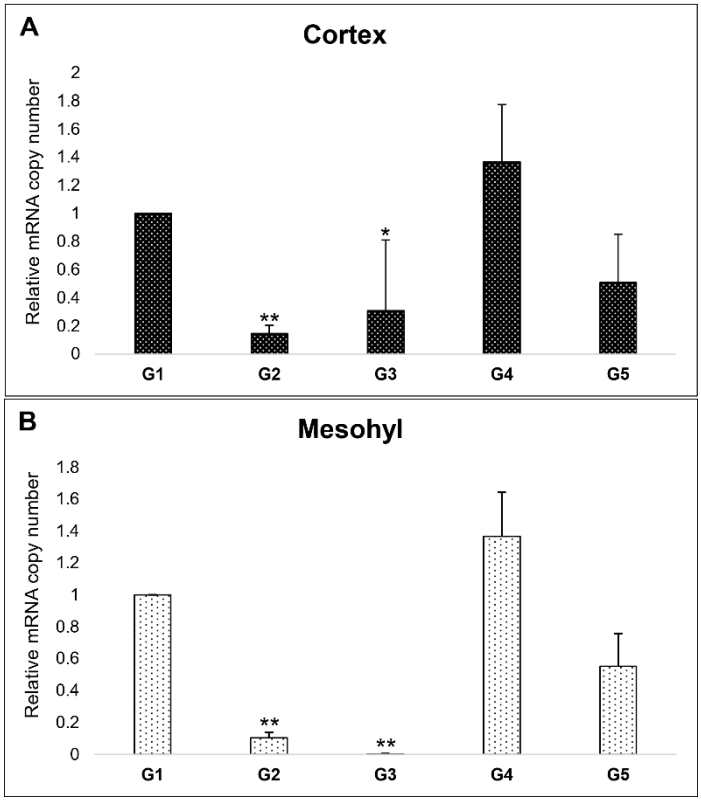


***Figure 33**: relative expression of fibrillar collagen gene in the cortical region (A) and in the mesohyle (B). *=p-value<0.5; **= p-value<0.01.*

3.3 Discussion

The identification and description of genes and proteins from non-model species pose significant challenges due to the reliance of mainstream bioinformatics tools and databases on data primarily derived from well-studied model organisms. These tools and databases are often calibrated and optimized for model species, such as *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, and *Arabidopsis thaliana*, among others. As a result, when studying non-model organisms, it is frequent to encounter limitations in the availability of reference genomes, annotated gene sequences, and comprehensive protein databases tailored specifically to the species of interest. The sort of bias towards model organisms can hinder, for instance, accurate annotation, functional prediction, and comparative analysis of genes and proteins in non-model species [Dowd, 2012]. By keeping this in consideration, this part of the doctoral research project aimed to individuate and characterize the genes coding for fibrillar collagen in the marine demosponge *Chondrosia reniformis*, because of the biotechnological interest which the collagen of this animal is known for. In total, five genes coding for fibrillar collagens were identified within *Chondrosia reniformis*' genome, which is publicly consultable on NCBI. In particular, two of the five fibrillar collagen genes – the ones that were named Gene 4 and Gene 5 – are located on Chromosome 2 of this species, while the other three – Gene 1, Gene 2, and Gene 3 – were found on Chromosome 13. G1, G3, and G4 are in minus position, i.e. their sequence has to be read in the opposite direction respect to G2 and G5; notably, G1 and G2 are in a "head-to-head" orientation and are really close to each other, as their coding sequences are separated by only 774 bp. The genomic locus of a gene pair is considered ''head-to-head'', or ''bidirectional'', when the two adjacent genes are divergently transcribed from opposite strands of DNA, consequently bringing the region between the two transcription start sites (TTSs) to be designated as a putative bidirectional promoter [Li et al., 2006]. Regarding collagens, similar situations can be observed also in Vertebrates, for example in the case of COL4A1 and COL4A2 genes in the murine genome, which are also in a head-to-head position and are distant 874 bp [Kaytes et al., 1988]. According to literature, this kind of gene arrangement does not occur randomly; instead, is an ancient and conserved type of organization, often being present even in higher Vertebrates for several genes. Head-to-head genes appear to be functionally related and subjected to a coordinated transcriptional regulation, offering an example of gene expression regulation based on genomic organization. The expression of bidirectional genes has been proven to be significantly correlated, where the correlation can be positive (more frequently), negative, or alternative depending on the experimental conditions [Li et al., 2006]. Therefore, considering the statistically more probable scenario, in *C. reniformis* G1 and G2 should be linked for what concerns their function and patter of expression; considering that their relative expression is maintained constant both in the exopinacoderm and in the critical region of the animal, it is possible to hypothesize that they could encode for two different α-chains constituting a heterotrimeric collagen fibril. The presence of heterotrimeric molecules of collagen could be suggested also by the relative abundance of the transcripts, where G1, G4, and G5 appear more transcripted than the others: this could imply, for instance, the occurrence of collagen molecules constituted by two alpha-helices deriving from the former genes, and one deriving from the less represented ones.

Once individuated all the fibrillar collagen genes within the genome of *Chondrosia reniformis*, the next step was to understand how many of them are expressed in the adult animal and to assess the intron-exon pattern of each gene; therefore, a transcriptome analysis was performed on a pool of almost 30 different animals. The transcriptome analysis yielded promising results, as the 89.31% of the transcripts could be traced back to the ortholog core genes present in the Metazoan genomes, and it was possible to individuate a transcript corresponding to each one of the previously identified genes,

confirming that there are not exclusively larval fibrillar collagens in this sponge. However, two of the five fibrillar collagen transcripts that were identified in the transcriptome were incomplete in their N-terminal region, and all of them lacked the 5'-UTR portions. According to our observations, even before this research project, this situation is not uncommon for what concerns fibrillar collagens; in fact, fibril forming collagen transcripts missing their N-terminal portions can be found in the supplementary material of a paper published by Riesgo and co-authors [Riesgo et al., 2014], for the marine sponge species *Petrosia ficiformis* and *Chondrilla nucula*. Similarly, Pozzolini and co-authors [Pozzolini et al., 2016] published a partial mRNA sequence of a fibrillar collagen obtained from a former transcriptome analysis (indeed, transcripted from the one that was named as "Gene 2" in this work), also in this case lacking almost half the entire sequence in the N-terminal region. We hypothesize that this recurrent condition may be due to technical challenges posed by the intrinsic nature of fibrillar collagen sequences themselves; in fact, it is known that RNA transcripts can display significant secondary structures which may hinder the action of the reverse transcriptase during cDNA production [Bustin, 2000]. Collagen transcripts contain repeated motifs especially at the level of the triple-helix coding sequence, which may form loops during the process of retrotranscription, hence somehow blocking the activity of reverse transcriptase, which is therefore prevented from arriving until the end of the mRNA sequence. The intron-exon assessment was made by pairing the transcripts' sequences to the coding sequences of their corresponding gene. In the case of G1, G2, and G3, it was possible to align the whole actual transcript to the gene, as the transcript was almost complete, except for the 5'-UTR; for G4 and G5, the alignment between the actual transcript and the gene was possible only partially, as the transcripts lacked, besides the 5'-UTR, also part of the N-terminal region; for this reason, the intron-exon assessment of that part was made following the *in silico* prediction by the Augustus software. Overall, with a few exceptions, it can be observed the classical "cassette" organization typical of the fibrillar collagen genes of higher organisms, where the exons that code for the triple helix have constant dimensions of 54 bp, 54-9 bp, or multiple of these numbers. This confirms the fact that the genetic structure of fibrillar collagens is highly conservated in Metazoans, as it occurs from poriferans to humans. Many studies have in fact demonstrated that fibrillar collagen genes have evolved from an ancestral multi-exon unit, and that once the gene was translated, a strong evolutionary pressure caused it to maintain this structure. It must be highlighted that G1, G4, and G5 have a low number of exons in the region encoding for the triple helix, but with larger dimensions in terms of base pairs; this suggests the probable occurring of an exon fusion phenomenon, like previously reported in other cases [Chu et al., 1984; Exposito et al., 1993].

A combined approach using both prediction systems and transcripts alignment on the genes made possible to precisely individuate the starting position of the coding sequence for all the genes, apart from G5; the reasons behind the lacking starting position of the CDS of G5 are still to be clarified. Knowing the starting position of the CDSs, an attempt was made to research for potential regulatory sites in the upstream region of the genes. To reconstruct and validate a complete gene regulation mechanism is a harsh challenge. In higher organisms, one of the molecular pathways involved in the regulation of fibrillar collagens expression is the TGFβ (transforming growth factor β) – Smad pathway [Verrecchia and Mauviel, 2007]. In short, TGFβ is a cytokine that, under certain conditions, binds to its cell receptors, known as the TGF-β receptors (TGFBRs). This binding determines the formation of a Smad complex which is transferred into the cell nucleus; here, it interacts with specific DNA sequences called Smad-binding elements (SBEs) within the promoters of target genes, including those encoding fibrillar collagens [Verrecchia and Mauviel, 2007]. In *Chondrosia reniformis* all the main actors of this pathway have been individuated *in silico*, in particular several members of the TGF superfamily, TGF-receptors [Pozzolini et al., 2018], and Smad proteins (unpublished observations). Considering that this molecular pathway in evolutionarily conserved and considering

that the administration of a TGF inhibitor blocked the restoration of the exopinacoderm layer [Pozzolini et al., 2018], the presence of SBEs in the upstream region of every individuated fibrillar collagen gene could be further evidence of the TGF-Smad system in this sponge; however, this should be experimentally validated.

The putative proteins coming from the translation of the transcripts were also analyzed. All putative proteins have similar dimensions, averagely of about 1500 amino acids, coherently with other Vertebrates and Invertebrates fibrillar collagens pre-proproteins consultable on NCBI (National Center for Biotechnology Information (nih.gov), last accessed March 30, 2024). The N-terminal region contains a minor triple helix in p-pP1, p-pP2, and p-pP4 (see paragraph 3.2.2), while no typical fibrillar collagen N-terminal domains have been found in p-pP3; as the N-terminal telopeptides are crucial to a normal fibril assembly and control of fibril shape [Hulmes, 2002; Malone et al., 2004], for p-pP3 it is possible to hypothesize the presence of other motifs that were not recognized by the tools used in this work, while for p-pP5 it is not possible to tell. Moreover, no thrombospondin, nor von Willebrand domains were observed. These kinds of telopeptide-related domains, involved as well in the assembling of collagen fibrils, have been found in all classes of Porifera [Tucker and Adams, 2024]; therefore, the reason behind their absence in *Chondrosia reniformis* is not ascribable to a further occurrence in the evolutionary path of poriferans, but it could represent a singular adaptation of this species, even if its implications and causes are still to be clarified. On the other hand, all putative proteins were provided of the C-terminal globular propeptide specific for fibrillar collagens, which is responsible for fundamental events of fibrillogenesis, holding sites necessary for the formation of cross-links, bonds with regulatory proteins and other post-translational modifications [Wess, 2005]. The C-terminal domain of fibrillar collagens appears to be very conserved among all the animal taxa. In particular, as reported in Figure, cysteine residues are present in almost all selected sequences, confirming their role in promoting the formation of disulfide bridges that will link the single pro-α chains at the level of the C-terminus during the formation of collagen molecules. Only p-pP4, among *C. reniformis*' sequences, appears slightly different in the C-terminal portion: its cysteine residues, while present, are in different positions and, in fact, the e-value of this domain is higher respect to that of the other proteins. However, the reasons behind this modification are not clear. Finally, in all putative proteins have been found RGD motifs; *in vivo*, the RGD motif serves as an important recognition site in various fibrillar collagens, such as types I, II, and III. Integrin receptors, notably integrin α5β1, recognize this motif: the binding of integrin α5β1 to the RGD motif can initiate intracellular signaling pathways, leading to epithelial-mesenchymal transition phenomena typical of the wound healing process [Yuan et al., 2019; Marconi et al., 2021]; therefore, this could be a positive property in the case of biotechnological applications of *C. reniformis'* collagen.

Considering the whole pre-proprotein, potential N-glycosylations and O-glycosylations sites were predicted; in particular, potential O-glycosylations sites are present in higher amounts. More specifically, potential N-glycosylations sites were individuated only at the level of the telopeptides, while potential O-glycosylations sites were present in every part of the pre-proprotein, but especially on the major triple helix region. Hence, considering only the triple helix as the main part of the future mature protein, only potential O-glycosylations sites can be individuated. In percentage, the most potentially glycosylated putative protein is p-pP4, with a value of 1.71%, while the less potentially glycosylated is p-pP3, with a value of 0.65%. Considering that O-glycosylations occurs on tyrosine and serine [Hema Thanka Christlet and Veluraja, 2001], it is possible to calculate the percentage of Tyr and Ser residues potentially glycosylated in the triple helix: in this case, p-pP1 has the highest potentially glycosylated Tyr+Ser/total Tyr +Ser *ratio*, with a value of 54.5%, while the lowest is again p-pP3, with 34.5%. Usually, at least in vertebrates' collagens, glycosylations mostly occur on

hydroxylysine residues [Boot-Handford and Tuckwell, 2003]; as we are speaking about the primary sequence of putative proteins, no information can be given about the amount of post-translational modifications which those proteins will be subjected *in vivo*. Hower, less common O-glycosylations on tyrosine and serine residues have been reported in the case of the vestimentiferan *Riftia pachyptila* [Mann et al., 1996] and predicted in the byssus [Engel, 1997]; in *R. pachyptila*, O-glycosylations have a dominant role in stabilizing the triple helix if compared to 4-Hyp. Both byssus and *R. pachyptila* fibrillar collagen are known for their stability; the prediction of this kind of glycosylations in *C. reniformis* fibrillar collagen could represent a possible explanation as well to its thermal resistance respect to other types of collagens, even considered that the low amount of Hyp in this sponge cannot be the only factor involved [Tassara et al., 2023].

In the mature protein, the telopeptides will be completely or almost completely cleavaged, thus having a small impact in the overall amino acid composition of the final product; therefore, to reason about the potential chemical-physical features of the putative protein, only the triple helix amino acid composition was considered. All the putative proteins contain an amount of glycine comparable to that of other fibrillar collagens from both Vertebrates and Invertebrates [Plez and Gross, 1959; Van B. Robertson, 1964; Kimura and Matsuura, 1974;], ranging from 33.3% (p-pP3) to 36.14% (p-pP5). The percentage of total proline is more variable, as it ranges from 14.1% in p-pP5 to 21.44% in p-pP1; only p-pP1 has a total proline content similar to that of mammalian collagen, while the other ones result comparable to fish collagen and marine invertebrates' collagens, like those from squids and crabs [Kimura and Matsuura, 1974; Simpson, 2012]. For what concerns lysine, in all putative proteins the percentages are quite lower if compared to vertebrates, but analogous to that of other marine invertebrates, and revolves around 3% [Kimura and Matsuura, 1974; Simpson, 2012]; only p-pP3 sticks out for a lysine content which is more than double the others, with a value of 7.5%. The amount of lysine content, however, is not the only factor which distinguish p-pP3 from the other p-pPs: first of all, while the other p-pPs have similar amounts of each amino acid, this protein has the tendency to be the one having the highest or the lowest amount of several amino acids; in particular, it has the lowest content of alanine, asparagine, aspartic acid, glutamine, serine, and threonine, while contains the highest amount (besides of lysine) of histidine, leucine, phenylalanine, tyrosine, and valine of them all. This is reflected in its isoelectric point: if the average predicted isoelectric point of the other p-pPs is 5, p-pP3 has a predicted value of 9.24 considering all the pre-proprotein and of 10.04 considering only the triple helix. This higher value depends on the presence of many residues of basic amino acids, especially histidine and lysine, and of few residues of some acidic amino acids, like aspartic acid. However, it is not possible to discuss about the impact on solubility of *C. reniformis* collagenous extract (which is particularly insoluble respect to more classical collagens, as explained in the introduction chapter), as we don't know in what *ratios* the different putative proteins assemble with each other to form fibrils, nor which ones, and how much of them, persist in the extract. The most interesting feature of p-pP3, however, is the fact that it contains four imperfections in the triple helix, where a fourth residue (of alanine ore serine) is added to the canonical Gly-Xaa-Yaa triplet, thus creating a Gly-Xaa-Yaa-Zaa quadruplet. Imperfections in the triple helix of a fibrillar collagen have already been documented in marine invertebrates, like sponges, worms, and abalones [Exposito et al., 2002]; it is thought that the presence of these imperfections in the triple helix does not affect molecular stability and fibrillar integrity, also considered the lower temperature at which marine animals live with respect to terrestrial animals, and – in the case of invertebrates – the higher amount of glycosylations with respect to vertebrates. The gene coding for p-pP3 has, moreover, a very different expression pattern if compared to the others. In *C. reniformis*, all fibrillar collagens display a similar expression profile, where they are more present in the mesohyle, besides G3, which is basically absent in the internal part of the sponge, but way more expressed than the others in the

cortical zone, almost seeming specific for this part of the sponge. These data, taken together with the fact that p-pP3 shows imperfections in the triple helix, seem to be particularly significant in the case of *C. reniformis*: in fact, considering that triple helix imperfections can improve the flexibility of the molecule [Exposito et al., 2002], p-pP3 could play a key role in the creeping phenomenon typical of this species, which takes place starting from the exopinacoderm of the animal. Nonetheless, considering the information obtained to this point, it is still not possible to confirm this hypothesis, which must be proven experimentally.

Chapter IV - FACTORS INFLUENCING COLLAGEN PRODUCTION IN *Chondrosia reniformis* (Nardo, 1847)

Reaching a sustainable production of *C. reniformis* able to meet market demands is an ambitious objective, as the growth of this animal is slow and, nevertheless of the efforts put in practice by many authors, aquaculture implants able to efficiently sustain this purpose still must be optimized. For this reason, this part of the research activity aimed to find a way to artificially improve the collagen biosynthesis in this sponge, so to get to a higher collagen yield in a shorter span of time, also possibly speeding up the regeneration process which normally occur in this sponge when it is subjected to a cut. To fulfil this goal, two different treatments were tested on *C. reniformis*; a first attempt was made testing the application of diatomaceous earth (provided by the Antarctica Museum of the Genoa University) on the sponge's body surface after a cut: this idea emerged because, during spring, a thin layer of diatoms was observed on the surface of many *C. reniformis* specimens in shallow waters by the Gallinara Island (Savona, Liguria, Italy) (personal unpublished observations). As the phenomenon appeared not accidental to the observer's eye, from this came the idea that the interaction between diatoms and the sponge could have some mutualistic significance for the two organisms. In fact, diatoms are characterized by structured cell walls made of nanopatterned silica (SiO_2) [Kröger and Poulsen, 2008], and it has previously been observed that silica induces collagen production in *C. reniformis* [Pozzolini et al., 2021; Pozzolini et al., 2015]. However, the experimental design wasn't effective, and no good results were obtained. A second attempt foresaw the irradiation of the animals with near-infrared 810 nm light. This brought to the publication of a paper ("**Near-Infrared 810 nm Light Affects Porifera *Chondrosia reniformis* (Nardo, 1847) Regeneration: Molecular Implications and Evolutionary Considerations of Photobiomodulation–Animal Cell Interaction**", Amaroli et al., 2022, https://doi.org/10.3390/ijms24010226), which essential parts are reported below (for complete references of this section, see the original work).

EFFECT OF PHOTOBIOMODULATION ON THE REGENERATION PROCESS IN *CHONDROSIA RENIFORMIS*

The interaction of visible (VIS) and near-infrared (NIR) light with living organisms has been extensively documented, particularly in plant cells, where photons serve as a primary energy source and are converted into chemical energy through photosynthesis. In contrast, animal cell evolution has been characterized by a preference for chemotrophic sources of energy, with limited light reception mediated by highly specialized molecules such as photoreceptors. Photo-biomodulation (PBM), also known as low-level laser therapy, has been described as a non-thermal and non-ionizing interaction between light and animal cells. In animal cells, both specialized photoreceptors and non-specialized molecules (known as photoacceptors) can interact with photons from VIS and NIR wavelengths emitted by light sources such as LEDs, lasers, and broadband light. These photoacceptors, including mitochondrial cytochromes, haem-containing proteins, and flavins, have the potential to modulate cellular metabolism. Recent studies have highlighted the ability of various life forms to respond to PBM. Despite being investigated extensively in the medical field, the evolutionary implications of PBM in terms of light-organism interactions and cellular pathways have not been thoroughly

discussed, partly due to inconsistent data. To address this gap, previous studies have utilized a standardized experimental setup to investigate the effects of 810 nm NIR light, referred to as 810 nm-1W PBM, across a diverse range of organisms, and have explored effects on cellular/tissue homeostasis, energetic metabolism, cell proliferation, differentiation, and disease recovery. In this work, the effect of 810 nm-1 W PBM was tested for the first time on Porifera, the simplest and oldest surviving metazoan lineages, using *Chondrosia reniformis* as experimental model. The gene expression profile related to inflammatory processes, apoptosis, heat stress, growth factors, collagen production, oxidative stress enzyme activity, and total prokaryotic symbiont content in *C. reniformis* exposed to 810 nm-1 W PBM were investigated.

4.1 Materials and methods

Unless differently stated, all reagents were acquired from Sigma-Aldrich (Milan, Italy).

4.1.1 Sponge harvesting

Several specimens of *C. reniformis* were collected by the Mount Portofino "Marine Natural Regional Park" (Liguria, Italy) at a depth between 10 to 20 m. The sponges were maintained at a temperature of 15°C in the same seawater of the site of collection until their arrival in the laboratory, where a short-term stabilization was performed; more specifically, the animals were kept at 14°C in a 200 L aquarium containing natural seawater collected in the same Portofino area with a salinity of 37‰ and equipped with an aeration system.

4.1.2 Experimental design

Three different specimens of *C. reniformis* (SA, SB, and SC) of medium-small dimensions (diameter: 6-8 cm) were divided into four equal parts with a sterile scalpel; this operation produced 12 sponge fragments, namely SA1, SA2, SA3 and SA4; SB1, SB2, SB3, and SB4; SC1, SC2, SC3, and SC4. For each specimen, two fragments were designated for the treatment (i.e., six fragments in total), while the other two fragments for each specimen (again, six fragments in total) were used as control. The treatment-designed samples were irradiated with an 810-nm diode laser (GaAlAs) device (Garda Laser, 7024 Negrar, Verona, Italy) provided with FT-HP, which can irradiate a spot area with a consistent energy distribution regardless of distance. Following the methodology of previous works, an 810nm-810-nm – 1 W PBM therapy at 1 W of power in CW mode was administered to the chosen samples for 60 seconds within a circular spot of 1 cm^2. These parameters allowed the generation of a power density of 1W/cm2 and a fluency of 60 J/cm2 (energy administered: 60 J); the control-designed samples, instead, were exposed to the same 810-nm diode laser and FT-HP, but the device was set to irradiate 0 W and 0 J for 60 s (810 nm-0 W PBM). A 635 nm red-light pointer (negligible power, <0.5 mW) was used in both cases to visualize the exposed area and keep the experiment blinded. During irradiation, the fragments were temporarily posed on a Petri dish, which was placed on an absorbent mat to avoid beam reflection, while the FT-HP was fixed to a stand 0.5 cm away from the sample. For each sponge, one treatment-designed fragment was irradiated once, while the other one was treated every 24 hours for four days, for a total of five irradiations at 0, 24, 48, and 96 hours. 60 s after each treatment, all the fragments were transferred again to the aquaria. The accuracy of the irradiated laser parameters was ensured with a Pronto-250 power meter (Gentec Electro-Optics, Inc. G2E Quebec City, Canada). Adverse events due to a possible undesirable thermal effect were avoided by monitoring the irradiation with a FLIR ONE Pro-iOS thermal camera (dynamic range: −20°C/+400°C; resolution: 0.1°C; FLIR Systems, Inc. designs, Portland, OR, USA). 24 hours after the first irradiation, the fragments treated once (short-term treated samples) and their respective

controls were collected; the regeneration edge was excised with a sterile scalpel and processed for RNA extraction, and genomic/prokaryotic DNA assessment, and histology. 24 h after the last of the five treatments (i.e., 120 hours after the first irradiation) the long-term treated samples were collected as well with their respective controls, and their regeneration edge was excised for the same purposes (see Figure 34A). The same experimental setup was also used to evaluate possible oxidative stress induced by PBM (see Figure 34B), but in this case the regeneration edge was excised only 6 hours after the first irradiation, to be subjected to protein extraction and enzymatic assays to evaluate oxidative stress markers.

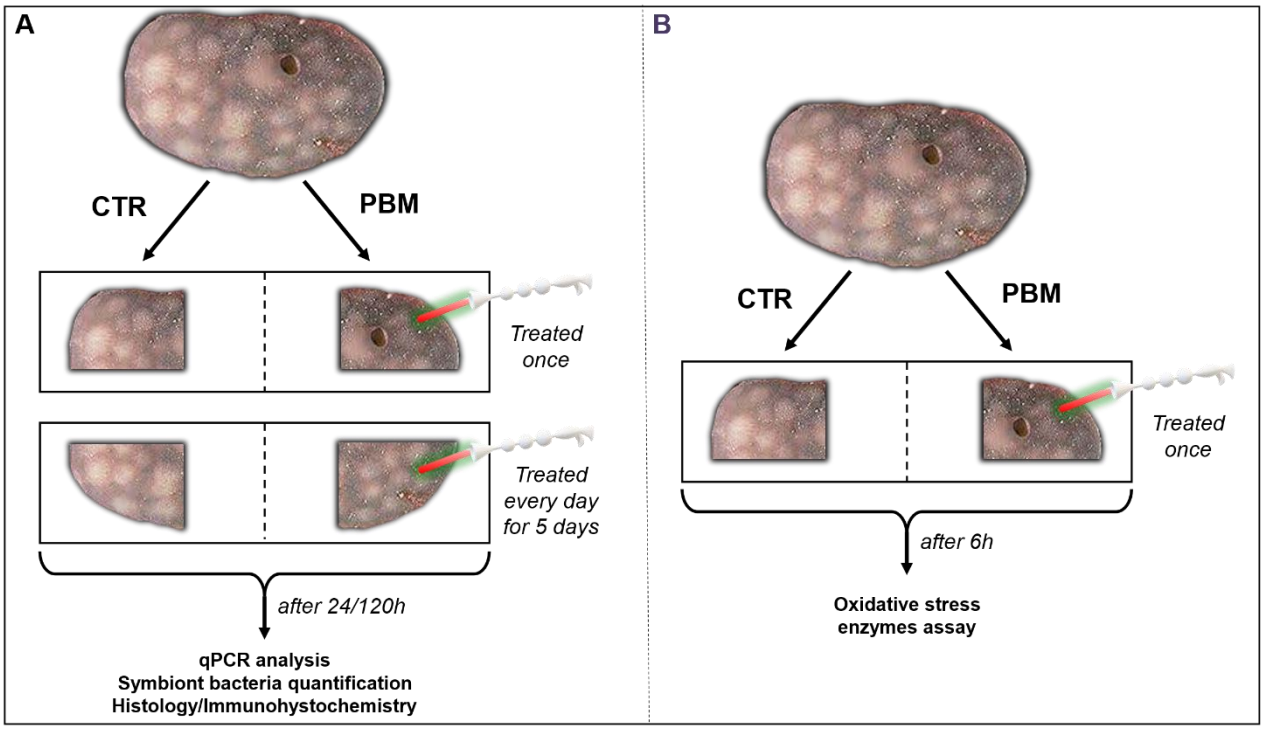


***Figure 34**: experimental design. Panel A shows the experimental design to obtain samples that, after 24 and/or 120 hours, will be used for qPCR analysis, symbiont bacteria quantification, histology, and immunohistochemistry; panel B shows the experimental design adopted to obtain samples for oxidative stress enzymes assays.*

4.1.3 Histology and Immunofluorescence

After the expected time, a fraction of the regenerating edge of both the control and PBM-treated fragments were embedded in paraffin, and cut transversally into 5 µm sections, used for histology analysis. The organization of the sponges' tissue was observed after Masson's trichrome staining. The presence of immunoreactivity for Msi-1, a marker of archaeocytes in sponge tissues, was evaluated by indirect immunofluorescence. Sections were incubated with polyclonal rabbit anti-Msi-1 antiserum (E-AB-17518-60; Elabscience, Houston, TX, USA), previously used on *C. reniformis*' tissues, and then with the secondary anti-rabbit Alexa Fluor 594 (Abcam, Cambridge, UK). Nuclei were counterstained with DAPI. Negative controls were performed by omitting the primary antiserum. Histological slides were then observed through a Leica DMRB light and epifluorescence microscope (Leica Microsystems, Wetzlar, Germany), provided with Moticam 3+ (Motic Europe, Barcelona, Spain). Finally, a semiquantitative analysis of Msi-1 immunoreactivity on histological slides of SA, SB, and SC was performed by capturing eight random photographs along the regenerating edges from two different slides for both the control and PBM-treated samples. The

percentage of red fluorescent pixels on tissue-occupied pixels in each photograph was evaluated by using the ImageJ software.

4.1.4 Gene Expression Evaluation

After the set time points, as explained in Figure 34, the regenerating edges of the sponge fragments were cut with a sterile scalpel, obtaining a thin layer of tissue about 3 mm thick that was promptly chopped into smaller pieces and put into an adequate volume of Isol-RNA Lysis Reagent (5 Prime GmbH, Hilden, Germany) to perform total RNA extraction. The poly-A fraction was then isolated with the FastTrack® MAG mRNA isolation kit (Life Technologies, Milan, Italy) following the manufacturer's instructions. Complementary DNA (cDNA) was synthesised by Revert Aid Reverse Transcriptase (Thermo Fisher Scientific, Milan, Italy), starting from 1 µg of purified RNA from each sample. Each PCR reaction was conducted in a final volume of 15 µL containing: 1× iQ SYBR®Green master mix (Bio-Rad, Milan, Italy), 0.2 µM of each primer, and 3 µL of a 1:5 dilution of cDNA. For each sample, the analysis was carried out in triplicate. The following thermal profile was used: initial denaturation at 95°C for 3 min, followed by 45 cycles with denaturation at 95°C for 15 s, and annealing and elongation at 60°C for 60 s. Fluorescence was measured at the end of each elongation step. The values were normalised to GAPDH expression (housekeeping gene). All the PCR primers (see Table) were designed using Beacon Designer 7.0 software (Premier Biosoft International, Palo Alto, CA, USA) and were provided by TibMolBiol (Genova, Italy). Data analyses were performed with the DNA Engine Opticon® 3 Real-Time Detection System Software program (3.03 version). To calculate the relative gene expression compared to an untreated (control) calibrator sample, the comparative threshold Ct method was used for gene expression analysis in the iCycler iQ Real-Time Detection System Software® (2004 Bio-Rad, Hercules, CA, USA). The data are presented as the mean ± standard deviation of three independent experiments performed in triplicate. Statistical analyses were performed using a one-way analysis of variance followed by Tukey's posthoc test (GraphPad Software, Inc., San Diego, CA, USA). A p-value < 0.05 was considered significant.

	Primer sequences	
Gene	*Forward*	*Reverse*
GAPDH	5'- AAGCCACCATCAAGAAGG-3'	5'-CCACCAGTTTCACAAAGC-3'
HSP60	5'-ACTATCATCTCGCCAATCG-3'	5'-TTATCCACTCCACTCAACATT-3'
HSP70	5'-TGTCAGTTCAGCAGTAGTG-3'	5'- TATTATCAACGAGCCAACAGC-3'
HSP90	5'-ACCGCATTCATCGCTTGG-3'	5'-GCTTCATCGTCACCTTCTTCC-3'
NOS	5'-CCACACTCTCCATAGCATTCT-3'	5'-CCGTCCTCSGCGTAACAC-3'
TNF	5′-AGAAATCGCCAGAAGCAAGTTG-3′	5′-GATGAGCACATTGATAGCAGACC-3'
Bcl-2	5'-ATATTCACCAAGCCAATGC-3'	5'-CCAGACACCACTCTATCG-3'
Msi-1	5'- ATTGGAATGCTTAGTGAAATGC-3'	5'- GTGTCTACCTTGCCTATTGC-3'
Wnt	5'-AGTGTGATGCTGTTAGTGA-3'	5'-TGCGACATTGACCTTCTATA-3'
ColF	5′-GTCATCCAGGTCGTCAAG-3′	5′-TAGGTCAATCTCTAATTCAGC-3′
FGF	5'-CCTTCGTATCTTAGAGAATGGA-3'	5'-ATGTTGCGGAACTTGAGA-3'
Tgf3	5'-TGCCAATACCGTCCAAGTC-3'	5'-TGCCTTCCTCCTCCTCTG-3'
Tgf4	5'-CATGAATGAAGCAATATCAACTG-3'	5'-ATTTATTGTTCGTGGTTTCTTTAC-3'
Tgf5	5'-ATTTATTGTTCGTGGTTTCTTTAC-3'	5'-AGTCCTTTCGTCTTTCCTTATCG-3'
Tgf6	5'-CCACCTCTGAACAACAACAATAAC-3'	5'-CTATGAGCACGGCAGCAAAG-3'

***Table 4**: primer sequences used for evaluating the expression profile of HSP60, HSP70, HSP90, NOS, TNF, Bcl-2, Msi-1, Wnt, ColF, FGF, Tgf3, Tgf4, Tfg5, and Tgf6, using GAPDH as a housekeeper gene.*

4.1.5 Oxidative Stress Enzyme Assays

To measure the activity of stress-related enzymes, a 2–3 mm thick section of tissue from the growing edge of each specimen was excised using a sterile scalpel (Figure methods). Subsequently, the tissue section was manually diced into small fragments and placed into cold, sterile phosphate-buffered saline (PBS) with a pH of 7.5. A bead-mill homogenization using the Tissue-Lyser system was conducted in two cycles lasting 1 minute each. Following homogenization, the suspensions underwent centrifugation at 18,000 × g at 4°C for 20 minutes to eliminate any insoluble components; the resulting supernatant was preserved at −80 °C for subsequent experiments. The concentration of extracted protein was then determined using a Bradford assay.

4.1.5.1 Catalase Activity

Catalase activity was assessed by monitoring the breakdown of hydrogen peroxide (H_2O_2) according to Claiborne's procedure. Each reaction mixture comprised 890 µL of 100 mM PBS (pH 7.5), 100 µL of 500 mM H_2O_2, and 10 µL of the sample, added immediately prior to analysis. The catalase activity within the sample was determined by measuring the absorbance at 240 nm every 15 seconds for a duration of 15 minutes using a Beckman spectrophotometer (DU 640, Indianapolis, IN, USA). The outcomes are expressed as µmol H_2O_2/min/mg proteins.

4.1.5.2 GST Activity

GST activity was evaluated using 1-chloro-2,4-dinitrobenzene (CDNB) and the reduced glutathione form (GSH) as substrates. CDNB was dissolved in 100% ethanol, while GSH was dissolved in 100 mM Tris-HCl (pH 7.4). The final assay mixture contained 1 mL of 100 mM PBS (pH 7.5), 50 of 1 mM CDNB, 50 µL of 5 mM GSH, and 50 µL of sample, added right before measurement. The GST activity was assessed by measuring the absorbance at 340 nm every 15 s for 5 min with a Beckman spectrophotometer (DU 640). The results are presented as nmol GSH-CDNB complex/min/mg proteins.

4.1.5.3 Malondialdehyde (MDA) Content Determination

MDA is a thiobarbituric acid reactive species (TBARS) resulting from lipid peroxidation processes; it was quantified in each sample preparing an assay mixture containing 1 mL of 0.67% thiobarbituric acid, 400 µL of 20% trichloroacetic acid, and 100 µL of the sample. Solutions were incubated at 100°C for 15 minutes and then centrifuged at 4000 × g at 4°C for 10 minutes. The supernatant was kept and finally read at 512 nm using a Beckman spectrophotometer (DU 640). The MDA content is expressed as nmol TBARS/mg proteins.

4.1.6 Symbiont Bacteria Quantification

Genomic DNA was extracted with the CTAB method from the slices of tissue of the *C. reniformis* specimens stimulated as previously explained. The relative number of bacterial genes in each sample was quantified by real-time PCR analysis with a Chromo4 instrument (MJ Research, Bio-Rad, Hercules, CA, USA), using GAPDH as the reference gene.

4.2 Results

4.2.1 Sponge Morphology during Regeneration

Comparing control samples and 810 nm-1 W PBM irradiated samples, no significant morphologic differences were observed during the first three days of regeneration. Instead, after 120 hours (five days), the treated samples displayed a more rounded shape respect to the controls, especially SB. More specifically, a smaller cutting edge is sign of advanced regeneration; as shown in Figure 35, the regeneration percentage of the three irradiated samples was averagely ~70%, while about 35% in the respective controls.

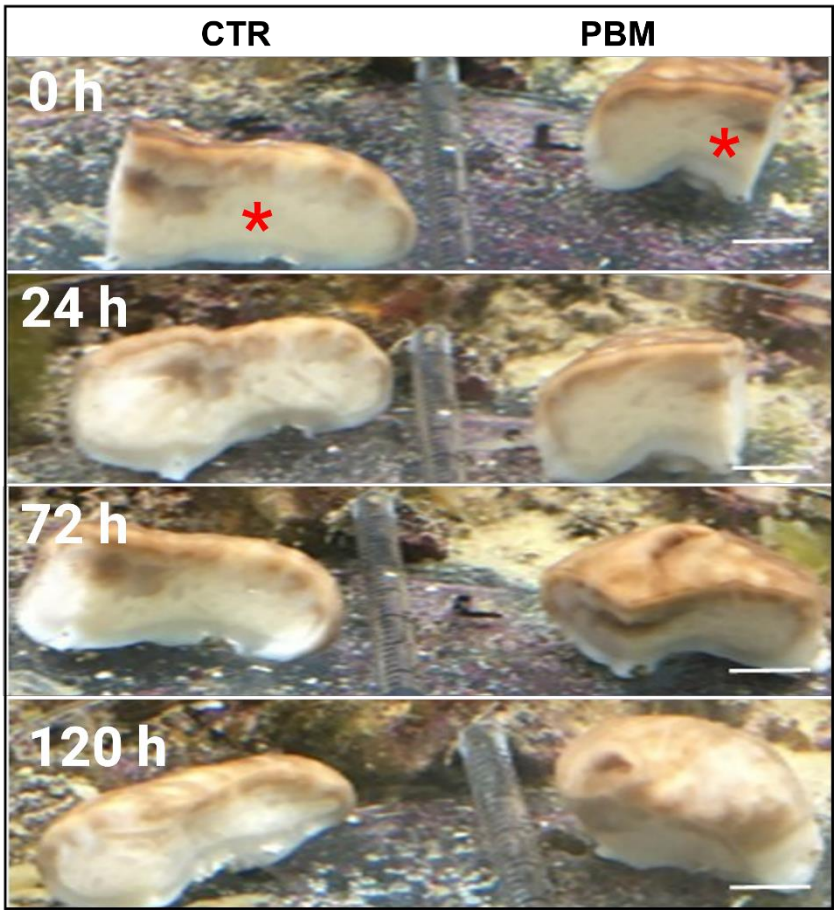


***Figure 35**: Effect of photo-biomodulation (PBM) on* Chondrosia reniformis *morphology during tissue regeneration. The regeneration profile of one representative* C. reniformis *specimen (notably, SB) immediately (t=0 h) and 24, 72, and 120 h after being cut into four parts, daily irradiated on the regenerating edge with 810 nm-1 W PBM. PBM = samples irradiated with 810 nm-1 W PBM; control = samples irradiated with 810 nm-0 W PBM. Scale bars are 1 cm. The centre of the irradiated area is indicated by the red asterisk.*

4.2.2 Histology and Archaeocyte Identification by Immunostaining

Masson's trichrome-stained slides revealed that 24 hours post-injury, both the control (see Figure 36A) and PBM-treated fragments (see Figure 36B) exhibited a regenerating edge covered by spherical cells and a sparse population of flattened cells, which resembled exopinacocytes. In particular, a higher number of cells, though smaller than what it is observed in the control, can be seen at the edge of the healing area. Immunofluorescence staining for Musashi 1 (Msi-1) (Figure 37A) depicted a significant abundance of archaeocytes within the initial few hundred micrometres from the regenerating edge in PBM-treated fragments, although archaeocytes were also detected in control fragments (Figure 37B and 37C).

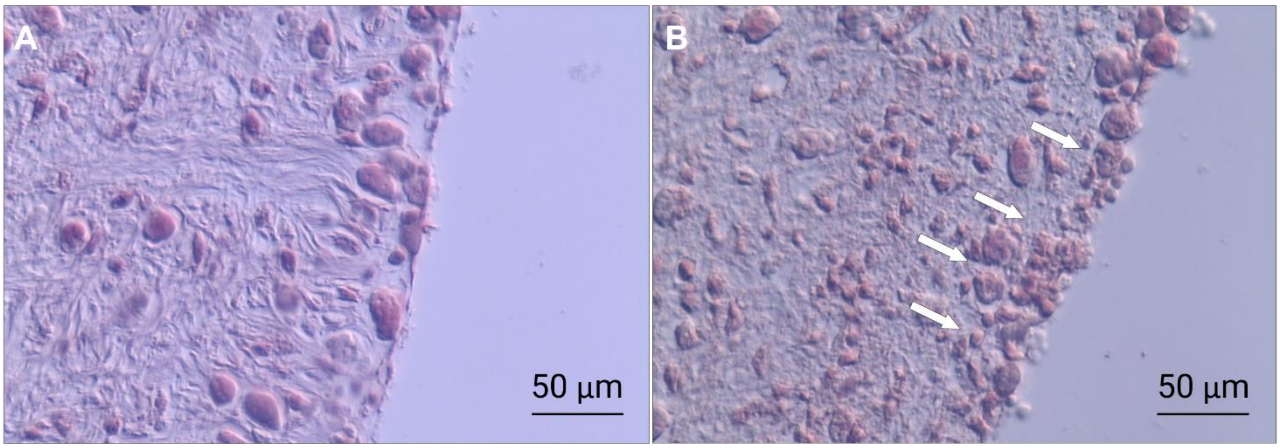


***Figure 36**: histological sections showing the regenerating edges of* Chondrosia reniformis *fragments 24 h after being cut. Panel A shows the control (810 nm-0 W PBM), and panel B shows the irradiated (810 nm-1 W PBM) fragments, both stained with Masson's trichrome. Some spherulous cells and exopinacocyte-like cells are visible on the regenerating edge.*

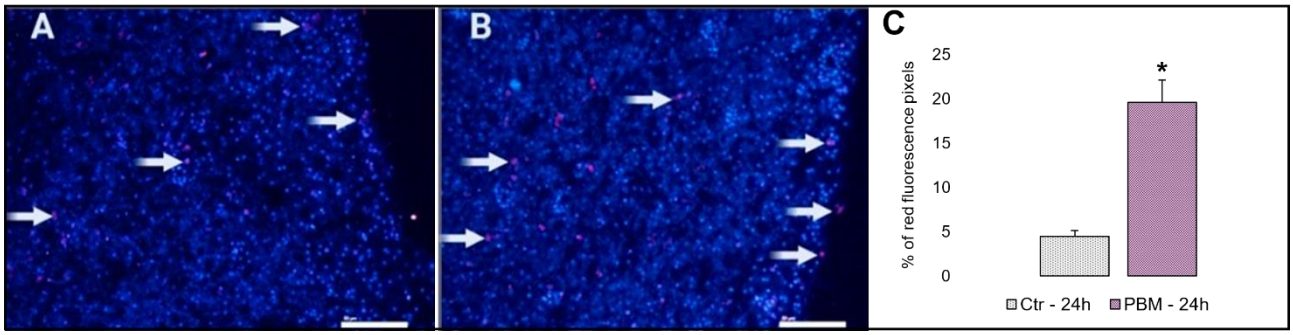


***Figure 37**: histological sections showing the regenerating edges of* Chondrosia reniformis *fragments 24 h after being cut. Panel A shows the immunofluorescence of the control (810 nm-0 W PBM) and panel B the irradiated (810 nm-1 W PBM) fragments, both stained for Msi-1 (red) and DAPI (blue). Archaeocytes are scattered in the sponge tissue (arrows). The scale bars are 50 µm. Panel C reports the semiquantitative Msi-1 immunofluorescence analysis: the percentage of red fluorescence pixels on total tissue-occupied pixels. Student's t-test, $p < 0.001$ (*). Each bar represents the mean ± standard deviation of three independent experiments performed in triplicate.*

4.2.3 Gene Expression Profile

Heat shock proteins (HSPs) and nitric oxide synthase (NOS) are known to be influenced by both inflammatory and heat-stress factors. Consequently, the expression patterns of these genes were investigated in the excised sponges. After 24 hours, HSP60 expression was significantly upregulated (30.47 ± 14.64-fold) in the 810 nm-1 W PBM-treated samples compared to the control, whereas after 120 hours, no significant differences were observed compared to the 120-hours control samples (Figure 38A). Notably, in the course of regeneration, the control samples exhibited a significant increase in HSP60 mRNA (22.0 ± 8.89-fold) at 120 hours post-excision, respect to 24 hours post-excision. Similarly, the expression pattern of the HSP70 gene (Figure 38B) mirrored that of HSP60, showing significant upregulation (35.33 ± 14.64-fold) at 24 hours post-excision; however, no significant differences were noted at 120 hours post-excision in the 810 nm-1 W PBM-treated samples compared to their respective controls. Nonetheless, at 120 hours post-excision, the non-irradiated samples exhibited significant upregulation (17.33 ± 5.51-fold) of the HSP70 gene compared to the untreated control samples at 24 hours. Taken together, these findings suggest that the PBM treatment primarily accelerates the response to HSP60 and HSP70 expression following tissue injury. Conversely, the expression of the HSP90 gene was significantly downregulated in the PBM-treated

samples at both 24 hours post-excision (0.45 ± 0.14-fold) and 120 hours post-excision (3.30 ± 1.00-fold) compared to their respective controls, suggesting a negative regulatory effect of the PBM treatment on the HSP90 gene in *C. reniformis* model. Consistent with the trends observed for HSP60 and HSP70 genes, HSP90 expression in the control samples increased by 3.15 ± 1.00-fold at 120 hours post-excision, if compared to 24 hours post-excision (Figure 38C). At both 24 hours and 120 hours post-cutting, no significant differences were observed in NOS gene expression between the irradiated and non-irradiated samples (Figure 38D). However, it's noteworthy that from 24 to 120 hours post-excision, NOS expression increased by 3.45 ± 0.75-fold in the control samples and 1.33 ± 0.39-fold in the irradiated samples. As reported in Figure 38E, at 24 hours post-excision the TNF mRNA had increased by 3.04 ± 0.72-fold in the PBM-treated samples compared to the controls, while at 120 hours post-excision a slight but significant downregulation of the TNF gene (0.75 ± 0.13) occurred compared to the control, indicating a return to pre-PBM baseline transcript levels. No differences were observed between the 24 hours and 120 hours control samples. The antiapoptotic protein Bcl-2 exhibited significant upregulation in the PBM-treated samples at both 24 and 120 hours post-excision, if compared to the control samples (3.33 ± 0.36-fold and 1.11 ± 0.4-fold, respectively), suggesting a reduction in apoptosis in the regenerating sponge tissue after irradiation. Conversely, no significant differences in Bcl-2 gene expression were noted between the 24 hours and 120 hours control samples (see Figure 38F).

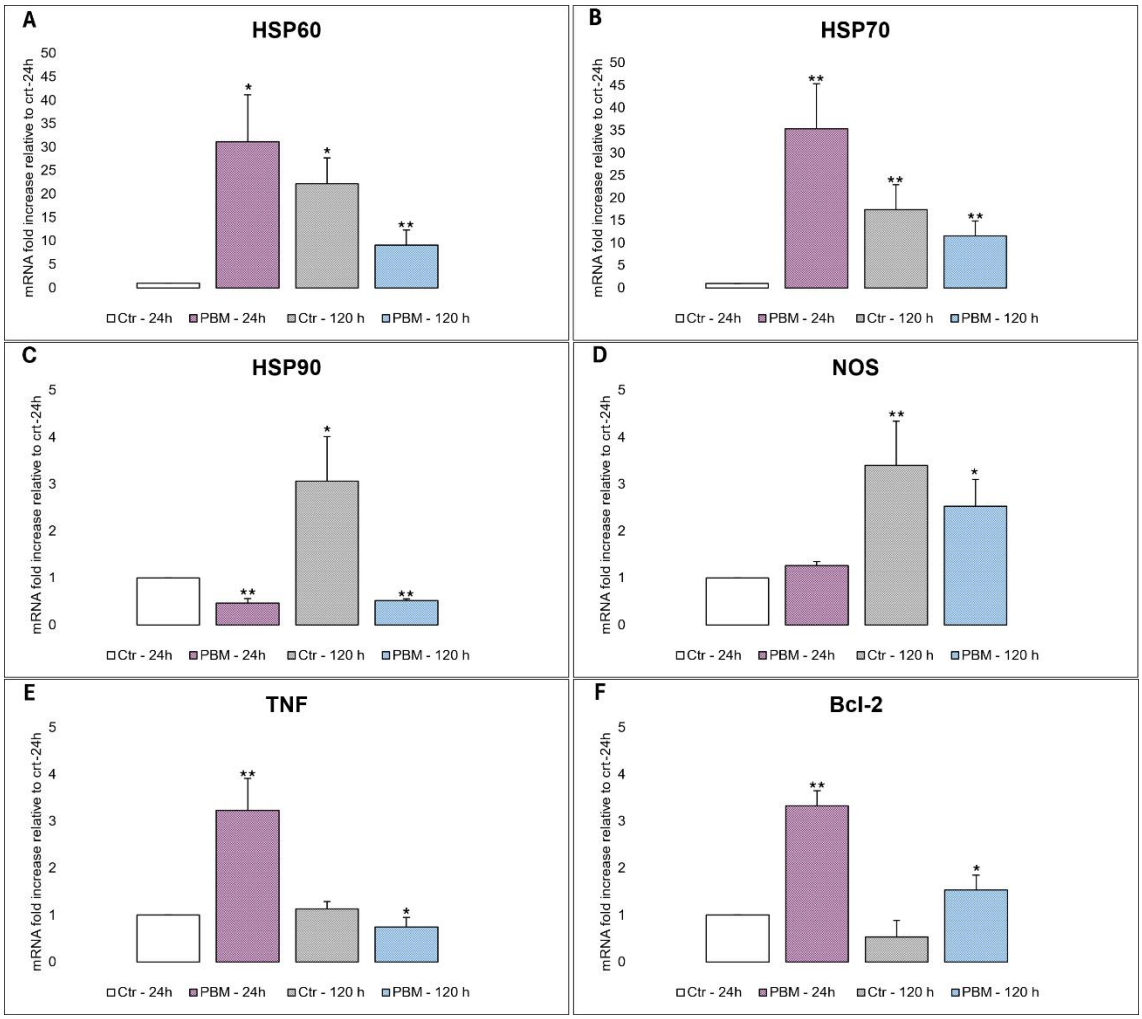


***Figure 38**: The effect of PBM on the expression of genes related to thermal stress, inflammation, and apoptosis. The real-time PCR gene expression analysis of heat shock proteins HSP60 (panel A), HSP70 (panel B), HSP90 (panel C), nitric oxide synthase NOS (panel D), tumour necrosis factor TNF (panel E), and Bcl-2 (panel F) in regenerating* Chondrosia reniformis *tissue 24 h (PBM-*

*24 h) or 120 h (PBM-120 h) after treatment. The control groups (Ctr-24 h and Ctr-120 h) are the respective 810 nm-0 W PBM nonirradiated samples (0 J). The data are expressed as the messenger RNA (mRNA) fold-increase relative to the Ctr-24 h sample and are normalised to GAPDH expression. Each bar represents the mean ± standard deviation of the three independent experiments performed in triplicate. The data were analysed with one-way analysis of variance (ANOVA) followed by Tukey's test (ANOVA p-values: (A) p = 0.013071; (B) p = 0.001023; (C) p = 0.000511; (D) p = 0.003392; (E) p = 0.000192; (F) p = 0.000014). The asterisks indicate significant differences between groups based on Tukey's test versus Ctr-24hrs; * p < 0.05; ** p < 0.01.*

To assess whether PBM could promote regeneration in sponges, the expression of Msi-1, known as an archaeocyte cell marker, was analyzed. As demonstrated in Figure 39A, at 24 hours post-excision, Msi-1 was upregulated by 5.19 ± 1.74-fold in the PBM-treated samples if compared to the control samples. However, at 120 hours post-excision, Msi-1 did not differ significantly between the two groups. Furthermore, in both the 120 hours control and PBM-treated samples, Msi-1 decreased significantly (0.3 ± 0.14 and 0.1 ± 0.007-fold, respectively) compared to the 24 hours control samples, indicating a transition from archaeocytes to exopinacocytes. In the injured tissue of higher animals, the Wnt (endogenous wingless protein) signal activates stem cells and contributes to tissue repair. As showed in Figure 39B, at 24 hours post-excision, Wnt-like mRNA increased by 1.7 ± 0.25-fold in the PBM-treated samples compared to the controls. There were no significant alterations in Wnt gene expression between the groups at 120 hours post-excision, but a notable increase in gene expression (1.61 ± 0.27-fold) was observed in the 120-hour control samples compared to the 24-hour controls. At 24 hours post-excision, fibrillar collagen exhibited a substantial increase by 49.35 ± 16.36-fold in the PBM-treated samples compared to the control samples (Figure 39C). However, at 120 hours post-excision, there were no discernible differences between the groups. In control samples, fibrillar collagen expression increased by 2.35 ± 0.73-fold from 24 to 120 hours post-excision, whereas the opposite trend was observed in the PBM-treated samples. The PBM treatment did not influence the expression of the fibroblast growth factor (FGF) homologue at either of the analyzed time points (Figure 39D). Several members of the transforming growth factor (TGF) family were also examined. At 24 hours post-excision, TGF3 was 0.68 ± 0.1-fold lower in the PBM-treated samples compared to the control samples. Although there was no significant difference between the groups at 120 hours post-excision, a decrease in TGF3 gene expression in the 120-hour control samples was observed compared to the 24-hour controls (0.46 ± 0.18-fold), indicating progression of the regeneration process (Figure 39E). TGF4 did not exhibit significant differences between the groups at either time point. However, in the control samples, TGF4 increased by 2.38 ± 0.51-fold from 24 to 120 hours post-excision (Figure 39F). At 24 hours post-excision, there was a slight but significant reduction in TGF5 expression in the PBM-treated samples compared to the controls (0.73 ± 0.14-fold lower). This reduction persisted over time, and by 120 hours post-excision, TGF5 gene expression was strongly downregulated by 0.46 ± 0.07-fold in the PBM-treated samples compared with the 24-hour control samples, resulting in significant downregulation in the 120-hour control samples (Figure 39G). Finally, in the PBM-treated samples, TGF6 expression was significantly upregulated by 1.36 ± 0.11-fold at 24 hours post-excision compared with the control samples. There were no significant differences between the groups at 120 hours post-excision. At that time point, TGF6 was markedly downregulated in the PBM-treated samples (0.4 ± 0.34-fold) compared to the 24-hour control samples (0.33 ± 0.19-fold) (Figure 39H).

It must be mentioned that, by the time of publication of this paper, the sequences of the five fibrillar collagen genes were not available yet, and that is the reason why their expression after laser treatment couldn't be quantified. The gene "ColF" corresponds to G2 (see Chapter III), as it was previously identified by [Pozzolini et al., 2015].

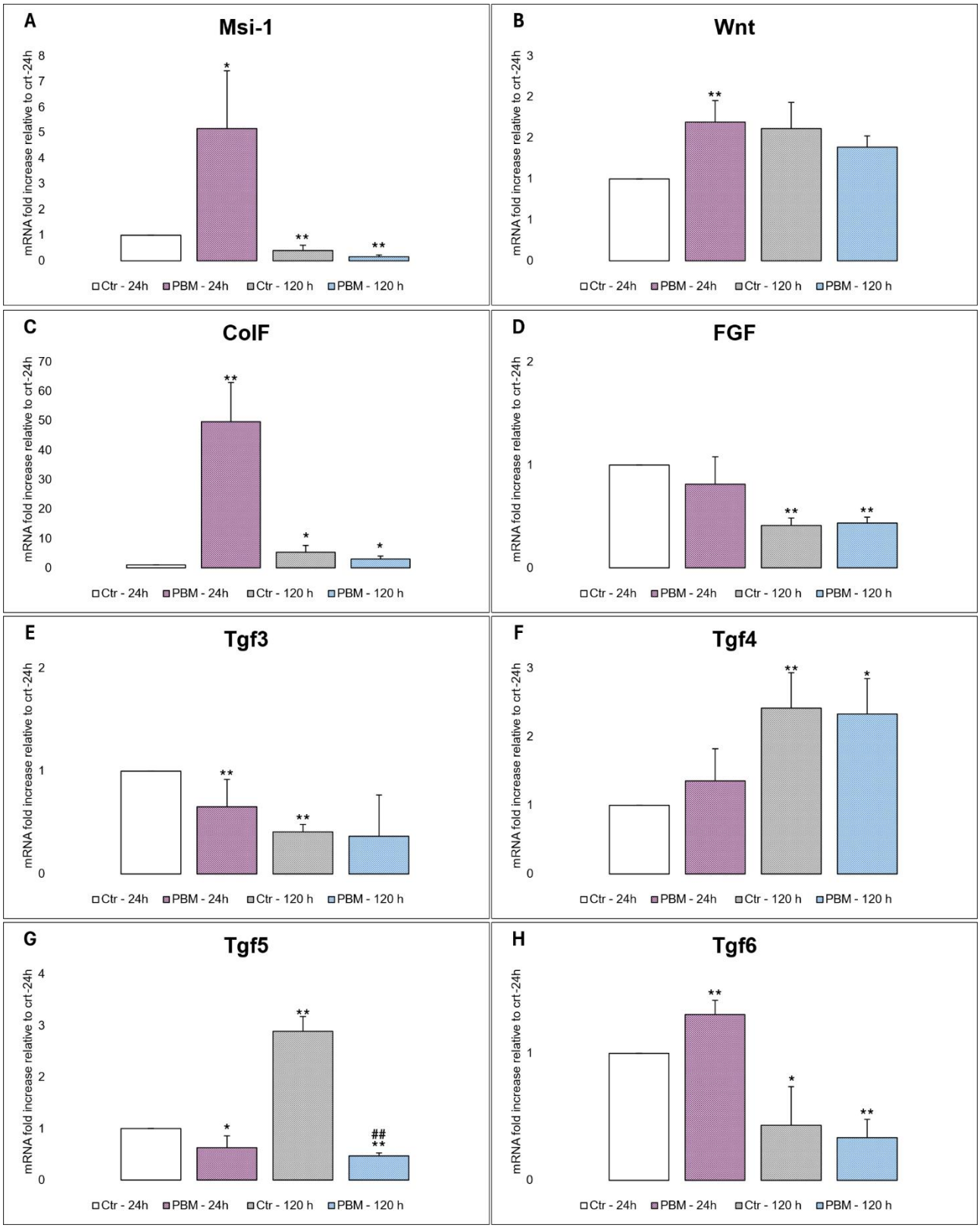


Figure 39*: PBM on the expression of proliferation/differentiation, extracellular matrix synthesis, and growth factor genes. The real-time PCR gene expression quantification of the pluripotent stem cell-inducer gene Musashi 1 Msi-1 (panel A), the endogenous wingless protein (Wnt) signalling pathway (panel B), fibrillar collagen (panel C), fibroblast growth factor FGF (panel D), and transforming growth factors TGF3, TGF4, TGF5, and TGF6 (panel E, F, G, H) in regenerating* Chondrosia reniformis *tissue 24 h (PBM-24 h) and 120 h (PBM-120 h) after treatment. The control groups (Ctr-24 h and Ctr-120 h) are the respective 810 nm-0 W PBM nonirradiated samples (0 J). The data are expressed as the messenger RNA (mRNA)-fold increase relative to Ctr-24 h and are normalised to GAPDH expression. Each bar represents the mean ± standard deviation of three independent experiments performed in triplicate. The data were analysed with one-way analysis of variance (ANOVA) followed by Tukey's test (ANOVA p-values: (A), p = 0.013071; (B), p = 0.001023; (C), p = 0.000511; (D), p = 0.003392; (E), p = 0.000192; (F), p = 0.000014). The asterisks*

*indicate significant differences between the groups based on Tukey's test versus CTR-24 h, * p < 0.05, ** p < 0.01. The hash symbol indicates significant differences between groups based on Tukey's test versus CTR-120 h, ## p < 0.01.*

4.2.4 Sample Temperature Analysis

Thermal camera measurements of the 810 nm-1 W PBM specimen surface registered 14.38 ± 0.21 °C before treatment and 16.57 ± 0.33 °C after treatment.

4.2.5 Oxidative Stress Evaluation

It is possible that he 810 nm-1 W PBM could enhance reactive oxygen species (ROS) generation *via* modulation of mitochondrial respiratory chain activity. Consequently, the activity of oxidative stress enzymes was assessed in regenerating *C. reniformis* samples 6 hours after irradiation; in particular, the levels of catalase and glutathione transferase (GST) activity were investigated, as well as the content of malondialdehyde, which is an oxidative stress marker. As outlined in Table 1, no significant differences were observed in these markers between the groups, suggesting that PBM does not induce oxidative stress in sponge tissues.

Sample	Catalase Activity (µmoles H_2O_2/min/mg Protein)	GST Activity (µmoles GST-CDNB/min/mg Protein)	Malondialdehyde Content (nmoles/mg Protein)
Ctr	6.03 ± 0.1	0.46 ± 0.19	172.8 ± 75.37
PBM	5.41 ± 0.14	0.52 ± 0.12	225.61 ± 88.01

Table 5: Oxidative stress enzyme activity and malondialdehyde content in Chondrosia reniformis *6 h post irradiation. The samples were normalized to total protein content. The data are presented as the mean ± standard deviation. Ctr = 810 nm-0 W nonirradiated specimens (control); GST = glutathione transferase; PBM = 810 nm-1 W irradiated specimens.*

4.2.6 Prokaryotic Symbiont Quantification

Considering that PBM interacts with prokaryotic cells, the microbial content in the regenerating edge of *C. reniformis* samples wax examinated. DNA was isolated from the samples 24 and 120 hours after cut, and the prokaryotic DNA/sponge DNA *ratio* was calculated. As shown in Figure 40, at 24 hours post-excision, the prokaryotic DNA slightly but significantly increased in the PBM-treated samples compared to the control samples (1.19 ± 0.11-fold). However, at 120 hours post-excision, no significant difference was observed between the groups. Notably, all samples exhibited a substantial increase in prokaryotic DNA relative to the 24-hour time point (control: 2.89 ± 0.81-fold; PBM: 2.52 ± 0.68-fold).

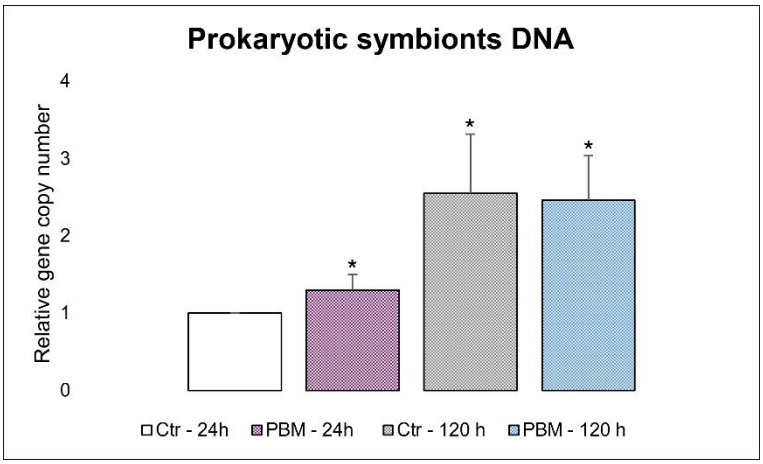


***Figure 40**: Prokaryotic symbiont DNA quantification. Analysis of the sponge prokaryotic symbiont population based on real-time PCR DNA quantification in the regenerating* Chondrosia reniformis *specimens 24 h post injury and the first dose of 810 nm-1 W*

*photo-biomodulation (PBM-24 h) and 120 h post injury (PBM-120 h). The control groups (Ctr-24 h and Ctr-120 h) are the respective non-irradiated samples. The data are expressed as the prokaryotic DNA gene copy number relative to Ctr-24 h and are normalised to the GAPDH gene copy number. Each bar represents the mean ± standard deviation of the three independent experiments performed in triplicate. The data were analysed with a one-way analysis of variance (ANOVA) followed by Tukey's test (ANOVA $p = 0.006047$). The asterisks indicate a significant difference between groups based on Tukey's test (versus CTR-24 h, * $p < 0.05$).*

4.3 Discussion

Poriferans represent the earliest branching *phylum* within the animal kingdom. Being a sister group to all other animal taxa and having a very simple body plan, these organisms have often been referred to as "living fossils" and are considered to be early animal ancestors. However, some authors argue that although the branching of Parazoa/Eumetazoa dates back millions of years, the Porifera *phylum* has been evolving for as long as any other modern animal lineage. Therefore, the remarkable morphological simplicity of sponges is coupled with a high degree of molecular complexity. Over the past two decades, in sponges have been identified various molecular pathways that are also typical of higher animals, including factors involved in innate immune responses and development/ differentiation processes. Considering this information, the effect of 810 nm-1 W PBM on sponge tissue regeneration was evaluated to assess whether the molecular response is conserved and has been maintained throughout evolution across the animal kingdom (for complete references of this sections, see the original work "**Near-Infrared 810 nm Light Affects Porifera *Chondrosia reniformis* (Nardo, 1847) Regeneration: Molecular Implications and Evolutionary Considerations of Photobiomodulation–Animal Cell Interaction**", Amaroli et al., 2022, https://doi.org/10.3390/ijms24010226. The qualitative analysis of wound healing 120 hours after cutting showed that the irradiated samples exhibited a more rounded shape and smaller cutting edge compared to the controls. Histological analysis of the regeneration edge at 24 hours post-injury supported this observation. It was confirmed that 810 nm-1 W PBM treatment enhanced the regeneration process in the sponge tissue at a molecular level, as evidenced by the robust upregulation of the fibrillar collagen gene 24 hours post-injury. The rapid and extensive deposition of new extracellular matrix components, such as fibrillar collagen, is crucial for the tissue regeneration of a marine sponge like *C. reniformis*. Furthermore, the early and significant upregulation of the Msi-1 gene – considered a sponge stem cell marker – also indicated an improved tissue regeneration induced by 810 nm-1 W PBM. The abundance of Msi-1-positive cells in the regenerating edge of the PBM-treated samples was further supported by immunostaining analysis. A set of TGF family members – TGF3 and TGF6 –, whose expression profiles may possess pro-stemness activity in *C. reniformis*, was previously illustrated. Their upregulation in the early regeneration phases and subsequent decrease during sponge tissue restoration were described. Additionally, TGF4 and TGF5 were considered to be pro-differentiating factors, as their expression increased only during the final phase of sponge tissue regeneration. A significant upregulation of TGF6 was observed in the PBM-treated samples compared to untreated control samples during the early phases of regeneration, supporting the positive effect of 810 nm-1 W PBM on sponge tissue recovery. However, in the irradiated samples, a downregulation of TGF3 was also noticed in the early phases, as well as a downregulation of TGF5 120 hours post-injury. These latter results may be attributed to a faster return to baseline expression levels of these genes in the PBM-treated samples, similarly to what was observed for HSP60 and HSP70. PBM may influence TGF3 and TGF5 by stimulating a more rapid response in gene expression. Conversely, the expression of a sponge FGF homologue and of TGF4 seemed not to be affected by irradiation, indicating that multiple independent pathways regulate these genes.

PBM has the potential to induce a selective photothermal effect with local overheating occurring in pigmented tissue. Although the treated samples were pigmented, no macroscopic temperature

increase was observed in the irradiated tissue and the surrounding medium. This finding is not unexpected: in *C. reniformis*, pigment deposition primarily occurs in the cortical region, while the irradiation was performed on the cutted mesohyle, which exhibits a uniform white appearance; moreover, the sponge surface temperature promptly returned to baseline in the aquarium. Therefore, the marked upregulation of HSP60 and HSP70 genes 24 hours post-injury can be interpreted as primarily due to the protein folding activity associated with the HSP family, rather than thermal stress. HSPs play a pivotal role in the healing process; for instance, in tendons, HSPs modulate the inflammatory response and reduce apoptosis by promoting the release of inflammatory cytokines and chemokines. PBM treatment for regenerating tendon tissues induces HSP70 upregulation, thereby improving wound healing. The upregulation of TNF and antiapoptotic Bcl-2 genes in this experimental model suggests that HSP70 and HSP60 may play a similar role in response to wound healing across the animal kingdom and are targeted by PBM even in lower metazoans. Overall, these findings indicate that PBM-induced inflammation is not detrimental, as it is concomitantly counteracted by the support of antiapoptotic membrane instability pathways. The dissociation between TNF signaling and apoptosis observed in this study is consistent with a previous work, where *C. reniformis* was exposed to quartz dust. This phenomenon could be explained by the structural differences in the intracellular region of the TNF receptor between sponges and higher animals.

Chapter V - NOVEL EMPLOYMENTS OF BIOMATERIALS OBTAINED FROM THE COLLAGEN OF *Chondrosia reniformis* (Nardo, 1847)

By the time of the beginning of this doctoral research activity, previous studies had examined the collagen from *Chondrosia reniformis* as a biomaterial, particularly focusing on its employment in bone repair applications [Palmer et al., 2015] and on its capacity to create biocompatible 2D-membranes [Pozzolini et al., 2018]. Nevertheless, its potential for biomedical applications related to skin repair had not been investigated before. This part of the doctoral research activity was thus focused on exploring the potential of *C. reniformis*-derived collagen 2D-membranes as skin-like substitutes devices in the treatment of wounds. The results obtained enabled the publication of a paper ("**2D Collagen Membranes from Marine Demosponge *Chondrosia reniformis* (Nardo, 1847) for Skin-Regenerative Medicine Applications: An *In Vitro* Evaluation**" – Tassara et al., 2023; https://doi.org/10.3390/md21080428) and the creation of a patent ("**Composition based on collagen of marine origin**" – Pozzolini, Tassara, Giovine, and Lova, 2023; Italian patent N. 102023000014598), which essential sections are reported below for both (for complete references, see the original works).

2D COLLAGEN MEMBRANES FROM MARINE DEMOSPONGE *CHONDROSIA RENIFORMIS* (NARDO, 1847) FOR SKIN-REGENERATIVE MEDICINE APPLICATIONS: AN *IN VITRO* EVALUATION.

Skin injuries, whether acute or chronic, are prevalent in both human and veterinary medicine and pose significant challenges to healthcare systems. Traditional treatments like skin grafts, though long-standing, present limitations in availability and can trigger strong rejection responses. Therefore, the field of regenerative medicine continually seeks advanced wound-dressing solutions and bioengineered substitutes resembling natural skin to promote tissue regeneration and expedite healing with minimal side effects. An ideal skin substitute should mimic skin's structure and function, providing a barrier against fluid loss and infections while being biocompatible and possessing adequate thermal and mechanical properties. This has led to growing interest in innovative polymers, with collagen being a prominent candidate due to its widespread availability. However, concerns persist regarding the immunogenicity and ethical considerations of "classical" collagens, to which marine collagen represents a promising alternative (see the general introduction section). This part of the doctoral research activity aimed to explore the suitability of 2D membranes derived from *C. reniformis* collagen in treating skin injuries, as skin-like substitutes (SLSs) devices. Assessments were conducted to evaluate membrane biodegradability, water and protein permeability, and bacterial retention capacity, crucial for creating an optimal wound environment. Additionally, cellular responses of fibroblast cells were examined, focusing on gene expression related to extracellular matrix proteins and mediators involved in ECM equilibrium maintenance.

5.1 Materials and methods

All reagents were obtained from SIGMA-ALDRICH (Milan, Italy), unless otherwise specified.

5.1.1 Sponge sampling

Several specimens of *Chondrosia reniformis* were taken in the waters of the Portofino Promontory (Liguria, Italy) at 10–20 m of depth by scuba diving, then maintained in thermic bags at a temperature of 14–15 °C until arrival in the laboratory, where the animals were allowed to acclimatize in an aquarium. A short stabulation was conducted before experiments by keeping the animals at 14 °C in a 200 L aquarium, containing natural seawater collected in the same area of the sampling and equipped with an aeration system, with a salinity of 38‰.

5.1.2 *Chondrosia reniformis* Fibrillar Collagen Extraction and Purification

A pool of fresh animals, weighing a total of 70 g, was processed to obtain the collagen fibrils in their native conformation, with some modifications. Briefly, sponges were rinsed twice with artificial seawater, cut into small pieces of a few mm^3, and put into 5 volumes of a solution with 0.1% trypsin and 100 mM ammonium-bicarbonate buffer (pH 8.5), then incubated overnight at 37 °C into a horizontal shaker. The liquid part was then removed by filtration with a metal strainer, while the solid material was soaked in 3 volumes of cold deionized water and incubated at 4 °C for three days in a rotating-disc agitator. This process allows to obtain a viscous suspension that was filtered with a metal strainer (keeping the remaining fragments of tissue for further cycles of extraction) and firstly centrifuged at 1500 × *g* for 10 minutes at 4 °C, to remove any residues of cells and sediment particles. The supernatant fluid containing the collagen suspension was finally recovered by centrifugation at 18,000× *g* for 35 min at 4 °C. The resulting pellet was washed twice with deionized water to eliminate any trypsin residues and resuspended in clean deionized water. This entire procedure was repeated six times until the fragments of sponge tissue were completely dissolved. The final fibrillar suspension, resulting from all the collagen collected from extraction cycles, has been stored at 4 °C until use. To improve the biocompatibility of the sample, some of the polysaccharidic components – which are inevitably co-extracted with sponge collagen – were partially removed by simply treating the suspension with an equal volume of 0.1 M NaOH for 6 h at 25 °C, then recovering again the collagen suspension by centrifugation at 18,000× *g* for 35 min, with two further washing steps. The ECM extract concentration (mg/mL) was then determined by drying and weighing a known volume of the suspension. The total collagen yield was calculated considering previous observations.

5.1.3 Production of 2D Collagen Membranes

The 2D collagen membranes (2D-CMs) were obtained by pouring 3 mL of a 3 mg/mL ECM fibrillar extract (ECM-FE), dissolved in distilled water, in rectangular silicon molds (25 × 28 mm) and let completely dry overnight at 37 °C, until thin bidimensional sheets with a mean weight of 8.5 ± 2.7 mg and a thickness of 0.1 mm were obtained. This 2D-CMs were used for permeability, bacteria infiltration, bacteriostatic activity, and *in vitro* biodegradability tests. For gene expression analysis, the ECM-FE was used to directly coat 6-well plates; each 2D-CM coating was sterilized with ethanol and under UV light for 30 min.

5.1.4 2D Collagen Membranes Permeability Tests

The ability of the *C. reniformis*-derived 2D-CMs to prevent liquid and protein loss was evaluated by testing their permeability to both distilled water and bovine serum albumin (BSA). The 2D-CMs were fixed into sterilized supports that were created *ad hoc* for this experimental activity (used as shown in Figure 41); the effectiveness of the experimental setup was confirmed with some preliminary trials. The experiments, performed in triplicate at room temperature, followed the example of [Ferrario et al., 2020] and were meant to recreate a dry and a moist skin wound condition.

In the former case, the term “dry” stands for a type of skin injury that does not release exudates; therefore, it was reproduced a situation in which one side of the 2D cm was exposed to air and the other side was wetted (“dry-wet” condition). In the latter case, the word “moist” indicates a type of skin injury that releases exudates; here, both sides of the 2D cm were hydrated (hence, “wet-wet” condition). To prevent long-term evaporation, in all cases the structures were kept in a humidified environment.

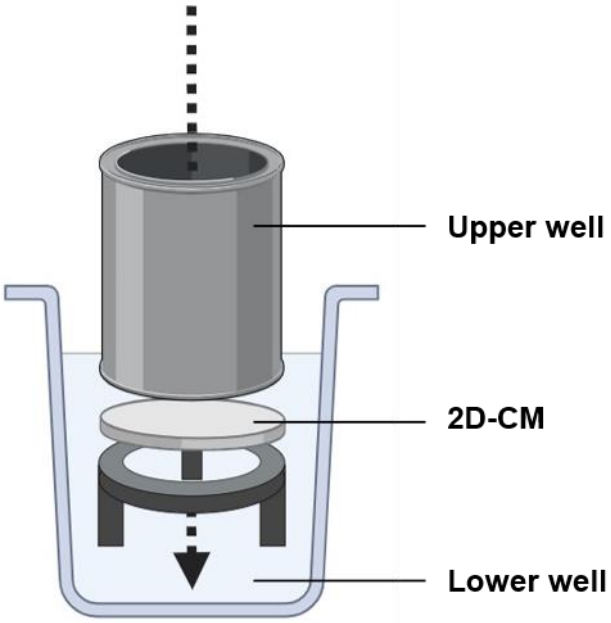


***Figure 41**: structure of the tools used for permeability and bacteria infiltration tests. Image created with www.Biorender.com).*

5.1.4.1 Water Permeability Test

“Dry-wet” condition: 2 mL of deionized water was put into the upper well, while the underlying well was left empty (see Figure 42A). The passage of water through the 2D cm was visually monitored at 0, 1, 3, 6, 24, 48 h, and 7 days, respectively.

“Wet-wet” condition: 2 mL of deionized water were put into the upper and the lower wells, in a way that both sides of the 2D-CM were in contact with the liquid (see Figure 42B). Water passage through the 2D-CMs was quantified by weighing the water in the upper well at the same time points set for the “dry-wet” condition. To prevent evaluation biases, the inserts were kept in a closed environment, and a control insert was used, which allowed us to keep track of possible evaporation. The percentage of deionized water that crossed the 2D-CMs after the stated time points was obtained as follows: [weight (g) of water in the upper well at a final time point/weight (g) of distilled water in the insert at time zero] × 100. The permeability of the 2D-CMs was expressed as the volume of water passing through the surface at each time-point (mL/cm^2).

5.1.4.2 Protein Permeability Test

In the “dry-wet” experiment condition with distilled water, no amount of liquid could pass through the membrane; given that solute cannot cross a barrier if its solvent is not able to do so in the first place, only the “wet-wet” condition experiment was performed for the protein permeability test (see Figure 42C). The upper well was filled with 2 mL of a 35 mg/mL bovine serum albumin (BSA) solution prepared in 50 mM PBS (pH 7.4), while the lower well contained only deionized water. At

specific time points (1, 3, 6, 24, 48 h, and 7 days, respectively), 100 µL of the lower well content was taken and used to perform a Bradford assay to quantify the protein amount in the solution. Samples were read at a 595 nm wavelength using a Beckman spectrophotometer (DU 640, (Beckman Coulter s.r.l., Milan, Italy). For each time-point, the amount of BSA occurring in the lower well (representing the protein permeability of the 2D-CMs) was calculated as a percentage as follows: 35 mg/mL: 100% = [BSA] in the lower wells: x.

5.1.5 Bacteria Infiltration Tests

Beside the ability to prevent water and protein loss, the epidermal layer should prevent pathogens from entering the body. The 2D-CMs were therefore tested to evaluate their effectiveness in avoiding bacteria infiltration. Three different strains of bacteria were used: *Staphylococcus aureus* ATCC 29273 (Gram-positive coccus), *Pseudomonas aeruginosa* ATCC 27853, and *Escherichia coli* ATCC 25404 (both Gram-negative bacilli). The 2D-CMs were fixed to the same type of support as created for water and protein permeability tests (Figure 42D). The upper well was filled with 2.5 mL of a 1× PBS solution (which doesn't support bacteria growing conditions) containing bacteria at a concentration of 10^7 cells/mL, while the lower well only contained the 1× PBS sterile solution. The two solutions could get in touch only through the membrane. The structures were left in incubation for 48 hours at 37 °C. After the set period, both the upper and lower wells content was collected to be analyzed (the former, to verify bacteria viability; the latter, to check an eventual bacterial infiltration through the 2D-CM). The solutions were centrifuged for 10 min at 5000 × *g* to pellet any bacteria; then, the pellet was re-suspended in 2.5 mL of 1× PBS. Finally, 100 µL of the samples were plated on LB agar plates, some undiluted and several-diluted in 1× PBS (10^{-3} to 10^{-6} dilutions), and left in incubation overnight at 37 °C. The number of colony-forming units (CFUs) was determined for each sample. The permeability of the 2D-CMs to bacteria was expressed as the percentage of bacteria found in the flow-through compared to the initial bacteria concentration (10^7 cells/mL). Accordingly, the percentage of retained bacteria was also calculated. The same procedure was carried out in triplicate for all the selected bacterial species.

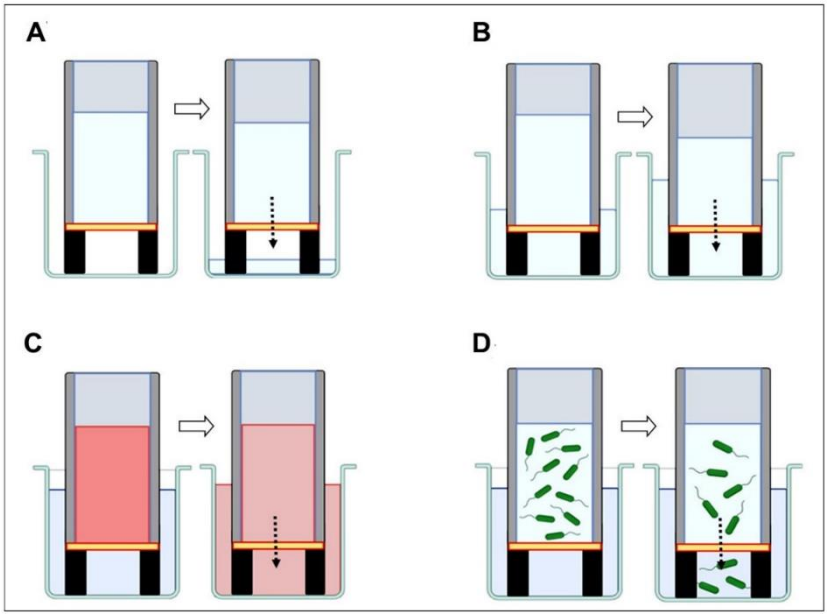


***Figure 42**: Experimental design for the* C. reniformis*-derived 2D-CMs permeability tests. Panel A: permeability to deionized water in "dry-wet" conditions; panel B: permeability to deionized water in "wet-wet" conditions. Panel C: permeability to proteins (BSA) in "wet-wet" conditions; panel D: bacteria infiltration test.*

5.1.6 Bacteriostaticity Test

The effect of *C. reniformis*-derived 2D-CMs was tested on the growth of *Staphylococcus aureus* ATCC 29273, *Pseudomonas aeruginosa* ATCC 27853, and *Escherichia coli* ATCC 25404 to evaluate if they could show bacteriostatic activity. It has been noticed, in fact, that collagen suspensions from this sponge do not develop bacterial contaminations and do not undergo bacterial degradation even if kept at 4 °C in non-sterile conditions (unpublished observations). One 25 × 28 mm membrane was put inside a tube filled with 10 mL of an LB solution with a bacterial concentration of 1×10^4 CFUs/mL and left in moderate agitation at 37 °C for 6 h, which was considered enough time for the bacteria to reach their exponential growth phase, as well as for 24 h. Control tubes with no membranes inside were also prepared. After the expected amount of time, 100 µL of each sample was plated, undiluted or several-diluted (10^{-1} to 10^{-4} dilutions), into LB agar plates and left in incubation overnight to monitor any difference in the CFU number between samples. The same experiment was repeated in triplicate.

5.1.7 *In Vitro* Biodegradability Test

A good biomaterial for SLSs should be thermally and mechanically stable, but it should also be biodegradable in the medium/long term. Thus, the stability of the 2D-CMs was *in vitro* evaluated. The initial weight (mg) of the dry membranes was registered (W_i), then they were soaked inside wells containing 5 mL of 1× PBS (pH 7.4) solution or of a 0.1 mg/mL *Clostridium histolyticum* collagenase solution prepared with 50 mM tricine, 10 mM $CaCl_2$, and 400 mM NaCl and left in incubation at 37 °C for 7, 14, and 21 days, respectively. Collagenase was renewed every 7 days. After the expected amount of time, the 2D-CMs were air-dried, and their weight was annotated again (W_f). The percentage (%) of average weight loss was expressed as follows: $[(W_f - W_i)/W_i] \times 100$.

5.1.7.1 Evaluation of the Material Released into the Incubation Media

The incubation media, in which the 2D-CMs were soaked for the in vitro biodegradability test, was recovered at 7, 14, and 21 days and used to evaluate the organic material released from the membranes, responsible for their weight loss over time. The protein component released in the media from the 2D-CMs in PBS or collagenase solution was measured using the bicinchoninic acid method after 7, 14, and 21 days using PBS or collagenase solution as a blank, respectively. Since the ECM-FE from which the 2D-CMs were made with mainly constitutes of fibrillar collagen and ECM-associated glycosaminoglycans, a hydroxyproline assay and a GAGs assay were also performed. To quantify the percentage of collagen-deriving material released in PBS or collagenase solution from the 2D-CMs during the experiment, the incubation media were hydrolysed in 2M NaOH at 120 °C for 20 min at 1 atm, then the hydroxyproline concentration was evaluated; the absorbance of each sample was read with Beckman spectrophotometer (DU 640), wavelength 550 nm, using a cis-4-hydroxy-l-proline as a standard curve. For each sample, the degraded collagen content was obtained using the proportion factor of 1 g of hydroxyproline per 6 g of collagen. Finally, the GAGs content within the incubation media was also quantified with an Alcian-Blue assay in comparison with a standard curve of 0.1 mg/mL chondroitin sulphate, reading the absorbance of each sample was read with a Beckman spectrophotometer (DU 640) at a wavelength of 620 nm.

5.1.8 Effect of 2D Collagen Membranes on Fibroblasts

5.1.8.1 Cell Cultures

The L929 mouse fibroblast cell line was obtained from the National Collection of Type Cultures (NCTC, Salisbury, UK). Cell cultures were maintained at 37 °C in a humidified, 5% CO2 atmosphere, in high glucose Dulbecco's modified Eagle's medium (D-MEM) with glutamax (Euroclone, Milan, Italy), which was supplemented with 10% FBS (Euroclone) and added with penicillin/streptomycin as antibiotics.

5.1.8.2 Fibroblast Gene Expression Analysis

To perform a further gene-expression analysis, L929 mouse fibroblast cells were seeded into 6-well plates that were or were not pre-coated with the *C. reniformis*-derived ECM-FE. Cells were seeded at a density of 300,000 cells/well for the 24 h gene expression analysis or 50,000 cells/well for the 120 h gene expression analysis and then incubated at 37 °C for the established amount of time. At the end of each experiment, total RNA was extracted with the RNeasy Mini Kit, (Qiagen, Milan, Italy) following the manufacturer's instructions. The cDNA was synthesized using the Revert Aid Reverse Transcriptase (Thermo Fisher Scientific, Milan, Italy) by taking 1 µg of purified total RNA from each sample. Every PCR reaction was performed in 15 µL containing 1× master mix iQ SYBR®Green (Bio-Rad, Milan, Italy), 0.2 µM of each primer, and 3 µL of a 1:5 diluted reverse transcription reaction buffer. Each sample was analyzed in triplicate. The following thermal conditions were used: initial denaturation at 95 °C for 3 min, followed by 45 cycles with denaturation at 95 °C for 15 s, and annealing and elongation at 60 °C for 60 s. Fluorescence was measured at the end of each elongation step. GAPDH was used as a reference gene and the values were normalized to its mRNA expression. All the PCR primers (See table) were designed through the Beacon Designer 7.0 software (Premier Biosoft International, Palo Alto, CA, USA) and acquired from TibMolBiol (Genova, Italy). Data analyses were obtained from the DNA Engine Opticon®3 Real-Time Detection System Software program (3.03 version) and, to calculate the relative gene expression compared to an untreated (control) calibrator sample, the comparative threshold Ct method was used within the gene expression analysis for iCycler iQ Real-Time Detection System Software® (2004 Bio-Rad, Milan, Italy). Data are means ± S.D. of three independent experiments performed in triplicate.

	Primer sequences	
Gene	*Forward*	*Reverse*
GAPDH	5'-TCTCCCTCACAATTTCCATCCCAG-3'	5'-GGGTGCAGCGAACTTTATTGATGG-3'
COL1A1	5'-CTGCTGGTCCTGCTGGTC-3'	5'-CCTTGTTCGCCTGTCTCAC-3'
FN	5'-CCAGTTCAGAGGAGCATCAG-3'	5'-GGCATTGTCGTTCAGAGTGTA-3'
MMP3	5'-TGACGATGATGAACGATGGA-3'	5'-CCTTGGCTGAGTGGTAGAG-3'
TGF-β	5'-AATTCCTGGCGTTACCTT-3	5'-CCTGTATTCCGTCTCCTT-3'
FGF2	5'-CTACAACTCCAAGCAGAAGA-3'	5'-GTTATTAGATTCCAGTCGTTCAA-3'
IL-1β	5'-AGTGATGAGAATGACCTGTT-3'	5'-GATACTGCCTGCCTGAAG-3'

***Table 6**: primer sequences used for evaluating the expression profile of COL1A1, fibronectin FN, matrix metalloproteinase 3 MMP3, transforming growth factor TGFβ, fibroblast growth factor FGF2, and interleukin IL-1 β in murine fibroblasts seeded on* C. reniformis*-derived 2D collagen coatings, using GAPDH as a housekeeper gene.*

5.2 Results

5.2.1 *Chondrosia reniformis* Fibrillar Extract Extraction and Purification

Using the extraction protocol reported in the Materials and Methods it was possible to obtain a fibrillar suspension with a main yield of 35% ± 1.5% (SD) of the total dry weight of the sponge tissues. Therefore, from 1 g of dried *C. reniformis* is possible to produce on average 0.35 g of ECM fibrillar extract. Considering that this suspension contains 56.81% of collagen, the total yield in collagen from the starting material results equal 19.9%.

5.2.2 2D Collagen Membrane Permeability Tests

As explained in the Materials and Methods, *C. reniformis*-derived 2D-CMs permeability to deionized water and proteins was evaluated over one week (see Figure 42).

5.2.2.1 Water Permeability Test

In the so-called "dry-wet" condition (with just one side of the membrane put in contact with water, resembling a dry wound), a visual monitoring denoted that the amount of water able to pass through the 2D-CMs after the stated period was equal to zero. In the "wet-wet" conditions (with both sides of the membrane put in contact with water, recreating a moist wound), after 1 day less than 0.1% ± 0.0043% (SD) of water was able to pass through the membrane, with a main permeability of 0.002 mL/cm^2, while after 7 days only 0.5% ± 0.0078% of water inside the upper well flowed to the lower well, with an average permeability of 0.01 mL/cm^2 (see Table 7). Hence, these 2D-CMs can be considered an effective liquid-retaining barrier.

Time Point	"Dry-Wet" Condition	"Wet-Wet" Condition
0 h	0	0
1 h	0	0
3 h	0	0
6 h	0	0
24 h	0	0.1 ± 0.0043
48 h	0	0.15 ± 0.27
7 days	0	0.5 ± 0.0078

***Table** 7: percentage (%) of water able to pass through the 2D-CMs, with relative standard deviations (SD), at different time points and in different experimental conditions.*

5.2.2.2 Protein Permeability Test

For protein permeability, less than the 0.03% (mean value: 0.01 mg/mL ± 0.0035 (SD)) of bovine serum albumin (BSA) (chosen as standard by following the example of [Ferrario et al., 2021]) passed through the membrane after 7 days, as shown in Figure 43. Thus, in this range of time, the 2D-CMs are almost impermeable to proteins.

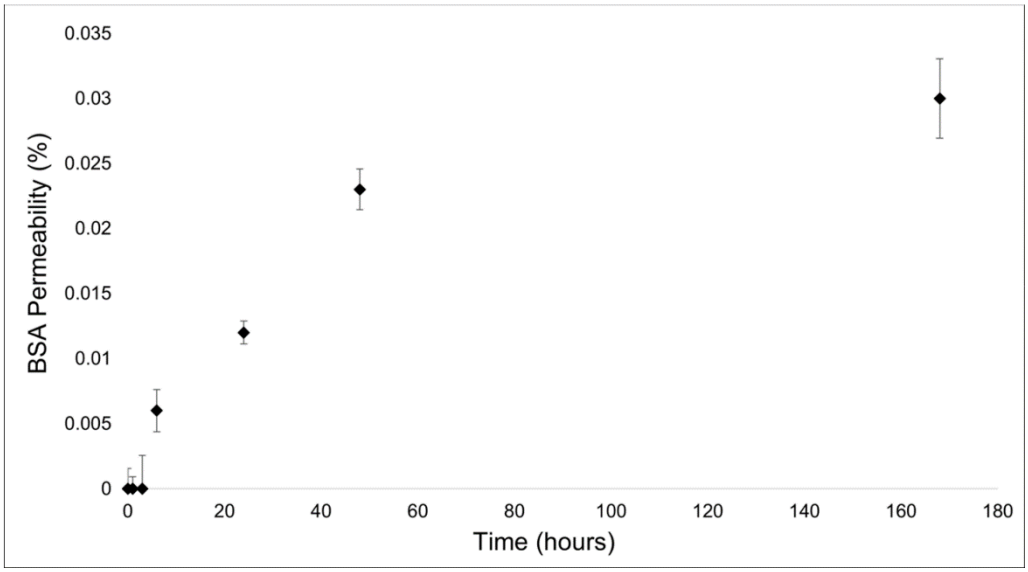


Figure 43: percentage of bovine serum albumin (BSA) passing through the 2D collagen membranes at specific time points. Time is expressed in hours. The graph shows mean values ± standard deviation (SD).

5.2.3 Bacteria Infiltration Tests

The efficiency of *C. reniformis*-derived 2D-CMs in preventing bacterial infiltration into a wound was evaluated as described in the Materials and Methods Section, by using three different bacterial species known to be common opportunistic pathogens in nosocomial environments, where often cause infections in hospitalized patients. The fibrillar network obtained from a 3 mg/mL collagen suspension was dense enough to retain almost all bacteria, reaching 100% retained bacteria for *S. aureus* in each experiment. Overall, less than an average of 0.00045% ± 0.0003 (SD) of the initiating bacteria (10^7 cells/mL) was able to pass through the membrane (see Table 8). In all cases, bacteria in the upper layer of the membrane remained completely viable, stating that the low number, or even the absence, of CFU from the solution taken from the lower well is due to an actual scarcity/absence of bacterial cells, rather than to a mortality phenomenon of the 2D-CMs; therefore, they acted as an extremely efficient barrier against bacteria infiltration.

Experiment	*S. aureus*	*P. aeruginosa*	*E. coli*
1st	0	0	0
2nd	0	30	20
3rd	0	10	120
Mean	0	20	70
Infiltrated bacteria (%)	0	0.0002	0.0007
±SD	0	0.000153	0.000643
Retained bacteria (%)	100	99.9998	99.9993

***Table 8**: S. aureus, P. aeruginosa, and E. coli infiltration. As described in the Materials and Methods, the bacterial suspensions were loaded onto the 2D-CMs at a concentration of 10^7 CFUs. The numbers indicate the CFUs that were recovered in the flow-through, therefore representing the bacteria that successfully crossed the 2D-CMs within a 48 h period. The percentage (%) of infiltrated bacteria, along with their respective standard deviations (SD), must be considered in relation to the initial bacterial concentration of 10^7 CFUs used in the experiments.*

5.2.4 Bacteriostaticity Test

As already stated in previous works, the membranes obtainable from the collagen extracts of *Chondrosia reniformis* have interesting properties, such as antioxidant activity. Additionally, it has been observed that the liquid fibrillar ECM extract can be conserved at 4 °C for a long time without showing any sign of bacterial contamination or degradation, even if kept in non-sterile conditions (operator observations). To evaluate their potential as SLSs and consider the susceptibility of injured skin to bacterial infections, the hypothesis by which this biomaterial could be intrinsically bacteriostatic was examined. However, 2D-CMs do not seem to have any kind of effect on slowing bacterial growth, at least in the chosen experimental conditions. Table 9 reports the average number of bacterial CFUs left following 6 h in agitation with and without the presence of a 2D-CM. Since no significant difference was found between samples, it is possible to state that the *C. reniformis*-derived fibrillar ECM extract in the form of a 2DCM does not affect the growth rate of common bacteria such as *S. aureus, P. aeruginosa,* and *E. coli*.

Sample	*S. aureus*	*P. aeruginosa*	*E. coli*
Control	5.6 × 108 ± 1.9 × 108	1.28 × 108 ± 1.9 × 108	1.41 × 108 ± 2.08 × 108
Presence of 2D-CM	6.1 × 108 ± 1.6 × 108	1.9 × 108 ± 0.7 × 108	1.47 × 108 ± 2.41 × 108

***Table 9**: Evaluation of the bacteriostatic activity of 2D-CMs on* S. aureus*,* P. aeruginosa*, and* E. coli *after a 6 h incubation in LB media, expressed with relative CFUs/mL registered at the end of the experiment.*

5.2.5 In Vitro Biodegradability Test

An initial evaluation of the degradation rate of 2D-CMs in PBS or collagenase solution was conducted by quantifying the percentage of their weight loss over time, as described in the Materials and Methods. Figure 44A shows the percentage of weight loss of 2D-CMs over 21 days in PBS and in collagenase solution, respectively. Both types of samples exhibited the same trend of degradation, with a major weight loss within the first 7 days, which tended to slow down over time, while still slowly progressing. However, given the collagenous nature of the biomaterial, the 2D-CMs in collagenase solution reach higher levels of degradation more rapidly, even if *C. reniformis*-derived collagen tends to be refractory to the enzymatic attack of common collagenases beyond a certain level. After 21 days, the samples in PBS solution on average lost 31.6% ± 0.11% (SD) of the weight, while reaching 44.2% ± 0.10% (SD) in collagenase solution in the same amount of time. Subsequently, the incubation media of the *in vitro* biodegradability test were analyzed to better understand which kind of macromolecules were released, causing the weight loss registered for the 2D-CMs. Figure 44B shows the release of total proteins within the media, while Figure 44C and Figure 44D represent collagen-related weight loss and the glycosaminoglycans-related weight loss, respectively. Again, in all cases, the trend is similar for both types of samples, but with higher values for those that were soaked into a collagenase solution. After 21 days, the major weight loss (up to 28.5% ± 5% (SD)) of the total weight loss in PBS samples, and up to 42.2% ± 8.5% (SD) in collagenase samples is attributable to total proteins (Figure 44B), of which collagen represents 4% ± 0.81% (SD) and 6% ± 0.9% (SD) in PBS and collagenase samples, respectively (Figure 44C). Glycosaminoglycans (GAGs), in Figure 44D, represent 13.2% ± 3.86% (SD) and 21.42% ± 2.93% (SD) of the total material released from the 2D-CMs soaked in PBS and in collagenase solution, respectively. It appears that a certain percentage of ECM components (roughly, 40% for membranes

soaked in PBS solution and 20% for those in collagenase solution) were responsible for the observed weight loss of the 2D-CMs, but this was not wholly identified.

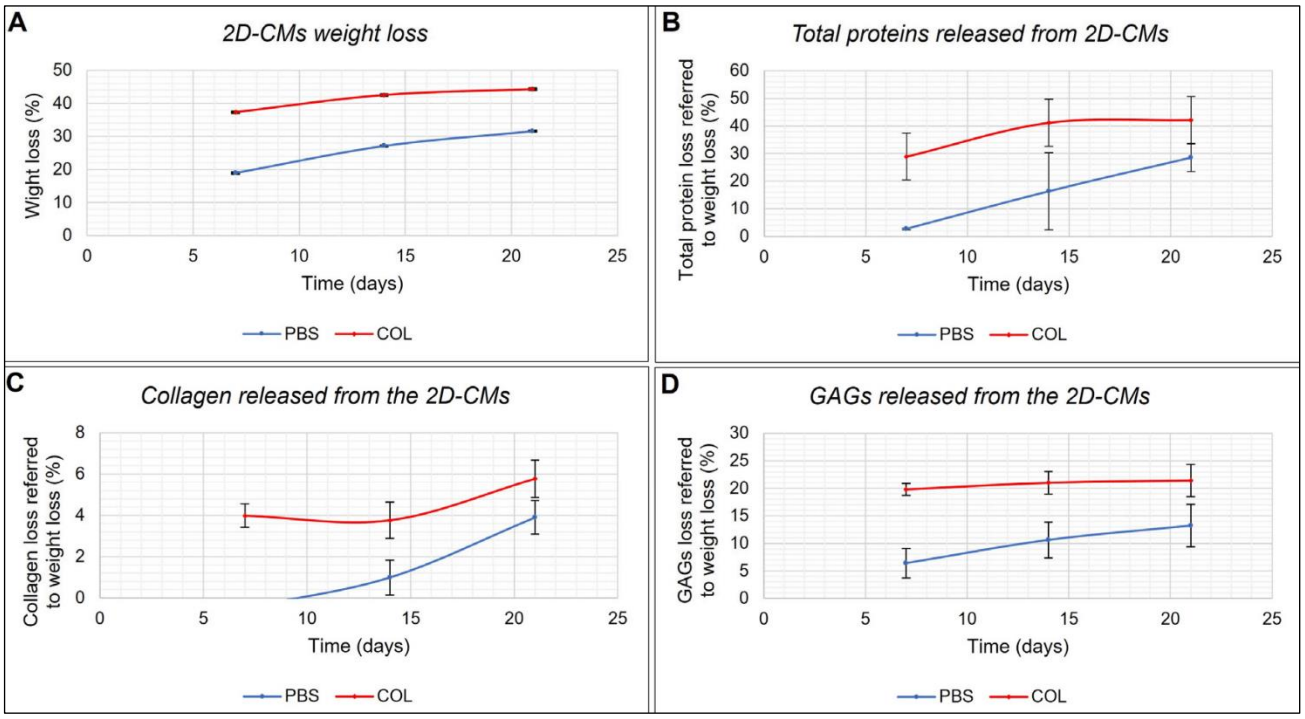


***Figure 44**: In vitro biodegradability test. (**A**) Percentage of degradation of the 2D-CMs in PBS and in 0.1 mg/mL collagenase solution calculated as 2D-CMs dry weight difference. The incubation media were used to determine the percentage of total proteins released from the 2D-CMs (**B**), the percentage of collagen release (**C**), and the percentage of GAGs release (**D**) referred to the total weight loss of the membranes.*

5.2.6 Effect of the 2D Collagen Membranes on Fibroblasts

Fibroblasts constitute many tissues in the body and are mainly responsible for the production and turnover of ECM fundamental constituents. During tissue repair processes, fibroblasts can assume a contractile phenotype involved both in increased ECM production and contraction phenomena. In a normal wound healing process, ECM molecules need to be rapidly synthesized for effective tissue remodeling, and their biosynthesis must be finely regulated. Excessive deposition of connective tissue could generate hypertrophic scars and keloids; therefore, a correct balance between connective tissue synthesis and breakdown is strictly necessary. Normally, this balance is controlled by the production of mediators like cytokines and growth factors, such as transforming growth factor (TGF-β) family, interleukins (such as IL-1β), tumor necrosis factor (TNF), the fibroblast growth factor (FGF) and others. To evaluate the response of fibroblasts in the presence of a 2D cm scaffold and, in particular, to assess if the 2D-CMs could induce the up-regulation of ECM production-/remodeling-related genes, α1-chain of collagen type I (COL1A1), fibronectin and metallopeptidase-3 (MMP3) gene expression profiles were evaluated by qPCR in L929 mouse fibroblasts growing on plates coated with *C. reniformis*-derived ECM fibrillar extract. On the same cells, a gene expression analysis was also performed to evaluate any kind of imbalance in the above-mentioned ECM regulating factors comparing them to the controls. Table 10 shows the gene expression levels of fundamental ECM-related genes in L929 fibroblast growing on 2D-CMs for 24 or 120 h. The gene expression levels of controls were considered normalized to 1. Gene expression of COL1A1, FGF, MMP3, and IL-1 does not show significant differences between controls and 2D cm coated samples,

neither after 24 h nor after 120 h. Conversely, fibronectin gene expression after 120 h, equal to 1.46 ± 0.13 folds, increases significantly, showing a 46% enhancement when compared to the uncoated controls. Instead, TGF-β gene expression after 24 h, equal to 0.79 ± 0.05 folds, is lower in 2D cm coated samples than in uncoated controls and diminishes by 21%.

Sample	**COL1A1**	**FN**	**MMP3**	**TGF-β**	**FGF**	**IL-1β**
24 h	0.94 ± 0.05	1.23 ± 0.2	1.25 ± 0.23	0.79 ± 0.05 *	0.66 ± 0.28	1.08 ± 0.21
120 h	1.27 ± 0.20	1.46 ± 0.13 *	1.20 ± 0.55	0.94 ± 0.24	0.76 ± 0.16	1.30 ± 0.25

Table 10*: the gene expression level of some of the main ECM-related genes: α1-chain of collagen type I (COL1A1), fibronectin (FN), metallopeptidase-3 (MMP3), transforming growth factor (TGF-β), fibroblast growth factor (FGF), and interleukin 1β (IL-1β) on L929 mouse fibroblasts growing on C. reniformis-derived 2D-CMs for 24 and 120 h. Significant differences between gene expression levels were calculated by comparing the two types of samples to an uncoated control. Asterisks indicate a significant difference versus the respective control (paired Tukey test, * = $p < 0.05$).*

USE OF COLLAGENASE-RESISTANT MARINE COLLAGEN FROM *CHONDROSIA RENIFORMIS* TO PRODUCE 2D-FILMS CONTAINING COLLAGENASE IMMOBILIZED IN BIO-SILICA NANOPARTICLES

Enzymatic removal of necrotic tissue resulting from skin wounds remains one of the preferred methodologies to promote tissue regeneration and healing. Enzymatic treatment is more effective compared to physical removal through surgical or mechanical intervention, as it is much more selective towards collagen, elastin, and fibrin present in devitalized tissues, thus preserving the structure of vital tissues more effectively. Moreover, this approach proves to be significantly more effective even in situations caused by skin damage from ulcers and injuries of various aetiologies, possibly non-infected. Currently, there are various formulations of creams and ointments containing enzyme mixtures that promote selective biodegradation of necrotic tissues. One of the key components in these formulations is collagenase, an enzyme crucial in degrading collagen structures but characterized by poor stability in aqueous environments and in the presence of metal ion chelators, essential for its functioning. In addition to the enzymatic component, maintaining the correct moisture level in the wound environment is also essential. The correct moisture level involves avoiding premature drying of the surface on one hand, and tissue maceration on the other. This additional requirement is usually met using various types of patches and dressings that allow proper wound transpiration while maintaining an appropriate moisture level. Adequate moisture also facilitates the removal and replacement of dressings during dressing procedures without causing lacerations to the newly formed tissue, episodes that can occur in case of excessive drying of the surfaces, with a risk of infections and pain during dressing removal. Nowadays, the market offers numerous types of dressings and patches that, combined with the creams and ointments mentioned above, allow for sustainable management of the wound healing process, albeit with significant room for improvement in both healing times and overall effectiveness of the approach. One of the ideal components for the production of patches is collagen, especially due to its high biocompatibility. However, one limiting factor in the use of collagens is their association with collagenases, as these enzymes normally degrade collagen. Conversely, the poor stability of collagenases remains an aspect that can be improved in the management approach of wound tissue regeneration. For this reason, in this part of the doctoral research activity, the collagenous extract obtained from *Chondrosia reniformis* was used to create a composite biomaterial with the aim to realize an innovative wound dressing device able to

enhance wound healing. Using biosilica as a linking agent, collagenase molecules were encapsuled and stabilized on the surface of *C. reniformis*-derived 2D collagen membranes; in fact, as extensively explained in the introduction section, the collagen of this sponge is highly resistant towards the attacks of commercial collagenases [Imhoff and Garrone, 1983], thus addressing a critical limitation in classical collagen-based dressings; furthermore, as stated in the previous paragraphs, it is still resorbable in the long-term. Silica particles were obtained using a sponge-derived silicatein, which is a silica-synthesizing hydrolase [Shimizu and Morse, 2018]; the silica component, besides serving as an enzyme-immobilization factor, has potential regenerative properties itself. Moreover, the presence of collagen-associated polysaccharides in the ECM extract of *C. reniformis* could also improve the maintenance of a proper hydration level in the interested treated area.

5.3 Materials and methods

Unless differently stated, all reagents were acquired from Sigma-Aldrich (Milan, Italy).

5.3.1 Collagenase encapsulation and immobilization on collagen fibrils

1 mg of commercial collagenase from *Clostridium histolyticum* (125 CDU/mg), 6 μg of *Petrosia ficiformis*-derived silicatein, 0.15 mg of *C. reniformis* collagen fiber extract (obtained as reported in the previously described work), 440 mM tetraethyl orthosilicate (TEOS), and 50 mM Tris-HCl buffer, pH 6.8, were incubated in a final volume of 0.5 mL for 3 hours at 4° C under continuous orbital agitation; after the set period of time, the sample was centrifuged at 20,000 x G at 4° C for 15 minutes. The precipitate of collagen fibrils, now containing silica and collagenase, was then washed twice with 0.5 mL of deionized H_2O, and finally suspended in 0.1 mL of a 50 mM tricine, 10 mM $CaCl_2$, and 400 mM NaCl, pH 7.5 buffer. The successful immobilization of collagenase in the fibril-silica composition obtained through the enzymatic condensation action of TEOS by silicatein was confirmed by analysing the absorbance at 280 nm of the supernatant obtained after centrifugation at the end of the 3 hours incubation. In addition, a protein content assay using the Bradford method [Bradford, 1976] was performed on the same sample to confirm the data in comparison with a negative control, where silicatein was not added (results not shown).

5.3.2 Production of composite collagen-silica-collagenase 2D membranes

3 mL of a 3 mg/mL *C. reniformis* collagen suspension were poured into 25 x 28 mm silicon moulds; right after, three pellets obtained as described before were incorporated in the suspension. The composite was let dry at 14° C for 48 hours and conserved at 4° C.

5.3.3 Characterization of composite collagen-silica-collagenase 2D membranes

5.3.3.1 Collagenase activity assay

To verify the integrity and availability of the immobilized collagenase in the composite membranes, the enzyme activity was evaluated using the FALGPA method [Van Wart and Steinbrink, 1981]. This assay is based on the principle by which the oligopeptide FALGPA (2-furanacryloyl-l-leucylglycyl-l-prolyl-l-alanine) is selectively and rapidly hydrolysed by *Clostridium histolyticum* collagenase, while is not cleavaged by any of other well-known proteinases such as trypsin, thermolysin, or elastase; the presence of intact FALGPA is spectrophotometrically detectable, and its hydrolysis causes a reduction of absorbance, thus being an index of collagenase activity. The composite membranes were re-hydrated with distilled water and put into a quartz cuvette together with a FALGPA solution; the absorbance was continuously measured for 10 minutes at a wavelength of 340 nm against a blank and

the results were compared with a negative control constituted by a 2D membrane where no collagenase pellet was incorporated.

5.3.3.2 Environmental Scanning Electron Microscope (ESEM) Observation

For ESEM observation, 2D composite membranes were firstly completely dehydrated by soaking them in a series of alcoholic solutions with an increasing concentration of ethanol up to 100%, then graphite was covered and examined. Images of the samples were observed and acquired with an ESEM Vega3–Tescan, type LMU (Tescan Brno s.r.o., Brno, Czech Republic) equipped with a microanalyzer system EDS-Apollo_x and EDS texture and elemental analytical microscopy software (TEAM^{TM} Analysis System, version 1.0, Coherent Scientific, Thebarton SA). Observation and acquisition of the four SCFMs images were performed with a FESEM Zeiss SUPRA 40 VP (Carl Zeiss AG, Oberkochen, Germany) and its associated software.

5.3.3.3 Thermogravimetric analysis (TGA)

To assess the percentages in which each component (in particular, collagen, silica, and loaded enzyme) is present within the 2D composite membranes, a thermogravimetric analysis (TGA) was performed. TGA was conducted on a membrane produced as described above, with a thermal ramp of 10 °C/min in a temperature range between 25 °C and 900 °C in presence of oxygen, using a TGA/DSC 1 STAR from Mettler Toledo.

5.3.3.4 Fourier transform infra-red (FTIR) analysis

The chemical composition of the composite membranes was investigated performing a Fourier Transform Infrared spectroscopy analysis, that was conducted in ATR (Attenuated Total Reflectance) mode over the frequency range of 400-4000 cm^-1 using a Vertex 70 spectrometer from Bruker.

5.3.3.5 Absorbance capacity

The water binding capacity of *C. reniformis*-derived 2D collagen membranes has been evaluated by [Pozzolini et al., 2018], which assessed that non-crosslinked *C. reniformis*-derived 2D-CMs have a high degree of hydrophilicity. This suggest that this biomaterial could be also capable of adsorbing compounds useful in the treatment of skin injuries, such as disinfectant, anaesthetic, and analgesic compounds. The adsorption capacity was determined for two active ingredients by immersing the composite membrane in 20 ml of methylene blue solution (MB, 20 mg/L) and ketorolac tromethamine (KT, 5 mg/L) for 15 hours under ambient conditions. This was followed by ten washes in deionized water. The adsorption of the active ingredients was then verified and quantified using optical spectroscopy in the spectral absorption range of the two molecules (400-800 nm for MB and 250-400 nm for KT).

5.4 Results

5.4.1 Collagenase Activity Assay on Composite Membranes

The presence of active collagenase on the composite 2D-CMs was assessed with a FALGPA assay, in comparison with a negative control constituted by a membrane on which no enzyme immobilization was performed. Figure 45 show the absorbance of FALGPA at 345 nm during the 10 minutes assay. In the negative control (Figure 45A), no significant variation in the absorbance values was detected; on the other hand, the composite 2D-CM loaded with collagenase determined a

decrease of absorbance during time (Figure 45B), confirming the functional presence of the enzyme on the membrane.

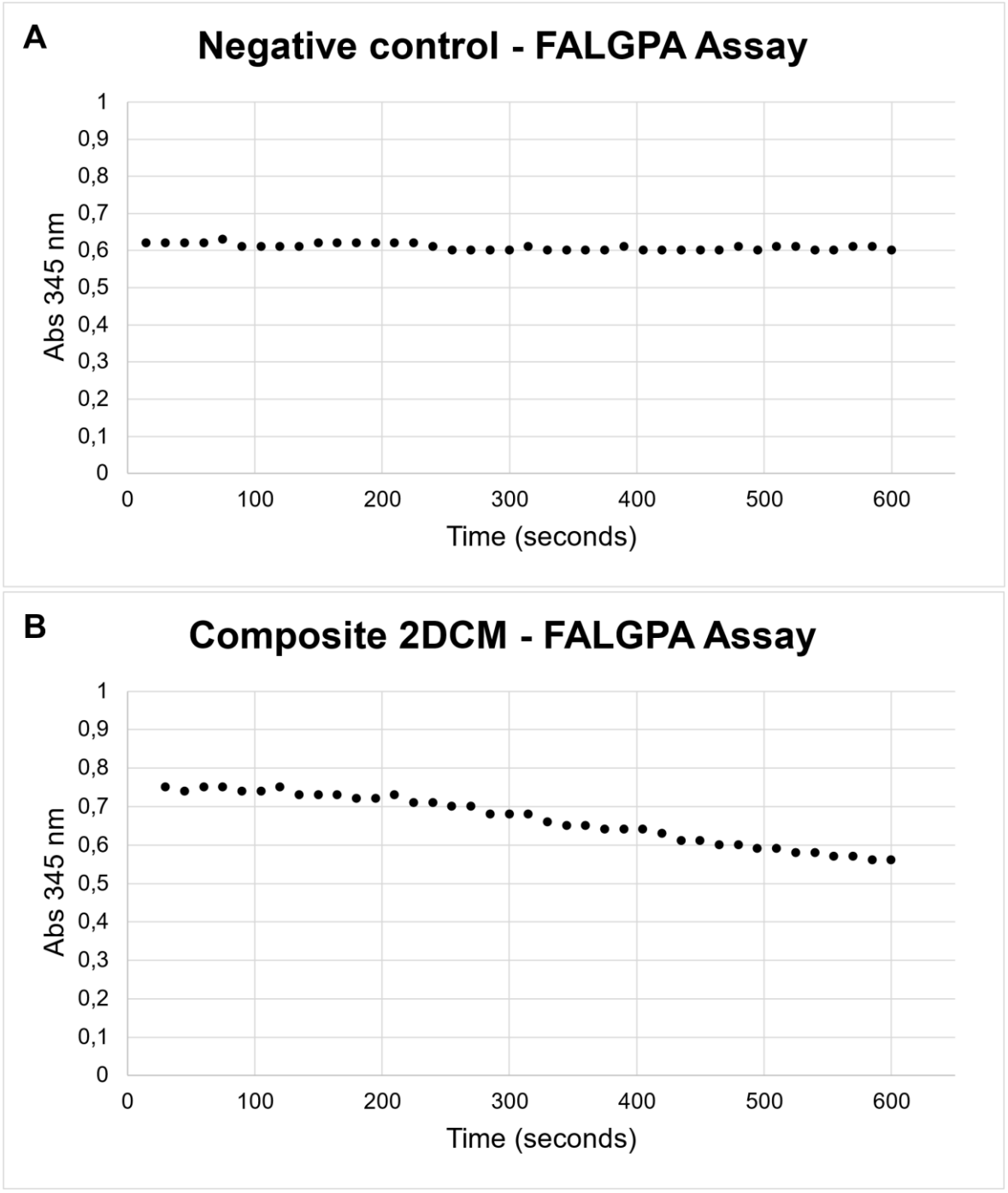


***Figure 45**: absorbance of FALGPA in presence of a negative control (panel A) and in presence of a collagenase-loaded 2D-CM (panel B).*

5.4.2 Environmental Scanning Electron Microscope (ESEM) Observation

The ESEM analysis (Figure 46) confirmed the presence of silica particles of variable dimensions on the 2D membranes, which appear incorporated within the fibrillar reticules.

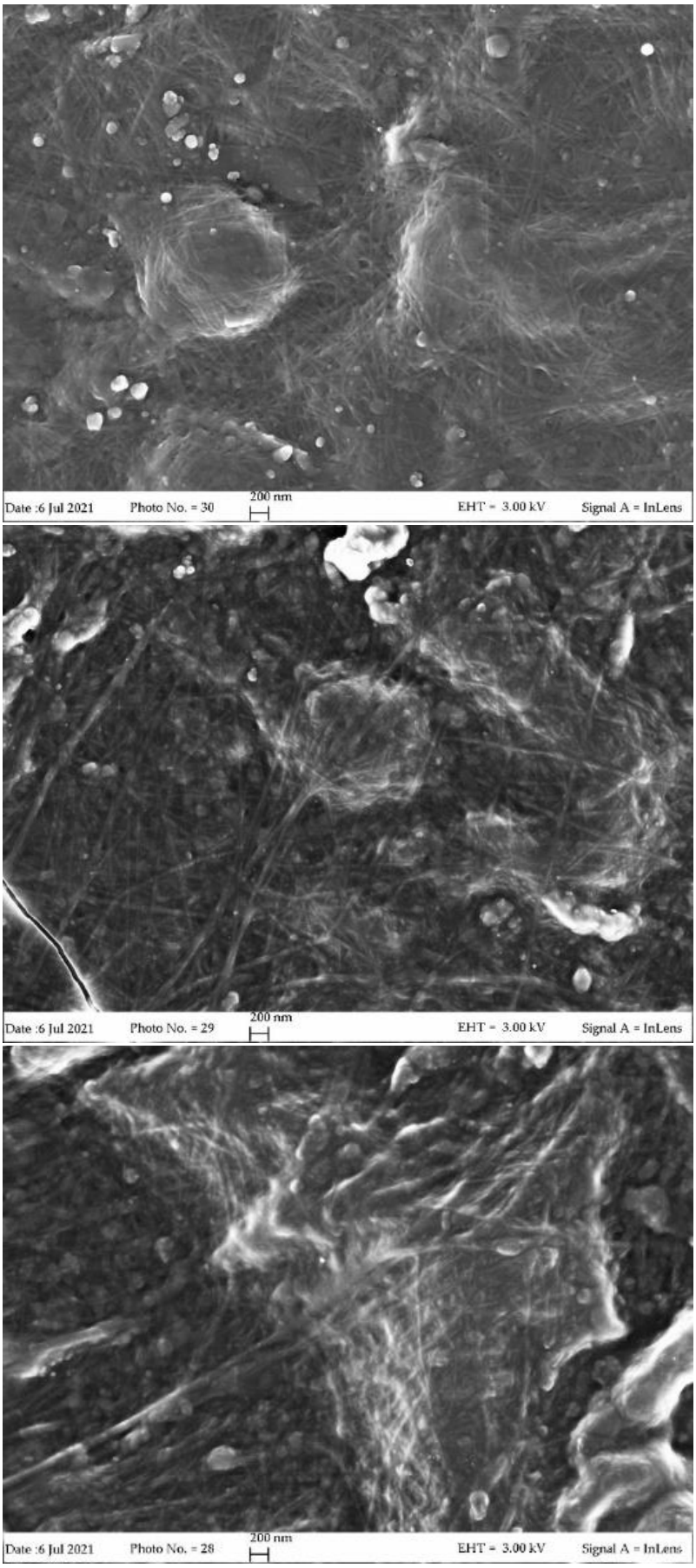


***Figure 46**: ESEM images of the composite C. reniformis-derived 2DCMs.*

5.4.3 Thermogravimetric analysis (TGA)

The TGA analysis results are shown in Figure 47. The sample mass decreases with complex behaviour characterized by the presence of multiple processes. In the range between 25 and 130°C, there is an initial mass decrease around 13% of the total (black line). This variation is associated with the release of water and ethanol adsorbed on the surface. Ethanol is a by-product of the enzymatic conversion of tetraethoxysilane, which, like water, is adsorbed into the material through weak bonds (hydrogen bonds and dipole-dipole interactions) with both collagen and silica, rich in surface hydroxyl groups. Between 100 and 430 °C, the organic components of the system are degraded. The mass loss between 130 and 250 °C is associated with the lighter component of the sample, around 11% of the total, which is associated with the TRIS buffer, with a boiling point of 219 °C. Its molecule, rich in hydroxyls, can be effectively adsorbed by the different components of the system through weak dipole-dipole and hydrogen bonds and can also participate in the condensation reaction of alkoxysilane. Between 250 and 430 °C, a mass loss of 31% is associated with the heavier free organic component, presumably collagen and polysaccharides. Note that the first derivative of the thermogram (grey line) shows at least two peaks in this temperature range at 276 and 320 °C, probably associated with two different components. Finally, between 430 and 600 °C, a massive loss of 38% is associated with the conversion of partially condensed alkoxysilane in the matrix with the release of ethanol and other organic components bound to the silica matrix during the condensation process. Such behaviour, where the organic component bound to silica degrades even up to 900°C, has also been shown in literature [Das et al., 2013]. The residual material at temperatures above 600 °C (8%) is associated with the inorganic component of the sample, namely silica, corresponding to a Silicon content slightly below 4%.

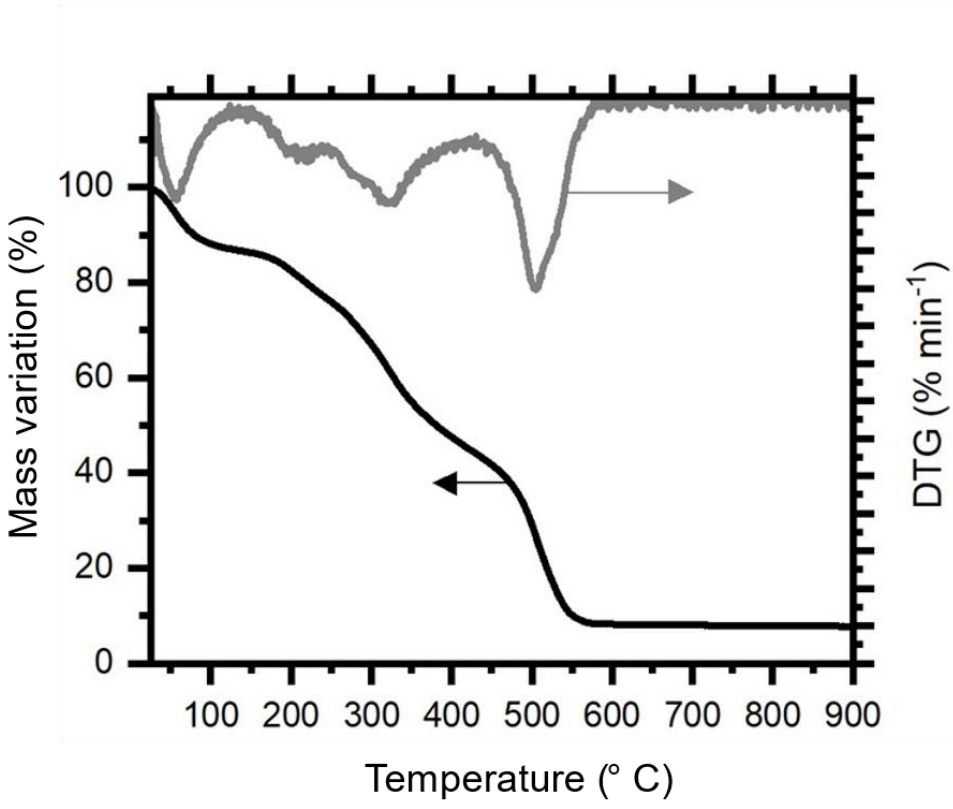


***Figure 47**: thermogravimetric analysis (TGA) of composite collagenase-loaded 2D-CM.*

5.4.4 Fourier transform infra-red (FT-IR) analysis

Figure 48 shows the results of the FT-IR analysis performed on the collagenase loaded 2D-CM. The spectra were collected from three portions of the same sample, demonstrating its relative homogeneity, as only minor differences in intensity are present. All spectra exhibit a highly structured signal in the range of 3800-2500 cm^{-1} associated with the stretching of O-H and N-H bonds. These bonds are abundant in the proteinaceous material as well as in the inorganic component; silica is indeed generally characterized by surface hydroxyl groups. Overlapping this signal is a cluster of peaks in the region of 2800-3000 cm^{-1} associated with the stretching of C-H bonds, with the same type of bond associated with the peak at 1451 cm^{-1} related to the bending of C-H2. The presence of silica is confirmed by the intense signal at 1053 cm^{-1} with a shoulder at 1120 cm^{-1} associated with the stretching of the Si-O-Si bond, and by the peak at 445 cm^{-1} associated with the rocking of the same bond. The high structuring of the peak at 1053 cm^{-1} indicates the presence of other species. Indeed, peaks in the region between 1000 and 1300 cm^{-1} (1026, 1198, and 1281 cm^{-1}) can be associated with the stretching of C-O-C bonds of polysaccharides present in the original extracellular matrix extract of the sponge, with which the membranes were created.

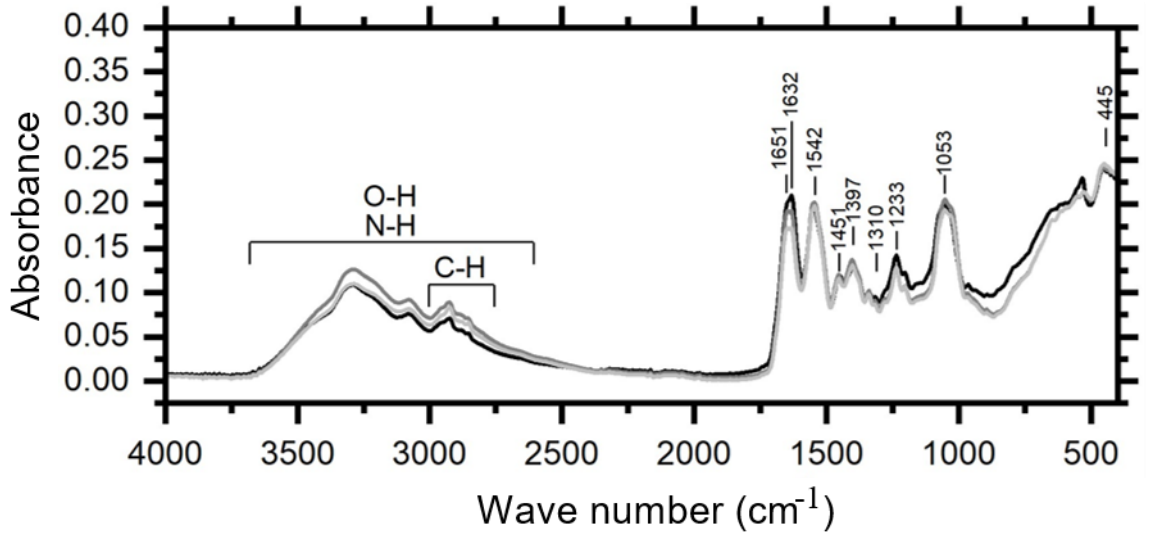


Figure 48: FT-IR analysis of composite collagenase-loaded 2D-CM.

5.4.5 Absorbance Capacity

Figure 49 shows the absorption spectrum of the virgin composite (grey) and after treatment in MB solution or Ketorolac thrometamine (black). The former shows no absorptions in the investigated range but only an increase in the spectral background as the wavelength of the incident radiation decreases, characteristic of light scattering phenomena. After treatment in BM, the material exhibits a clear absorption peak at 556 nm with a shoulder at 616 nm, characteristic of the adsorbed molecule. The dashed line in Figure 49A shows a spectrum calculated as the difference between that of the sample containing BM and that of the virgin sample, confirming the absorption of the molecule. From this absorption, it is possible to derive a value of the concentration of adsorbed material, which is around some $\mu g/cm^3$. Regarding the second active ingredient, it was analyzed in the spectral range of 250-500 nm, where the KT molecule shows two absorption peaks at about 320 and 250 nm (Figure 49B). As in the previous case, in this range, the spectral background shows an increase as the wavelength decreases. The virgin sample also shows intense absorption starting from about 300 nm.

The spectrum of the sample treated with KM is largely overlapping with these, with increases in the absorption value around 315 and 250 nm. The difference between the two spectra (dashed line) reveals indeed a maximum peak at 250 nm and a second structured one at about 315 nm, corresponding to the absorption of the active ingredient. In this case, the concentration of adsorbed material is around tens of $\mu g/cm^3$.

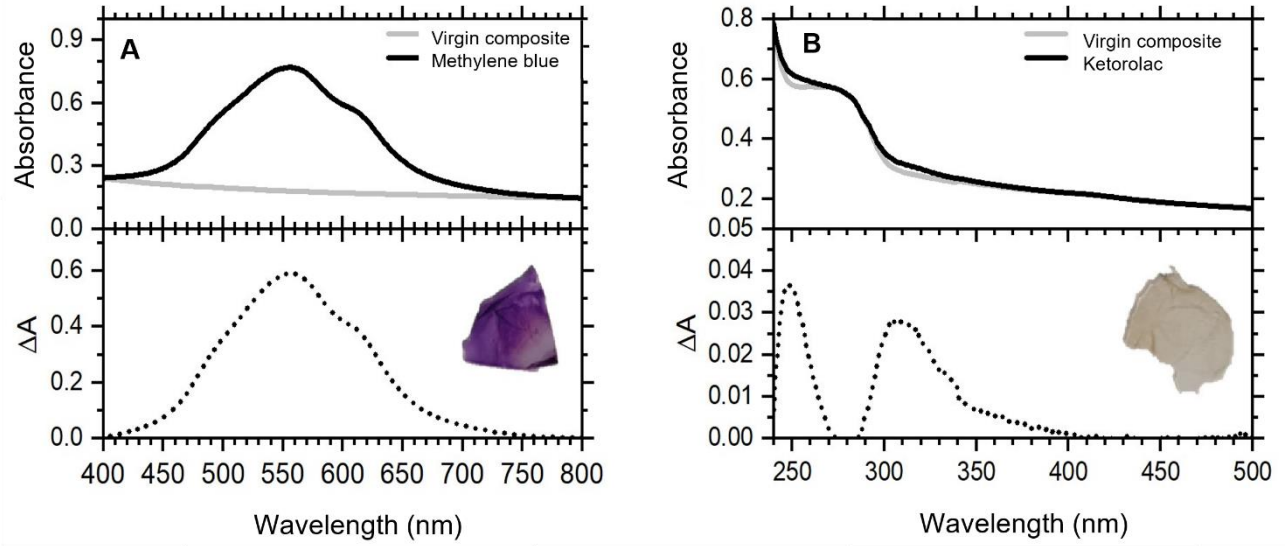


Figure 49: absorption spectrum of the virgin composite (grey) and after treatment in MB solution or Ketorolac tromethamine (black).

5.5 Discussion

Collagen is one of the most employed natural polymers for the production of skin substitutes. It has been reported that some collagens of marine origin display unique chemical and physical properties and are suitable for regenerative medicine. Among them, collagen from marine sponges has been previously suggested as a viable biomaterial for constructing scaffolds intended for bone regeneration. Conversely, there is still a lack of knowledge about the usage of Porifera-derived collagens to produce skin substitutes. This part of the doctoral research activity aimed to evaluate the suitability of 2D membranes obtained using an intact fibrillar extract derived from the marine sponge *C. reniformis* for skin regenerative medicine purposes. Using the extraction procedure described in the Materials and Methods Section, it was possible to averagely obtain 0.35 g of ECM fibrillar extract from 1 g of dried sponge tissue. This result indicates that increasing the number of extraction cycles from two, as previously described, to six, i.e., until reaching the complete dissolution of the sponge body, does not seem to increase the extraction yield. The collagen fibers obtained in the first isolation cycles may be more insoluble and easier to recover than the ones extracted in the last cycles, and, therefore, could have been lost during the repeated washing steps. As stated in 5.2, the total yield of collagen starting from fresh tissue is equal to 19.9%. The isolation of collagen from the same organism using a carbon dioxide extraction approach provided a yield of only 10%, aside from it being faster. Given the richness of the collagen in this animal, with the solvent-free extraction procedure described here it is possible to obtain an overall good quantity of intact collagen fibers. Being a purification method based on trypsin digestion and scaling up the extraction system with a view to a develop a circular economy, it could be possible to replace the

more expensive purified trypsin with fish stomach extracts deriving from waste material from fish processing. When 2D membranes are designed as skin substitutes, they must ensure limited evaporation of water and loss of proteins from the injured area. For the *C. reniformis* derived 2D-CMs, no water passage was registered in the "dry-wet" condition, whereas when a moist wound was recreated in the "wet-wet" conditions a main permeability of only 0.01 mL/cm^2 was detected after 7 days. The water permeability across the collagen membranes is affected by porosity, pores diameter, cross-links level, and cross-links hydrophilicity. An increased diffusion coefficient is related to a reduced cross-linking state. The strong water insolubility observed in the *C. reniformis* fibrillar extract is evidence of a high level of interchain crosslinking, which allows the obtaining of 2D membranes through directly casting the collagen extract without any chemical crosslinking steps. The high crosslinking rate gives this biomaterial a reduced permeability to water. The water permeability rate detected in *C. reniformis* collagen-derived membranes was about three-fold lower than those evaluated in 2D membranes derived from sea urchin collagen. Although these two types of membranes are similar in composition, the reduced water permeability detected in this study can be ascribed to a higher polymer concentration used for casting the membranes. Even if a biomaterial for wound dressing must guarantee a limited loss of water during healing, a certain permeability level is necessary, as excessive water retention may cause wound exudate accumulation. Therefore, to obtain optimal water permeability performance in 2D membranes designed as skin substitutes, the polymer concentration and, accordingly, its porosity level must be modulated properly. Moreover, *C. reniformis*-derived 2D-CMs showed low permeability to proteins, preventing the loss of growth factors, and ensuring maintenance of the optimal microenvironment in the regeneration region. *C. reniformis*-derived 2D-CMs proved to be a strong barrier against bacterial infiltration, against both Gram-positive and Gram-negative strains. This seems to be particularly evident for Gram-positive strains, where the percentage of bacteria retention is 100%. Although the sponge-derived 2D-CMs constitute a thick barrier that prevents the infiltration of bacteria into the injured area, the contact of the tested microorganisms with the membranes' surface did not affect their viability; furthermore, the 2D-CMs do not seem to release any soluble substance with a bacteriostatic activity. Sponges are populated by numerous symbiotic microorganisms; therefore, across evolution, they had to develop a manifold molecular strategy to contain bacteria over-growth, and, for this reason, marine sponges are important sources of natural compounds with antibiotic and antiviral action. However, the extraction procedures used in the present study allowed us to obtain an extract of intact collagen fibers, free from any sponge-derived secondary metabolites. As poriferan natural compounds exhibit antitumor and cytotoxic activities, the previously attested biocompatibility of 2D-CMs on *in vitro* cell lines further confirms this issue.

Biomaterials for tissue engineering must display a low degradation rate but, on the other hand, once the regenerative process is finished, they must be resorbable. Previous studies have demonstrated the peculiar resistance of *C. reniformis*' collagen extract to collagenase digestion. Furthermore, after 6 days in collagenase solution, the *C. reniformis* derived 2D-CMs remained intact if compared to commercial membranes. In this study, a 2D-CMs time-course stability test within 21 days was performed. Although within 21 days of incubation the sponge-derived 2D-CMs seemed to maintain their integrity, a certain percentage of weight loss was detected in both saline solution and collagenase solution. The higher degradation rate observed in collagenase-treated samples indicates that this biomaterial is not completely resistant to enzyme digestion and, hence, is potentially resorbable, even if its long-term resistance to collagenases was confirmed, as the membranes lost some of their weight, but they didn't dissolve as happens with commercial collagen membranes. A remarkable contribution to membrane weight loss derived from GAGs: the amount of glycosaminoglycans released in the medium was constant and directly proportional to the weight loss

along the entire observation time, as well as being higher in the presence of collagenase activity. This suggests that these sulphurated polysaccharides are intimately associated with collagenous proteins that, once enzymatically attacked, cause the GAGs release. If the GAGs adhesion to the 2D-CMs is subjected to collagen proteins, the collagen release is also affected by GAGs. After seven days, when the main percentage of GAGs was released, a significant increase in collagen and non-collagen proteins was detected. It was experimentally shown that the addition of GAGs to collagen scaffolds can preserve them from collagenase digestion, improving the stability of the biomaterial. *C. reniformis*-derived 2D-CMs are naturally GAGs decorated, and this explains their peculiar stability to collagenase treatment. As stated in the Results Section, a certain percentage of components in the 2D-CMs related to their weight loss during the stability tests was not characterized, but we suggest that, reasonably, it could be constituted by polysaccharidic elements; however, until a complete molecular characterization of the *C. reniformis*-derived ECM extract is reached, it will not be possible to perform specific tests to accurately clarify the nature of the above-mentioned elements. The biocompatibility of *C. reniformis*-derived 2D-CMs has been previously attested *in vitro* on cell lines. Additionally, in this study, an expression profile of some key genes of the wound healing process was created. During skin wound healing various matrix metallopeptidase proteins (MMPs) are involved in ECM degradation. Matrix metallopeptidase 3 (MMP-3) is considered essential for the initial phase of tissue repair. Here, in the L929 fibroblasts cell line, the MMP-3 gene expression was not significantly affected, indicating that the sponge-derived biomaterial does not interfere with the regular process of skin regeneration. Typically, during the recovery processes of skin injuries, ECM components such as collagens and fibronectin are released. An imbalance in collagen deposition during wound healing may induce scar tissue formation over time. When fibroblast cell lines were left to grow on *C. reniformis*-derived 2D-membranes, no significant difference in collagen (COL1A1) gene expression was observed for both the incubation times of 24 and 120 h, confirming how the contact of mouse fibroblasts with sponge collagen extract does not affect the fibrogenesis process preventing the risk of scaring process. Comparing these results with gene expression profiles of L929 cell lines growing on other sponge-derived 2D-CMs, the same response was observed when the mouse fibroblasts were let grow on *S. foetidus*-derived 2D-CMs, while a different response was detected in fibroblast grown on *I. oros* derived-2D-CMs, as here the contact with the biomaterial induced after 24 h a significant collagen up-regulation. These data suggest that fibrogenesis mediated by contact with these types of materials is extremely heterogeneous and linked to the species-specific molecular composition and structure of the tested biomaterial. Conversely, after 120 h, L929 mouse fibroblasts growing on sponge-derived 2D-CMs showed a slight but significant up-regulation of the fibronectin gene. One of the main factors able to induce fibronectin up-regulation is tumor growth factor (TGF-β); however, this seems not to be the case, as TGF-β expression was low during the first stage of the experiments. Instead, fibronectin up-regulation could be here related to the interaction of cells with marine collagen. Collagens, including sponges' collagens, are known for containing RGD fibronectin-binding sites and the presence of this motif is confirmed in the fibrillar collagen of *Chondrosia reniformis* in which the aminoacidic sequence has already been partially described. The RGD motif is an important recognition site found in several fibrillar collagens, including collagen types I, II, and III. This motif is recognized by integrin receptors, particularly integrin α5β1. The binding of integrin α5β1 to the RGD motif can activate intracellular signaling pathways to induce epithelial-mesenchymal transition phenomena typical of the wound healing process. As collagen fibronectin is a key ECM protein released during the early phases of skin repair, the interaction with 2D-CMs could be considered able to stimulate the wound-healing process in fibroblast cell lines. Finally, 2D-CMs seem not to affect the expression of neither FGF nor cytokines genes such as IL-1β that are involved in skin wound healing response. These results indicate that, although *C. reniformis* derived 2D-CMs when compared to commercial collagen membranes were able to

improve L929 cells proliferation within the first 3 days of incubation and fibronectin gene expression after 120 h, these responses are not under the control of the above-mentioned growth factors and cytokines.

Once confirmed that *C. reniformis* derived 2D-CMs are potentially suitable as skin-like substitute devices, being biocompatible, non-scarring-inducing, and biodegradable on the long term, the next step was that to create a functionally active membrane with the same biomaterial, with the aim of developing an innovative device employable in the treatment of wounds. The main advantage in choosing *C. reniformis*' collagen stands in the fact that it is resistant to the attack of collagenases, but not completely. This consent on one hand to create functionalized membranes loaded with enzymes useful in the treatment of skin wounds, such as collagenases (as they remove necrotic tissue), obtaining a composite material that can be applied on wounds without being rapidly degraded by the enzyme itself, but that can, the other hand, be resorbed on the long term. During this part of the research activity, a new composite biomaterial made of *C. reniformis'* collagen intimately mixed with biosilica, which immobilizes and stabilizes the *Clostridium histolyticum* collagenase enzyme, was created. The composition of the device results in an optimal combination of marine collagen, associated with biosilica as a potential inducer of the biosynthesis of new extracellular matrix and complemented by the presence of collagenase immobilized on its surface the device through encapsulation in enzymatically produced silica nanoparticles. The silica nanoparticles effectively stabilize the collagenase, as the activity of the enzyme has been confirmed and detected on the membrane up to five day later its creation; this suggests the possibility to functionalize the membrane with also other enzymes, that could work in combination with collagenases during tissue repair processes (e.g., catalases or superoxide dismutase). The presence of GAGs and other polysaccharides, revealed by the FT-IR analysis, could furtherly contribute not only to the resistance of the 2D film to degradation by the collagenases, but also to the stability of the enzyme itself, while simultaneously acting as important contributors to maintain the correct hydration level of the interface microenvironment with the exposed tissues. Finally, giving the good rate of absorption observed for the non-steroidal anti-inflammatory drug Ketorolac tromethamine (see 5.4.5), it can be stated that a biomedical device created as described in the Methods' section could also be used in combination with other drugs and/or bioactive compounds such as antibiotics, antioxidants, growth factors (e.g., TGF) and/or anti-inflammatory molecules.

The data obtained so far allowed the creation of a patent (see the first paragraph of this section) and paved the way for further applications of this collagen following the same approach; however, this part of research is still ongoing. For complete references, see the original works "**2D Collagen Membranes from Marine Demosponge *Chondrosia reniformis* (Nardo, 1847) for Skin-Regenerative Medicine Applications: An *In Vitro* Evaluation**" – Tassara et al., 2023; https://doi.org/10.3390/md21080428) and "**Composition based on collagen of marine origin**" – Pozzolini, Tassara, Giovine, and Lova, 2023; Italian patent N. 102023000014598.

Conclusions

The demosponge *Chondrosia reniformis* is considered a promising source of marine collagen for biomedical applications. During this doctoral research activity, the molecular features of the fibrillar collagen genes of this species have been studied, as well as factors that could enhance collagen biosynthesis within this species, and novel biotechnological applications for this biomaterial have been explored.

Five genes encoding for fibrillar collagen have been individuated in the genome of *C. reniformis* and their molecular features have been described, together with their pattern of expression in physiological conditions. However, the 5'-UTR of these genes is still missing by the time of the writing of this thesis, as the transcriptomic analysis approach couldn't effectively provide a complete sequence of those regions (see Chapter III). For this reason, the next step will foresee the reconstruction of the 5'-UTRs via a 5'-RACE approach, in order to obtain additional insights into the regulation mechanisms of the biosynthesis of *C. reniformis*' fibrillar collagens. One gene, in particular, differs from the others for the structure of its translation product and for the specificity of its expression location in the cortical zone of the sponge: further studies could better investigate the role of this molecule in physiological processes typical of *C. reniformis*, such as the creeping phenomenon. Moreover, a proteomic analysis on the collagenous extract could furtherly improve the understanding how the different gene products affect the biopolymer obtainable from this species.

It was observed that photobiomodulation (PBM) administered at 810 nm, 1 W, and 60 J and applied for 60 s with an FT-HP on a circular area of 1 cm^2 affects *C. reniformis* regeneration, which appears to occur faster than in the control samples. The effect is consistent with our previous data on a wide range of animals and humans, showing cell pathways conserved through the evolution of life forms, which appear to be influenced by NIR light. However, these finding could be improved by a more in-depth analysis on how PBM affects the expression of the different fibrillar collagen genes in this species. Concerning future perspectives, the increment of collagen deposition induced by PBM treatment observed in our study could open its usage to sponge mariculture techniques for biotechnological purposes. Although the 810 nm wavelength of the sun does not usually reach *C. reniformis* in nature, the evidence of interactions between light and sponge tissue suggests a deeper investigation of the influence of sunlight wavelengths on sponge metabolism, growth, and recovery.

Finally, it was introduced for the first time the novel use of a marine sponge-derived collagen extract from *C. reniformis* as a viable biomaterial for producing 2D membranes applicable in skin regenerative medicine. Despite inherent limitations such as slight variability and fixed fiber dimensions, these membranes show promising potential as skin-like substitutes (SLSs): in fact, it was demonstrated that these devices are biocompatible, as no induction of scar tissue or inflammatory responses over time were detected; moreover, they provide a great barrier against water and protein loss and bacterial infiltration, and have a slow degradation rate, ensuring long-term resilience. These findings brought to the production of functionalized 2D membranes, where collagenase molecules were immobilized *via* biosilica precipitation, which can potentially be employed in the treatment of wounds. Through these investigations, the aim is to advance understanding and development of marine collagen films for burn and chronic wound treatment, potentially revolutionizing wound care practices. Further analyses, particularly *in vivo* or on regenerated tissues, are necessary to optimize their functionality and suitability for practical use of these devices.

An important limitation in the employment of *Chondrosia reniformis* collagen is still the lack of sustainability of this source: however, together with further attempts in optimizing aquaculture

techniques effective for a consistent production of this biomaterial, the in-depth molecular description can serve as a starting point for biomimetic applications that could address the sustainability issue associated with the sourcing of this marine collagen.

References

Aarskog, N.; Vedeler, C. Real-Time Quantitative Polymerase Chain Reaction. *Hum Genet* **2000**, *107*, 494–498, doi:10.1007/s004390000399.

Abraham, L.C.; Zuena, E.; Perez-Ramirez, B.; Kaplan, D.L. Guide to Collagen Characterization for Biomaterial Studies. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **2008**, *87B*, 264–285, doi:10.1002/jbm.b.31078.

Addad, S.; Exposito, J.-Y.; Faye, C.; Ricard-Blum, S.; Lethias, C. Isolation, Characterization and Biological Evaluation of Jellyfish Collagen for Use in Biomedical Applications. *Marine Drugs* **2011**, *9*, 967–983, doi:10.3390/md9060967.

Aluko, R.E. Antihypertensive Peptides from Food Proteins. *Annual Review of Food Science and Technology* **2015**, *6*, 235–262, doi:10.1146/annurev-food-022814-015520.

Amaroli, A.; Tassara, E.; Ferrando, S.; Aicardi, S.; Pasquale, C.; Giovine, M.; Bertolino, M.; Zekiy, A.; Pozzolini, M. Near-Infrared 810 Nm Light Affects Porifera Chondrosia Reniformis (Nardo, 1847) Regeneration: Molecular Implications and Evolutionary Considerations of Photobiomodulation–Animal Cell Interaction. *International Journal of Molecular Sciences* **2023**, *24*, 226, doi:10.3390/ijms24010226.

Aouacheria, A.; Geourjon, C.; Aghajari, N.; Navratil, V.; Deléage, G.; Lethias, C.; Exposito, J.-Y. Insights into Early Extracellular Matrix Evolution: Spongin Short Chain Collagen-Related Proteins Are Homologous to Basement Membrane Type IV Collagens and Form a Novel Family Widely Distributed in Invertebrates. *Molecular Biology and Evolution* **2006**, *23*, 2288–2302, doi:10.1093/molbev/msl100.

Arnone, M.I.; Byrne, M.; Martinez, P. Echinodermata. In *Evolutionary Developmental Biology of Invertebrates 6: Deuterostomia*; Wanninger, A., Ed.; Springer: Vienna, 2015; pp. 1–58 ISBN 978-3-7091-1856-6.

Avila Rodríguez, M.I.; Rodríguez Barroso, L.G.; Sánchez, M.L. Collagen: A Review on Its Sources and Potential Cosmetic Applications. *Journal of Cosmetic Dermatology* **2018**, *17*, 20–26, doi:10.1111/jocd.12450.

Awang, N.A.; Amid, A.; Arshad, Z.I. Method for Purification of Collagen: A Systematic Review. *APJMBB* **2020**, 99–112, doi:10.35118/apjmbb.2020.028.3.10.

Bairati, A. *Biology of Invertebrate and Lower Vertebrate Collagens*; Springer Science & Business Media, 2013; ISBN 978-1-4684-7636-1.

Banerjee, P.; Shanthi, C. Cryptic Peptides from Collagen: A Critical Review. *Protein and Peptide Letters* **2016**, *23*, 664–672.

Bavestrello, G.; Arillo, A.; Benatti, U.; Cerrano, C.; Cattaneo-Vietti, R.; Cortesognoi, L.; Gaggero, L.; Giovine, M.; Tonetti, M.; Sarà, M. Quartz Dissolution by the Sponge Chondrosia Reniformis (Porifera, Demospongiae). *Nature* **1995**, *378*, 374–376, doi:10.1038/378374a0.

Bavestrello, G.; Arillo, A.; Calcinai, B.; Cerrano, C.; Lanza, S.; Sarà, M.; Cattaneo-Vietti, R.; Gaino, E. Siliceous Particles Incorporation in Chondrosia Reniformis (Porifera, Demospongiae). *Italian Journal of Zoology* **1998b**, *65*, 343–348, doi:10.1080/11250009809386771.

Bavestrello, G.; Benatti, U.; Calcinai, B.; Cattaneo-Vietti, R.; Cerrano, C.; Favre, A.; Giovine, M.; Lanza, S.; Pronzato, R.; Sara, M. Body Polarity and Mineral Selectivity in the Demosponge Chondrosia Reniformis. *The Biological Bulletin* **1998**, *195*, 120–125, doi:10.2307/1542819.

Bavestrello, G.; Benatti, U.; Calcinai, B.; Cattaneo-Vietti, R.; Cerrano, C.; Favre, A.; Giovine, M.; Lanza, S.; Pronzato, R.; Sara, M. Body Polarity and Mineral Selectivity in the Demosponge Chondrosia Reniformis. *The Biological Bulletin* **1998c**, *195*, 120–125, doi:10.2307/1542819.

Bavestrello, G.; Burlando, B.; Sara', M. The Architecture of the Canal Systems of Petrosia Ficiformis and Chondrosia Reniformis Studied by Corrosion Casts (Porifera, Demospongiae). *Zoomorphology* **1988a**, *108*, 161–166, doi:10.1007/BF00363932.

Belbachir, K.; Noreen, R.; Gouspillou, G.; Petibois, C. Collagen Types Analysis and Differentiation by FTIR Spectroscopy. *Anal Bioanal Chem* **2009**, *395*, 829–837, doi:10.1007/s00216-009-3019-y.

Bermueller, C.; Schwarz, S.; Elsaesser, A.F.; Sewing, J.; Baur, N.; von Bomhard, A.; Scheithauer, M.; Notbohm, H.; Rotter, N. Marine Collagen Scaffolds for Nasal Cartilage Repair: Prevention of Nasal Septal Perforations in a New Orthotopic Rat Model Using Tissue Engineering Techniques. *Tissue Engineering Part A* **2013**, *19*, 2201–2214, doi:10.1089/ten.tea.2012.0650.

Bhatnagar, R.S.; Gough, C.A. Circular Dichroism of Collagen and Related Polypeptides. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G.D., Ed.; Springer US: Boston, MA, 1996; pp. 183–199 ISBN 978-1-4757-2508-7.

Bielajew, B.J.; Hu, J.C.; Athanasiou, K.A. Collagen: Quantification, Biomechanics and Role of Minor Subtypes in Cartilage. *Nat Rev Mater* **2020**, *5*, 730–747, doi:10.1038/s41578-020-0213-1.

Birk, D.E.; Brückner, P. Collagens, Suprastructures, and Collagen Fibril Assembly. In *The Extracellular Matrix: an Overview*; Mecham, R.P., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp. 77–115 ISBN 978-3-642-16554-2.

Bonasoro, F.; Wilkie, I.C.; Bavestrello, G.; Cerrano, C.; Carnevali, M.D.C. Dynamic Structure of the Mesohyl in the Sponge Chondrosia Reniformis (Porifera, Demospongiae). *Zoomorphology* **2001**, *121*, 109–121, doi:10.1007/PL00008497.

Boo, Y.C. Ascorbic Acid (Vitamin C) as a Cosmeceutical to Increase Dermal Collagen for Skin Antiaging Purposes: Emerging Combination Therapies. *Antioxidants* **2022**, *11*, 1663, doi:10.3390/antiox11091663.

Boot-Handford, R.P.; Tuckwell, D.S. Fibrillar Collagen: The Key to Vertebrate Evolution? A Tale of Molecular Incest. *BioEssays* **2003**, *25*, 142–151, doi:10.1002/bies.10230.

Boyd, C. *Extracellular Matrix Genes*; Elsevier, 2012; ISBN 978-0-323-15766-7.

Bradford, M.M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry* **1976**, *72*, 248–254, doi:10.1016/0003-2697(76)90527-3.

Burgeson, R.E. Type VII Collagen, Anchoring Fibrils, and Epidermolysis Bullosa. *Journal of Investigative Dermatology* **1993**, *101*, 252–255, doi:10.1111/1523-1747.ep12365129.

Bustin, S.A. Absolute Quantification of mRNA Using Real-Time Reverse Transcription Polymerase Chain Reaction Assays. *Journal of Molecular Endocrinology* **2000**, *25*, 169–193, doi:10.1677/jme.0.0250169.

Cao, H.; Chen, M.-M.; Liu, Y.; Liu, Y.-Y.; Huang, Y.-Q.; Wang, J.-H.; Chen, J.-D.; Zhang, Q.-Q. Fish Collagen-Based Scaffold Containing PLGA Microspheres for Controlled Growth Factor Delivery in Skin Tissue Engineering. *Colloids and Surfaces B: Biointerfaces* **2015**, *136*, 1098–1106, doi:10.1016/j.colsurfb.2015.10.022.

Capati, M.L.F.; Nakazono, A.; Yamamoto, K.; Sugimoto, K.; Yanagiguchi, K.; Yamada, S.; Hayashi, Y. Fish Collagen Promotes the Expression of Genes Related to Osteoblastic Activity. *International Journal of Polymer Science* **2016**, *2016*, e5785819, doi:10.1155/2016/5785819.

Carolo, A.; Melotti, L.; Zivelonghi, G.; Sacchetto, R.; Akyürek, E.E.; Martinello, T.; Venerando, A.; Iacopetti, I.; Sugni, M.; Martinelli, G.; et al. Mutable Collagenous Tissue Isolated from Echinoderms Leads to the Production of a Dermal Template That Is Biocompatible and Effective for Wound Healing in Rats. *Marine Drugs* **2023**, *21*, 506, doi:10.3390/md21100506.

Cary, S.C.; Shank, T.; Stein, J. Worms Bask in Extreme Temperatures. *Nature* **1998**, *391*, 545–546, doi:10.1038/35286.

Chai, H.-J.; Li, J.-H.; Huang, H.-N.; Li, T.-L.; Chan, Y.-L.; Shiau, C.-Y.; Wu, C.-J. Effects of Sizes and Conformations of Fish-Scale Collagen Peptides on Facial Skin Qualities and Transdermal Penetration Efficiency. *BioMed Research International* **2010**, *2010*, e757301, doi:10.1155/2010/757301.

Chak, V.; Kumar, D.; Visht, S. A Review on Collagen Based Drug Delivery Systems. *International Journal of Pharmacy Teaching & Practices* **2013**, *4*, 811–820.

Chan, A.; Ma, S.; Pearson, B.J.; Chan, D. Collagen IV Differentially Regulates Planarian Stem Cell Potency and Lineage Progression. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2021251118, doi:10.1073/pnas.2021251118.

Chandika, P.; Ko, S.-C.; Jung, W.-K. Marine-Derived Biological Macromolecule-Based Biomaterials for Wound Healing and Skin Tissue Regeneration. *International Journal of Biological Macromolecules* **2015**, *77*, 24–35, doi:10.1016/j.ijbiomac.2015.02.050.

Chattopadhyay, S.; Raines, R.T. Collagen-Based Biomaterials for Wound Healing. *Biopolymers* **2014**, *101*, 821–833, doi:10.1002/bip.22486.

Chen, Q.; Chen, O.; M. Martins, I.; Hou, H.; Zhao, X.; B. Blumberg, J.; Li, B. Collagen Peptides Ameliorate Intestinal Epithelial Barrier Dysfunction in Immunostimulatory Caco-2 Cell Monolayers via Enhancing Tight Junctions. *Food & Function* **2017**, *8*, 1144–1151, doi:10.1039/C6FO01347C.

Chen, Y.-P.; Liang, C.-H.; Wu, H.-T.; Pang, H.-Y.; Chen, C.; Wang, G.-H.; Chan, L.-P. Antioxidant and Anti-Inflammatory Capacities of Collagen Peptides from Milkfish (*Chanos chanos*) Scales. *J Food Sci Technol* **2018**, *55*, 2310–2317, doi:10.1007/s13197-018-3148-4.

Choi, S.Y.; Ko, E.J.; Lee, Y.H.; Kim, B.G.; Shin, H.J.; Seo, D.B.; Lee, S.J.; Kim, B.J.; Kim, M.N. Effects of Collagen Tripeptide Supplement on Skin Properties: A Prospective, Randomized, Controlled Study. *Journal of Cosmetic and Laser Therapy* **2014**, *16*, 132–137, doi:10.3109/14764172.2013.854119.

Christiano, A.M.; Hoffman, G.G.; Chung-Honet, L.C.; Lee, S.; Cheng, W.; Uitto, J.; Greenspan, D.S. Structural Organization of the Human Type VII Collagen Gene (COL7A1), Composed of More Exons Than Any Previously Characterized Gene. *Genomics* **1994**, *21*, 169–179, doi:10.1006/geno.1994.1239.

Chu, M.-L.; de Wet, W.; Bernard, M.; Ding, J.-F.; Morabito, M.; Myers, J.; Williams, C.; Ramirez, F. Human Proα1(I) Collagen Gene Structure Reveals Evolutionary Conservation of a Pattern of Introns and Exons. *Nature* **1984**, *310*, 337–340, doi:10.1038/310337a0.

Churches, N.; Swalla, B.J.; Kohn, A.; Moroz, L. The Collagens of the Ctenophore Pleurobrachia bachei. **2012**.

Cohen-Solal, L.; Castanet, J.; Meunier, F.J.; Glimcher, M.J. Proline Hydroxylation of Collagens Synthesized at Different Temperatures in Vivo by Two Poikilothermic Species. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **1986**, *83*, 483–486, doi:10.1016/0305-0491(86)90399-8.

Constantine, V.S.; Mowry, R.W. Selective Staining of Human Dermal Collagen: II. The Use of Picrosirius Red F3BA with Polarization Microscopy*. *Journal of Investigative Dermatology* **1968**, *50*, 419–423, doi:10.1038/jid.1968.68.

Coppola, D.; Oliviero, M.; Vitale, G.A.; Lauritano, C.; D'Ambra, I.; Iannace, S.; de Pascale, D. Marine Collagen from Alternative and Sustainable Sources: Extraction, Processing and Applications. *Marine Drugs* **2020**, *18*, 214, doi:10.3390/md18040214.

Cox, G.N. Molecular and Biochemical Aspects of Nematode Collagens. *The Journal of Parasitology* **1992**, *78*, 1–15, doi:10.2307/3283678.

Coyne, K.J.; Qin, X.-X.; Waite, J.H. Extensible Collagen in Mussel Byssus: A Natural Block Copolymer. *Science* **1997**, *277*, 1830–1832, doi:10.1126/science.277.5333.1830.

Dai, M.; Liu, X.; Wang, N.; Sun, J. Squid Type II Collagen as a Novel Biomaterial: Isolation, Characterization, Immunogenicity and Relieving Effect on Degenerative Osteoarthritis *via* Inhibiting STAT1 Signaling in pro-Inflammatory Macrophages. *Materials Science and Engineering: C* **2018**, *89*, 283–294, doi:10.1016/j.msec.2018.04.021.

David, C.N.; Özbek, S.; Adamczyk, P.; Meier, S.; Pauly, B.; Chapman, J.; Hwang, J.S.; Gojobori, T.; Holstein, T.W. Evolution of Complex Structures: Minicollagens Shape the Cnidarian Nematocyst. *Trends in Genetics* **2008**, *24*, 431–438, doi:10.1016/j.tig.2008.07.001.

Davison-Kotler, E.; Marshall, W.S.; García-Gareta, E. Sources of Collagen for Biomaterials in Skin Wound Healing. *Bioengineering* **2019**, *6*, 56, doi:10.3390/bioengineering6030056.

De Domenico, S.; De Rinaldis, G.; Paulmery, M.; Piraino, S.; Leone, A. Barrel Jellyfish (*Rhyzostoma pulmo*) as Source of Antioxidant Peptides. *Marine Drugs* **2019**, *17*, 134, doi:10.3390/md17020134.

de Voogd, N.J.; Alvarez, B.; Boury-Esnault, N.; Carballo, J.L.; Cárdenas, P.; Díaz, M.-C.; Dohrmann, M.; Downey, R.; Goodwin, C.; Hajdu, E.; Hooper, J.N.A.; Kelly, M.; Klautau, M.; Lim, S.C.; Manconi, R.; Morrow, C.; Pinheiro, U.; Pisera, A.B.; Ríos, P.; Rützler, K.; Schönberg, C.; Vacelet, J.; van Soest, R.W.M.; Xavier, J. (2023). World Porifera Database. *Chondrosia reniformis* Nardo, 1847. Accessed at: https://www.marinespecies.org/porifera/porifera.php?p=taxdetails&id=134112 on 2023-09-18.

Deming, T.J. Mussel Byssus and Biomolecular Materials. *Current Opinion in Chemical Biology* **1999**, *3*, 100–105, doi:10.1016/S1367-5931(99)80018-0.

Derkus, B.; Arslan, Y.E.; Bayrac, A.T.; Kantarcioglu, I.; Emregul, K.C.; Emregul, E. Development of a Novel Aptasensor Using Jellyfish Collagen as Matrix and Thrombin Detection in Blood Samples Obtained from Patients with Various Neurodisease. *Sensors and Actuators B: Chemical* **2016**, *228*, 725–736, doi:10.1016/j.snb.2016.01.095.

Di Camillo, C.G.; Coppari, M.; Bartolucci, I.; Bo, M.; Betti, F.; Bertolino, M.; Calcinai, B.; Cerrano, C.; De Grandis, G.; Bavestrello, G. Temporal Variations in Growth and Reproduction of Tedania Anhelans and Chondrosia Reniformis in the North Adriatic Sea. In Ancient Animals, New Challenges: Developments in Sponge Research; Maldonado, M., Turon, X., Becerro, M., Jesús Uriz, M., Eds.; Developments in Hydrobiology; Springer Netherlands: Dordrecht, 2012; pp. 299–313 ISBN 978-94-007-4688-6.

Dieudonne, J.; Carroll, J.M. The Impacts of Boring Sponges on Oyster Health across Multiple Sites and Tidal Heights. *Estuaries and Coasts* **2022**, *45*, 213–224, doi:10.1007/s12237-021-00942-1.

Dinescu, S.; Kaya, M.; Chitoiu, L.; Nazarie, S.; Kaya, D.; Costache, M. Collagen-Based Hydrogels and Their Applications for Tissue Engineering and Regenerative Medicine. In; 2019; pp. 1643–1664 ISBN 978-3-319-77829-7.

Dolmatov, I.Y.; Nizhnichenko, V.A. Extracellular Matrix of Echinoderms. *Marine Drugs* **2023**, *21*, 417, doi:10.3390/md21070417.

Dowd, W.W. Challenges for Biological Interpretation of Environmental Proteomics Data in Non-Model Organisms. *Integrative and Comparative Biology* **2012**, *52*, 705–720, doi:10.1093/icb/ics093.

Draper, G.W.; Shoemark, D.K.; Adams, J.C. Modelling the Early Evolution of Extracellular Matrix from Modern Ctenophores and Sponges. *Essays in Biochemistry* **2019**, *63*, 389–405, doi:10.1042/EBC20180048.

Ehrlich, H. Chitin and Collagen as Universal and Alternative Templates in Biomineralization. *International Geology Review* **2010**, *52*, 661–699, doi:10.1080/00206811003679521.

Ehrlich, H.; Wysokowski, M.; Żółtowska-Aksamitowska, S.; Petrenko, I.; Jesionowski, T. Collagens of Poriferan Origin. *Marine Drugs* **2018**, *16*, 79, doi:10.3390/md16030079.

Engel, J. Versatile Collagens in Invertebrates. *Science* **1997**, *277*, 1785–1786, doi:10.1126/science.277.5333.1785.

Erwin, P.M.; Coma, R.; López-Sendino, P.; Serrano, E.; Ribes, M. Stable symbionts across the HMA-LMA dichotomy: Low seasonal and interannual variation in sponge-associated bacteria from taxonomically diverse hosts. *FEMS Microbiol. Ecol.* **2015**, *91*, 10.

Exposito, J.-Y.; Cluzel, C.; Garrone, R.; Lethias, C. Evolution of Collagens. *The Anatomical Record* **2002**, *268*, 302–316, doi:10.1002/ar.10162.

Exposito, J.-Y.; Larroux, C.; Cluzel, C.; Valcourt, U.; Lethias, C.; Degnan, B.M. Demosponge and Sea Anemone Fibrillar Collagen Diversity Reveals the Early Emergence of A/C Clades and the Maintenance of the Modular Structure of Type V/XI Collagens from Sponge to Human *. *Journal of Biological Chemistry* **2008**, *283*, 28226–28235, doi:10.1074/jbc.M804573200.

Exposito, J.-Y.; Lethias, C. Invertebrate and Vertebrate Collagens. In *Evolution of Extracellular Matrix*; Keeley, F.W., Mecham, R.P., Eds.; Biology of Extracellular Matrix; Springer: Berlin, Heidelberg, 2013; pp. 39–72 ISBN 978-3-642-36002-2.

Exposito, J.-Y.; van der Rest, M.; Garrone, R. The Complete Intron/Exon Structure of Ephydatia mülleri Fibrillar Collagen Gene Suggests a Mechanism for the Evolution of an Ancestral Gene Module. *J Mol Evol* **1993**, *37*, 254–259, doi:10.1007/BF00175502.

Fassini, D.; Duarte, A.R.C.; Reis, R.L.; Silva, T.H. Bioinspiring Chondrosia Reniformis (Nardo, 1847) Collagen-Based Hydrogel: A New Extraction Method to Obtain a Sticky and Self-Healing Collagenous Material. *Marine Drugs* **2017**, *15*, 380, doi:10.3390/md15120380.

Fassini, D.; Parma, L.; Lembo, F.; Candia Carnevali, M.D.; Wilkie, I.C.; Bonasoro, F. The Reaction of the Sponge *Chondrosia reniformis* to Mechanical Stimulation Is Mediated by the Outer Epithelium and the Release of Stiffening Factor(s). *Zoology* **2014**, *117*, 282–291, doi:10.1016/j.zool.2014.03.003.

Fassini, D.; Parma, L.; Wilkie, I.C.; Bavestrello, G.; Bonasoro, F.; Candia Carnevali, M.D. Ecophysiology of Mesohyl Creep in the Demosponge *Chondrosia reniformis* (Porifera: Chondrosida). *J. Exp. Mar. Biol. Ecol.* **2012**, *428*, 24–31.

Fassini, D.; Wilkie, I.C.; Pozzolini, M.; Ferrario, C.; Sugni, M.; Rocha, M.S.; Giovine, M.; Bonasoro, F.; Silva, T.H.; Reis, R.L. Diverse and Productive Source of Biopolymer Inspiration: Marine Collagens. *Biomacromolecules* **2021**, *22*, 1815–1834, doi:10.1021/acs.biomac.1c00013.

Felician, F.F.; Xia, C.; Qi, W.; Xu, H. Collagen from Marine Biological Sources and Medical Applications. *Chemistry & Biodiversity* **2018**, *15*, e1700557, doi:10.1002/cbdv.201700557.

Fenoglio, I.; Martra, G.; Coluccia, S.; Fubini, B. Possible Role of Ascorbic Acid in the Oxidative Damage Induced by Inhaled Crystalline Silica Particles. *Chem. Res. Toxicol.* **2000**, *13*, 971–975, doi:10.1021/tx000125h.

Ferrario, C.; Leggio, L.; Leone, R.; Di Benedetto, C.; Guidetti, L.; Coccè, V.; Ascagni, M.; Bonasoro, F.; La Porta, C.A.M.; Candia Carnevali, M.D.; et al. Marine-Derived Collagen Biomaterials from Echinoderm Connective Tissues. *Marine Environmental Research* **2017**, *128*, 46–57, doi:10.1016/j.marenvres.2016.03.007.

Fessas, D.; Signorelli, M.; Schiraldi, A.; Kennedy, C.J.; Wess, T.J.; Hassel, B.; Nielsen, K. Thermal Analysis on Parchments I: DSC and TGA Combined Approach for Heat Damage Assessment. *Thermochimica Acta* **2006**, *447*, 30–35, doi:10.1016/j.tca.2006.04.007.

Fibrous Proteins: Structures and Mechanisms; Parry, D.A.D., Squire, J.M., Eds.; Subcellular Biochemistry; Springer International Publishing: Cham, 2017; Vol. 82; ISBN 978-3-319-49672-6.

Fidler, A.L.; Darris, C.E.; Chetyrkin, S.V.; Pedchenko, V.K.; Boudko, S.P.; Brown, K.L.; Gray Jerome, W.; Hudson, J.K.; Rokas, A.; Hudson, B.G. Collagen IV and Basement Membrane at the Evolutionary Dawn of Metazoan Tissues. *eLife* **2017**, *6*, e24176, doi:10.7554/eLife.24176.

Francois, J.; Herbage, D.; Junqua, S. Cockroach Collagen: Isolation, Biochemical and Biophysical Characterization. *European Journal of Biochemistry* **1980**, *112*, 389–396, doi:10.1111/j.1432-1033.1980.tb07217.x.

Friess, W. Collagen – Biomaterial for Drug Delivery1. *European Journal of Pharmaceutics and Biopharmaceutics* **1998**, *45*, 113–136, doi:10.1016/S0939-6411(98)00017-4.

Fronchek, E.V.; Grigoryan, A.Y.; Blatun, L.A. Innovative biologically active wound healing and haemostatic agents based on chitosan and collagen: development stages, medical and technical

characteristics. *Wounds and wound infections. The prof. B.M. Kostyuchenok journal* **2020**, *5*, 14–21, doi:10.25199/2408-9613-2018-5-4-14-21.

Fu, Y.; Therkildsen, M.; Aluko, R.E.; Lametsch, R. Exploration of Collagen Recovered from Animal By-Products as a Precursor of Bioactive Peptides: Successes and Challenges. *Critical Reviews in Food Science and Nutrition* **2019**, *59*, 2011–2027, doi:10.1080/10408398.2018.1436038.

Fukai, N.; Apte, S.S.; Olsen, B.R. [1] Nonfibrillar Collagens. In *Methods in Enzymology*; Extracellular Matrix Components; Academic Press, 1994; Vol. 245, pp. 3–28.

Gaill, F.; Bouligand, Y. Supercoil of Collagen Fibrils in the Integument of *Alvinella*, an Abyssal Annelid. *Tissue and Cell* **1987**, *19*, 625–642, doi:10.1016/0040-8166(87)90070-X.

Garrone, R.; Huc, A.; Junqua, S. Fine Structure and Physicochemical Studies on the Collagen of the Marine Sponge *Chondrosia reniformis* Nardo. *Journal of Ultrastructure Research* **1975**, *52*, 261–275, doi:10.1016/S0022-5320(75)80117-1.

Garrone, R.; Vacelet, J.; Ceccatty, M. An Unusual Collagen Structure: The Filament of Horny Sponges Ircinia. Ultrastructural, Physicochemical and Biochemical Study [FORMATION COLLAGENE PARTICULIERE: LES FILAMENTS DES EPONGES CORNEES IRCINIA. ETUDE ULTRASTRUCTURALE, PHYSICO CHIMIQUE ET BIOCHIMIQUE]. *Journal of Microscopy* **1973**, *17*, 241–260.

Geddis, A.E.; Prockop, D.J. Expression of Human COL1A1 Gene in Stably Transfected HT1080 Cells: The Production of a Thermostable Homotrimer of Type I Collagen in a Recombinant System. *Matrix* **1993**, *13*, 399–405, doi:10.1016/S0934-8832(11)80045-4.

Gelse, K.; Pöschl, E.; Aigner, T. Collagens—Structure, Function, and Biosynthesis. *Advanced Drug Delivery Reviews* **2003**, *55*, 1531–1546, doi:10.1016/j.addr.2003.08.002.

Giovine, M.; Scarfì, S.; Pozzolini, M.; Penna, A.; Cerrano, C. Cell Reactivity to Different Silica. In *Biomedical Inorganic Polymers: Bioactivity and Applications of Natural and Synthetic Polymeric Inorganic Molecules*; Müller, W.E.G., Wang, X., Schröder, H.C., Eds.; Progress in Molecular and Subcellular Biology; Springer: Berlin, Heidelberg, 2013; pp. 143–174 ISBN 978-3-642-41004-8.

Gökalp, M.; Wijgerde, T.; Murk, A.; Osinga, R. Design for Large-Scale Maricultures of the Mediterranean Demosponge Chondrosia Reniformis Nardo, 1847 for Collagen Production. *Aquaculture* **2022**, *548*, 737702, doi:10.1016/j.aquaculture.2021.737702.

Gordon, M.K.; Hahn, R.A. Collagens. *Cell Tissue Res* **2010**, *339*, 247–257, doi:10.1007/s00441-009-0844-4.

Göthel H. Guide De La Faune Sous-Marine: La Méditerranée. Invertébrés Marins Et Poisons. (Eygen Ulmer GmbH & Co., **1992**).

Gross, J.; Sokal, Z.; Rougvie, M. Structural and chemical studies on the connective tissue of marine sponges. *J Histochem Cytochem.* **1956**, *4*, 227–246, doi:10.1177/4.3.227.

Haq, F.; Ahmed, N.; Qasim, M. Comparative Genomic Analysis of Collagen Gene Diversity. *3 Biotech* **2019**, *9*, 83, doi:10.1007/s13205-019-1616-9.

Hashim, P.; Mohd Ridzwan, M. S.; Bakar, J.; Mat Hashim; D. Collagen in food and beverage industries. *International Food Research Journal* **2015**, 22, 1-8.

Heinemann, S.; Coradin, T.; F. Desimone, M. Bio-Inspired Silica–Collagen Materials: Applications and Perspectives in the Medical Field. *Biomaterials Science* **2013**, *1*, 688–702, doi:10.1039/C3BM00014A.

Heinemann, S.; Ehrlich, H.; Douglas, T.; Heinemann, C.; Worch, H.; Schatton, W.; Hanke, T. Ultrastructural Studies on the Collagen of the Marine Sponge Chondrosia Reniformis Nardo. *Biomacromolecules* **2007**, *8*, 3452–3457, doi:10.1021/bm700574y.

Heino, J.; Kapyla, J. Cellular Receptors of Extracellular Matrix Molecules. *Current Pharmaceutical Design* **2009**, *15*, 1309–1317, doi:10.2174/138161209787846720.

Helling, A.L.; Tsekoura, E.K.; Biggs, M.; Bayon, Y.; Pandit, A.; Zeugolis, D.I. In Vitro Enzymatic Degradation of Tissue Grafts and Collagen Biomaterials by Matrix Metalloproteinases: Improving the Collagenase Assay. *ACS Biomater. Sci. Eng.* **2017**, *3*, 1922–1932, doi:10.1021/acsbiomaterials.5b00563.

Hema Thanka Christlet, T.; Veluraja, K. Database Analysis of O-Glycosylation Sites in Proteins. *Biophysical Journal* **2001**, *80*, 952–960, doi:10.1016/S0006-3495(01)76074-2.

Holm Nielsen, S.; Crespo-Bravo, M.; Karsdal, M.A. Chapter 24 - Type XXIV Collagen. In *Biochemistry of Collagens, Laminins and Elastin (Third Edition)*; Karsdal, M.A., Ed.; Academic Press, 2024; pp. 203–205 ISBN 978-0-443-15617-5.

Holmes, D.F.; Graham, H.K.; Trotter, J.A.; Kadler, K.E. STEM/TEM Studies of Collagen Fibril Assembly. *Micron* **2001**, *32*, 273–285, doi:10.1016/S0968-4328(00)00040-8.

Hu, J.; Song, Y.; Zhang, C.; Huang, W.; Chen, A.; He, H.; Zhang, S.; Chen, Y.; Tu, C.; Liu, J.; et al. Highly Aligned Electrospun Collagen/Polycaprolactone Surgical Sutures with Sustained Release of Growth Factors for Wound Regeneration. *ACS Appl. Bio Mater.* **2020**, *3*, 965–976, doi:10.1021/acsabm.9b01000.

Hudecki, A.; Kiryczyński, G.; Łos, M.J. Chapter 7 - Biomaterials, Definition, Overview. In *Stem Cells and Biomaterials for Regenerative Medicine*; Łos, M.J., Hudecki, A., Wiecheć, E., Eds.; Academic Press, 2019; pp. 85–98 ISBN 978-0-12-812258-7.

Hulmes, D.J.S. Building Collagen Molecules, Fibrils, and Suprafibrillar Structures. *Journal of Structural Biology* **2002**, *137*, 2–10, doi:10.1006/jsbi.2002.4450.

Illidge, C.; Kielty, C.; Shuttleworth, A. Type VIII Collagen: Heterotrimeric Chain Association. *The International Journal of Biochemistry & Cell Biology* **2001**, *33*, 521–529, doi:10.1016/S1357-2725(01)00013-9.

Imhoff, J.M.; Garrone, R. Solubilization and Characterization of Chondrosia Reniformis Sponge Collagen. *Connective Tissue Research* **1983**, doi:10.3109/03008208309004855. 1.

In Jeong, S.; Kim, S.Y.; Cho, S.K.; Chong, M.S.; Kim, K.S.; Kim, H.; Lee, S.B.; Lee, Y.M. Tissue-Engineered Vascular Grafts Composed of Marine Collagen and PLGA Fibers Using Pulsatile Perfusion Bioreactors. *Biomaterials* **2007**, *28*, 1115–1122, doi:10.1016/j.biomaterials.2006.10.025.

Ishibashi, Y.; Haraguchi, R.; Aoki, S.; Oishi, Y.; Narita, T. Effect of UV Irradiation of Pre-Gel Solutions on the Formation of Collagen Gel Tubes. *Gels* **2023**, *9*, 458, doi:10.3390/gels9060458.

Iwamoto, Y.; Haraguchi, R.; Nakao, R.; Aoki, S.; Oishi, Y.; Narita, T. One-Pot Preparation of Collagen Tubes Using Diffusing Gelation. *ACS Omega* **2022**, *7*, 22872–22878, doi:10.1021/acsomega.2c02623.

Jenkins, C.L.; Bretscher, L.E.; Guzei, I.A.; Raines, R.T. Effect of 3-Hydroxyproline Residues on Collagen Stability. *J. Am. Chem. Soc.* **2003**, *125*, 6422–6427, doi:10.1021/ja034015j.

K. Das, S.; R. Khan, M.M.; K. Guha, A.; Naskar, N. Bio-Inspired Fabrication of Silver Nanoparticles on Nanostructured Silica: Characterization and Application as a Highly Efficient Hydrogenation Catalyst. *Green Chemistry* **2013**, *15*, 2548–2557, doi:10.1039/C3GC40310F.

Kabiraj, A.; Gupta, J.; Khaitan, T.; Bhattacharya, P.T. Principle and techniques of immunohistochemistry – a review. *International Journal of Biological & Medical Research* **2015**, *6*, 5204–5210.

Kaytes, P.; Wood, L.; Theriault, N.; Kurkinen, M.; Vogeli, G. Head-to-Head Arrangement of Murine Type IV Collagen Genes. *Journal of Biological Chemistry* **1988**, *263*, 19274–19277, doi:10.1016/S0021-9258(19)77629-7.

Keller, L.; Pugliano, M. Combined Jellyfish Collagen Type II, Human Stem Cells and TGFß-B3 as a Therapeutic Implant for Cartilage Repair. *Journal of Stem Cell Research & Therapy* **2017**, *07*, doi:10.4172/2157-7633.1000382.

Kershaw, D.R. Phylum Arthropoda. In *Animal Diversity*; Kershaw, D.R., Ed.; Springer Netherlands: Dordrecht, 1983; pp. 121–168 ISBN 978-94-011-6035-3.

Kesava Reddy, G.; Enwemeka, C.S. A Simplified Method for the Analysis of Hydroxyproline in Biological Tissues. *Clinical Biochemistry* **1996**, *29*, 225–229, doi:10.1016/0009-9120(96)00003-6.

Khong, N.M.H.; Yusoff, F.Md.; Jamilah, B.; Basri, M.; Maznah, I.; Chan, K.W.; Armania, N.; Nishikawa, J. Improved Collagen Extraction from Jellyfish (*Acromitus hardenbergi*) with Increased Physical-Induced Solubilization Processes. *Food Chemistry* **2018**, *251*, 41–50, doi:10.1016/j.foodchem.2017.12.083.

Kimura, S.; Kobayashi, N.; Kubota, M. Studies on Marine Invertebrate Collagens-VI. *Nippon Suisan Gakkaishi* **1972**, *38*, 1163–1169, doi:10.2331/suisan.38.1163.

Kimura, S.; Matsuura, F. The Chain Compositions of Several Invertebrate Collagens. *The Journal of Biochemistry* **1974**, *75*, 1231–1240, doi:10.1093/oxfordjournals.jbchem.a130506.

Kiontke, K.; Fitch, D.H.A. Nematodes. *Current Biology* **2013**, *23*, R862–R864, doi:10.1016/j.cub.2013.08.009.

Kitts, D.D.; Weiler, K. Bioactive Proteins and Peptides from Food Sources. Applications of Bioprocesses Used in Isolation and Recovery. *Current Pharmaceutical Design* **2003**, *9*, 1309–1323, doi:10.2174/1381612033454883.

Lang, K.; Schmid, F.X.; Fischer, G. Catalysis of Protein Folding by Prolyl Isomerase. *Nature* **1987**, *329*, 268–270, doi:10.1038/329268a0.

Lazoski, C.; Solé-Cava, A.; Boury-Esnault, N.; Klautau, M.; Russo, C. Cryptic Speciation in a High Gene Flow Scenario in the Oviparous Marine Sponge Chondrosia Reniformis. *Marine Biology* **2001**, *139*, 421–429, doi:10.1007/s002270100542.

Le Bris, N.; Gaill, F. How Does the Annelid Alvinella Pompejana Deal with an Extreme Hydrothermal Environment? *Rev Environ Sci Biotechnol* **2007**, *6*, 197–221, doi:10.1007/s11157-006-9112-1. 1.

LeBleu, V.S.; MacDonald, B.; Kalluri, R. Structure and Function of Basement Membranes. *Exp Biol Med (Maywood)* **2007**, *232*, 1121–1129, doi:10.3181/0703-MR-72.

Lee, S.-W.; Kim, H.-K.; Naidansuren, P.; Ham, K.A.; Choi, H.S.; Ahn, H.-Y.; Kim, M.; Kang, D.H.; Kang, S.W.; Joe, Y.A. Peroxidasin Is Essential for Endothelial Cell Survival and Growth Signaling by Sulfilimine Crosslink-Dependent Matrix Assembly. *The FASEB Journal* **2020**, *34*, 10228–10241, doi:10.1096/fj.201902899R.

Lévi, C. and Lévi, P. Embryogenèse de Chondrosia reniformis (Nardo), Démosponge ovipare, et transmission des bactéries symbiotiques. Ann. Sc. Nat., Zool. **1976**, *18*, 367-380.

Li, Y.-Y.; Yu, H.; Guo, Z.-M.; Guo, T.-Q.; Tu, K.; Li, Y.-X. Systematic Analysis of Head-to-Head Gene Organization: Evolutionary Conservation and Potential Biological Relevance. *PLOS Computational Biology* **2006**, *2*, e74, doi:10.1371/journal.pcbi.0020074.

Liaci, L. S. and Sciscioli, M. Sexual cycles of some marine Porifera. *Pubbl. Stn. Zool. Napoli* **1975**, *39*, 307–316.

Lin, C.C.; Chiou, T.K.; Sung, W.C. Characteristics of Gelatin from Giant Grouper (*Epinephelus lanceolatus*) Skin. *International Journal of Food Properties* **2015**, *18*, 2339–2348, doi:10.1080/10942912.2014.980947.

Lin, K.; Zhang, D.; Macedo, M.H.; Cui, W.; Sarmento, B.; Shen, G. Advanced Collagen-Based Biomaterials for Regenerative Biomedicine. *Advanced Functional Materials* **2019**, *29*, 1804943, doi:10.1002/adfm.201804943.

Lin, Y.-J.; Le, G.-W.; Wang, J.-Y.; Li, Y.-X.; Shi, Y.-H.; Sun, J. Antioxidative Peptides Derived from Enzyme Hydrolysis of Bone Collagen after Microwave Assisted Acid Pre-Treatment and Nitrogen Protection. *International Journal of Molecular Sciences* **2010**, *11*, 4297–4308, doi:10.3390/ijms11114297.

Liu, X.; Zhang, M.; Shi, Y.; Qiao, R.; Tang, W.; Sun, Z. Production of the Angiotensin I Converting Enzyme Inhibitory Peptides and Isolation of Four Novel Peptides from Jellyfish (*Rhopilema esculentum*) Protein Hydrolysate. *Journal of the Science of Food and Agriculture* **2016**, *96*, 3240–3248, doi:10.1002/jsfa.7507.

Liu, Y.; Griffith, M.; Watsky, M.A.; Forrester, J.V.; Kuffová, L.; Grant, D.; Merrett, K.; Carlsson, D.J. Properties of Porcine and Recombinant Human Collagen Matrices for Optically Clear Tissue Engineering Applications. *Biomacromolecules* **2006**, *7*, 1819–1828, doi:10.1021/bm060160o.

M. Pawelec, K.; M. Best, S.; E. Cameron, R. Collagen: A Network for Regenerative Medicine. *Journal of Materials Chemistry B* **2016**, *4*, 6484–6496, doi:10.1039/C6TB00807K.

Ma, F.; Xiao, Z.; Chen, B.; Hou, X.; Dai, J.; Xu, R. Linear Ordered Collagen Scaffolds Loaded with Collagen-Binding Basic Fibroblast Growth Factor Facilitate Recovery of Sciatic Nerve Injury in Rats. *Tissue Engineering Part A* **2014**, *20*, 1253–1262, doi:10.1089/ten.tea.2013.0158.

Mackie, G.O.; Mills, C.E.; Singla, C.L. Structure and Function of the Prehensile Tentilla of Euplokamis (Ctenophora, Cydippida). *Zoomorphology* **1988**, *107*, 319–337, doi:10.1007/BF00312216.

Malone, J.P.; George, A.; Veis, A. Type I Collagen N-Telopeptides Adopt an Ordered Structure When Docked to Their Helix Receptor during Fibrillogenesis*. *Proteins: Structure, Function, and Bioinformatics* **2004**, *54*, 206–215, doi:10.1002/prot.10526.

Mann, K.; Mechling, D.E.; Bächinger, H.P.; Eckerskorn, C.; Gaill, F.; Timpl, R. Glycosylated Threonine but Not 4-Hydroxyproline Dominates the Triple Helix Stabilizing Positions in the Sequence of a Hydrothermal Vent Worm Cuticle Collagen. *Journal of Molecular Biology* **1996**, *261*, 255–266, doi:10.1006/jmbi.1996.0457.

Marconi, G.D.; Fonticoli, L.; Rajan, T.S.; Pierdomenico, S.D.; Trubiani, O.; Pizzicannella, J.; Diomede, F. Epithelial-Mesenchymal Transition (EMT): The Type-2 EMT in Wound Healing, Tissue Regeneration and Organ Fibrosis. *Cells* **2021**, *10*, 1587, doi:10.3390/cells10071587.

Marin, F. Mollusc Shellomes: Past, Present and Future. *Journal of Structural Biology* **2020**, *212*, 107583, doi:10.1016/j.jsb.2020.107583.

Martin-Malpartida, P.; Batet, M.; Kaczmarska, Z.; Freier, R.; Gomes, T.; Aragón, E.; Zou, Y.; Wang, Q.; Xi, Q.; Ruiz, L.; et al. Structural Basis for Genome Wide Recognition of 5-Bp GC Motifs by SMAD Transcription Factors. *Nat Commun* **2017**, *8*, 2070, doi:10.1038/s41467-017-02054-6.

Matinong, A.M.E.; Chisti, Y.; Pickering, K.L.; Haverkamp, R.G. Collagen Extraction from Animal Skin. *Biology* **2022**, *11*, 905, doi:10.3390/biology11060905.

Matsuda, N.; Koyama, Y.; Hosaka, Y.; Ueda, H.; Watanabe, T.; Araya, T.; Irie, S.; Takehana, K. Effects of Ingestion of Collagen Peptide on Collagen Fibrils and Glycosaminoglycans in the Dermis. *Journal of Nutritional Science and Vitaminology* **2006**, *52*, 211–215, doi:10.3177/jnsv.52.211.

Matsumoto, R.; Uemura, T.; Xu, Z.; Yamaguchi, I.; Ikoma, T.; Tanaka, J. Rapid Oriented Fibril Formation of Fish Scale Collagen Facilitates Early Osteoblastic Differentiation of Human Mesenchymal Stem Cells. *Journal of Biomedical Materials Research Part A* **2015**, *103*, 2531–2539, doi:10.1002/jbm.a.35387.

Mazzola, P.G.; Lopes, A.M.; Hasmann, F.A.; Jozala, A.F.; Penna, T.C.; Magalhaes, P.O.; Rangel-Yagui, C.O.; Pessoa Jr, A. Liquid–Liquid Extraction of Biomolecules: An Overview and Update of the Main Techniques. *Journal of Chemical Technology & Biotechnology* **2008**, *83*, 143–157, doi:10.1002/jctb.1794.

Miyazaki, H.; Hayashi, K. Tensile Tests of Collagen Fibers Obtained from the Rabbit Patellar Tendon. *Biomedical Microdevices* **1999**, *2*, 151–157, doi:10.1023/A:1009953805658.

Mizuta, S.; Isobe, S.; Yoshinaka, R. Existence of Two Molecular Species of Collagen in the Muscle Layer of the Ascidian (*Halocynthia Roretzi*). *Food Chemistry* **2002**, *79*, 9–13, oi:10.1016/S0308-8146(02)00167-X.

Mizuta, S.; Yoshinaka, R.; Sato, M.; Sakaguchi, M. Characterization of Collagen in Muscle of Several Crustacean Species. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **1994**, *107*, 365–370, doi:10.1016/0305-0491(94)90060-4.

Monboisse, J.C.; Oudart, J.B.; Ramont, L.; Brassart-Pasco, S.; Maquart, F.X. Matrikines from Basement Membrane Collagens: A New Anti-Cancer Strategy. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2014**, *1840*, 2589–2598, doi:10.1016/j.bbagen.2013.12.029.

Müller, W.E.G.; Boreiko, A.; Wang, X.; Belikov, S.I.; Wiens, M.; Grebenjuk, V.A.; Schloβmacher, U.; Schröder, H.C. Silicateins, the Major Biosilica Forming Enzymes Present in Demosponges:

Protein Analysis and Phylogenetic Relationship. *Gene* **2007**, *395*, 62–71, doi:10.1016/j.gene.2007.02.014.

Murray, L.W.; Tanzer, M.L.; Cooke, P. *Nereis* Cuticle Collagen: Relationship of Fiber Ultrastructure to Biochemical and Biophysical Properties. *Journal of Ultrastructure Research* **1981**, *76*, 27–45, doi:10.1016/S0022-5320(81)80048-2.

Nardo, G.D. Osservazioni anatomiche sopra l'animale marino detto volgarmente Rognone di mare. *Atti delle Adunanze dell'I. R. Istituto Veneto di Scienze, Lettere ed Arti* **1847** 6: 267-276.

Nickel, M. Evolutionary Emergence of Synaptic Nervous Systems: What Can We Learn from the Non-Synaptic, Nerveless Porifera? *Invertebrate Biology* **2010**, *129*, 1–16, doi:10.1111/j.1744-7410.2010.00193.x.

Nickel, M.; Brümmer, F. In Vitro Sponge Fragment Culture of Chondrosia Reniformis (Nardo, 1847). *Journal of Biotechnology* **2003**, *100*, 147–159, doi:10.1016/S0168-1656(02)00256-0.

Novelli, J.; Page, A.P.; Hodgkin, J. The C Terminus of Collagen SQT-3 Has Complex and Essential Functions in Nematode Collagen Assembly. *Genetics* **2006**, *172*, 2253–2267, doi:10.1534/genetics.105.053637.

O'Leary, L.E.R.; Fallas, J.A.; Bakota, E.L.; Kang, M.K.; Hartgerink, J.D. Multi-Hierarchical Self-Assembly of a Collagen Mimetic Peptide from Triple Helix to Nanofibre and Hydrogel. *Nature Chem* **2011**, *3*, 821–828, doi:10.1038/nchem.1123.

Øiseth, S.K.; Delahunty, C.; Cochet, M.; Lundin, L. Why Is Abalone So Chewy? Structural Characterization and Relationship to Textural Attributes. *shre* **2013**, *32*, 73–79, doi:10.2983/035.032.0113.

Onursal, C.; Dick, E.; Angelidis, I.; Schiller, H.B.; Staab-Weijnitz, C.A. Collagen Biosynthesis, Processing, and Maturation in Lung Ageing. *Frontiers in Medicine* **2021**, *8*.

Orel, B.; Giovine, M.; Ilan, M. On the Path to Thermo-Stable Collagen: Culturing the Versatile Sponge Chondrosia Reniformis. *Marine Drugs* **2021**, *19*, 669, doi:10.3390/md19120669.

Orgel, J.P.R.O.; Sella, I.; Madhurapantula, R.S.; Antipova, O.; Mandelberg, Y.; Kashman, Y.; Benayahu, D.; Benayahu, Y. Molecular and Ultrastructural Studies of a Fibrillar Collagen from Octocoral (Cnidaria). *Journal of Experimental Biology* **2017**, *220*, 3327–3335, doi:10.1242/jeb.163824.

Owczarzy, A.; Kurasiński, R.; Kulig, K.; Rogóż, W.; Szkudlarek, A.; Maciążek-Jurczyk, M. Collagen - Structure, Properties and Application. *Engineering of Biomaterials* **2020**, *Vol. 23*, doi:10.34821/eng.biomat.156.2020.17-23.

Özbek, S. The Cnidarian Nematocyst: A Miniature Extracellular Matrix within a Secretory Vesicle. *Protoplasma* **2011**, *248*, 635–640, doi:10.1007/s00709-010-0219-4.

Pal, P.; Srivas, P.K.; Dadhich, P.; Das, B.; Maity, P.P.; Moulik, D.; Dhara, S. Accelerating Full Thickness Wound Healing Using Collagen Sponge of Mrigal Fish (*Cirrhinus cirrhosus)* Scale Origin. *International Journal of Biological Macromolecules* **2016**, *93*, 1507–1518, doi:10.1016/j.ijbiomac.2016.04.032.

Pallela, R.; Bojja, S.; Janapala, V.R. Biochemical and Biophysical Characterization of Collagens of Marine Sponge, *Ircinia Fusca* (Porifera: Demospongiae: Irciniidae). *International Journal of Biological Macromolecules* **2011**, *49*, 85–92, doi:10.1016/j.ijbiomac.2011.03.019.

Palmer, I.; Nelson, J.; Schatton, W.; Dunne, N.J.; Buchanan, F.J.; Clarke, S.A. Biocompatibility of Calcium Phosphate Bone Cement with Optimized Mechanical Properties. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **2016**, *104*, 308–315, doi:10.1002/jbm.b.33370.

Panduranga Rao, K. Recent Developments of Collagen-Based Materials for Medical Applications and Drug Delivery Systems. *Journal of Biomaterials Science, Polymer Edition* **1996**, *7*, 623–645, doi:10.1163/156856295X00526.

Paschalis, E.P.; Verdelis, K.; Doty, S.B.; Boskey, A.L.; Mendelsohn, R.; Yamauchi, M. Spectroscopic Characterization of Collagen Cross-Links in Bone. *Journal of Bone and Mineral Research* **2001**, *16*, 1821–1828, doi:10.1359/jbmr.2001.16.10.1821.

Patino, M.G.; Neiders, M.E.; Andreana, S.; Noble, B.; Cohen, R.E. Collagen: An Overview. *Implant Dentistry* **2002**, *11*, 280.

Pita, L.; Rix, L.; Slaby, B.M.; Franke, A.; Hentschel, U. The Sponge Holobiont in a Changing Ocean: From Microbes to Ecosystems. *Microbiome* **2018**, *6*, 46, doi:10.1186/s40168-018-0428-1.

Plez, K.A.; Gross, J. The Amino Acid Composition and Morphology of Some Invertebrate and Vertebrate Collagens. *Biochimica et Biophysica Acta* **1959**, *34*, 24–39, doi:10.1016/0006-3002(59)90229-X.

Plumb, D.A.; Ferrara, L.; Torbica, T.; Knowles, L.; Jr, A.M.; Kadler, K.E.; Briggs, M.D.; Boot-Handford, R.P. Collagen XXVII Organises the Pericellular Matrix in the Growth Plate. *PLOS ONE* **2011**, *6*, e29422, doi:10.1371/journal.pone.0029422.

Pokidysheva, E.; Boudko, S.; Vranka, J.; Zientek, K.; Maddox, K.; Moser, M.; Fässler, R.; Ware, J.; Bächinger, H.P. Biological Role of Prolyl 3-Hydroxylation in Type IV Collagen. *Proceedings of the National Academy of Sciences* **2014**, *111*, 161–166, doi:10.1073/pnas.1307597111.

Pozzolini, M.; Bruzzone, F.; Berilli, V.; Mussino, F.; Cerrano, C.; Benatti, U.; Giovine, M. Molecular Characterization of a Nonfibrillar Collagen from the Marine Sponge Chondrosia Reniformis Nardo 1847 and Positive Effects of Soluble Silicates on Its Expression. *Mar Biotechnol* **2012**, *14*, 281–293, doi:10.1007/s10126-011-9415-2.

Pozzolini, M.; Ferrando, S.; Gallus, L.; Gambardella, C.; Ghignone, S.; Giovine, M. Aquaporin in Chondrosia Reniformis Nardo, 1847 and Its Possible Role in the Interaction Between Cells and Engulfed Siliceous Particles. *The Biological Bulletin* **2016**, *230*, 220–232, doi:10.1086/BBLv230n3p220.

Pozzolini, M.; Gallus, L.; Ghignone, S.; Ferrando, S.; Candiani, S.; Bozzo, M.; Bertolino, M.; Costa, G.; Bavestrello, G.; Scarfi, S. Insights into the Evolution of Metazoan Regenerative Mechanisms: Roles of TGF Superfamily Members in Tissue Regeneration of the Marine Sponge Chondrosia Reniformis. *Journal of Experimental Biology* **2019**, *222*, jeb207894, doi:10.1242/jeb.207894.

Pozzolini, M.; Scarfi, S.; Gallus, L.; Castellano, M.; Vicini, S.; Cortese, K.; Gagliani, M.C.; Bertolino, M.; Costa, G.; Giovine, M. Production, Characterization and Biocompatibility Evaluation of Collagen Membranes Derived from Marine Sponge Chondrosia Reniformis Nardo, 1847. *Marine Drugs* **2018a**, *16*, 111, doi:10.3390/md16040111.

Pozzolini, M.; Scarfi, S.; Ghignone, S.; Mussino, F.; Vezzulli, L.; Cerrano, C.; Giovine, M. Molecular Characterization and Expression Analysis of the First Porifera Tumor Necrosis Factor

Superfamily Member and of Its Putative Receptor in the Marine Sponge Chondrosia Reniformis. *Dev Comp Immunol* **2016**, *57*, 88–98, doi:10.1016/j.dci.2015.12.011.

Pozzolini, M.; Scarfì, S.; Giovine, M. Marine Collagen and Its Biotechnological Applications. In *Encyclopaedia of Marine Biotechnology*; John Wiley & Sons, Ltd, 2020; pp. 1007–1030 ISBN 978-1-119-14380-2.

Pozzolini, M.; Scarfì, S.; Mussino, F.; Salis, A.; Damonte, G.; Benatti, U.; Giovine, M. *Pichia Pastoris* Production of a Prolyl 4-Hydroxylase Derived from *Chondrosia Reniformis* Sponge: A New Biotechnological Tool for the Recombinant Production of Marine Collagen. *Journal of Biotechnology* **2015**, *208*, 28–36, doi:10.1016/j.jbiotec.2015.05.007.

Pozzolini, M.; Tassara, E.; Dodero, A.; Castellano, M.; Vicini, S.; Ferrando, S.; Aicardi, S.; Cavallo, D.; Bertolino, M.; Petrenko, I.; et al. Potential Biomedical Applications of Collagen Filaments Derived from the Marine Demosponges Ircinia Oros (Schmidt, 1864) and Sarcotragus Foetidus (Schmidt, 1862). *Marine Drugs* **2021**, *19*, 563, doi:10.3390/md19100563.

Prade, I.; Schröpfer, M.; Seidel, C.; Krumbiegel, C.; Hille, T.; Sonntag, F.; Behrens, S.; Schmieder, F.; Voigt, B.; Meyer, M. Human Endothelial Cells Form an Endothelium in Freestanding Collagen Hollow Filaments Fabricated by Direct Extrusion Printing. *Biomaterials and Biosystems* **2022**, *8*, 100067, doi:10.1016/j.bbiosy.2022.100067.

Pradillon, F.; Gaill, F. Adaptation to deep-sea hydrothermal vents: some molecular and developmental aspects. *Journal of Marine Science and Technology* **2007**, *15*, doi:10.51400/2709-6998.2057.

Prajaputra, V.; Isnaini, N.; Maryam, S.; Ernawati, E.; Deliana, F.; Haridhi, H.A.; Fadli, N.; Karina, S.; Agustina, S.; Nurfadillah, N.; et al. Exploring Marine Collagen: Sustainable Sourcing, Extraction Methods, and Cosmetic Applications. *South African Journal of Chemical Engineering* **2024**, *47*, 197–211, doi:10.1016/j.sajce.2023.11.006.

Proksch, E.; Schunck, M.; Zague, V.; Segger, D.; Degwert, J.; Oesser, S. Oral Intake of Specific Bioactive Collagen Peptides Reduces Skin Wrinkles and Increases Dermal Matrix Synthesis. *Skin Pharmacology and Physiology* **2013**, *27*, 113–119, doi:10.1159/000355523.

Quigley, A.S.; Bancelin, S.; Deska-Gauthier, D.; Légaré, F.; Veres, S.P.; Kreplak, L. Combining Tensile Testing and Structural Analysis at the Single Collagen Fibril Level. *Sci Data* **2018**, *5*, 180229, doi:10.1038/sdata.2018.229.

Ramirez, F.; Boast, S.; Dalessio, M.; Lee, B.; Prince, J.; Su, M.; Vissing, H.; Yoshioka, H. Fibrillar Collagen Genes - Structure and Expression in Normal and Diseased States. *Ann. N.Y. Acad. Sci.* **1990**, *580*, 74–80, doi:10.1111/j.1749-6632.1990.tb17919.x.

Ran, X.-G.; Wang, L.-Y. Use of Ultrasonic and Pepsin Treatment in Tandem for Collagen Extraction from Meat Industry By-Products. *Journal of the Science of Food and Agriculture* **2014**, *94*, 585–590, doi:10.1002/jsfa.6299.

Rastian, Z.; Pütz, S.; Wang, Y.; Kumar, S.; Fleissner, F.; Weidner, T.; Parekh, S.H. Type I Collagen from Jellyfish *Catostylus mosaicus* for Biomaterial Applications. *ACS Biomater. Sci. Eng.* **2018**, *4*, 2115–2125, doi:10.1021/acsbiomaterials.7b00979.

Reitner, J.; Mehl, D. Monophyly of the Porifera. *Verhandlungen des Naturwissenschaftlichen Vereins in Hamburg* **1996**, 5–32, doi:10.23689/fidgeo-2563.

Ren, Y.; Fan, L.; Alkildani, S.; Liu, L.; Emmert, S.; Najman, S.; Rimashevskiy, D.; Schnettler, R.; Jung, O.; Xiong, X.; et al. Barrier Membranes for Guided Bone Regeneration (GBR): A Focus on Recent Advances in Collagen Membranes. *International Journal of Molecular Sciences* **2022**, *23*, 14987, doi:10.3390/ijms232314987.

Rezvani Ghomi, E.; Nourbakhsh, N.; Akbari Kenari, M.; Zare, M.; Ramakrishna, S. Collagen-Based Biomaterials for Biomedical Applications. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **2021**, *109*, 1986–1999, doi:10.1002/jbm.b.34881.

Ricard-Blum, S. The Collagen Family. *Cold Spring Harbor Perspectives in Biology* **2011**, *3*, a004978–a004978, doi:10.1101/cshperspect.a004978.

Ricard-Blum, S.; Ballut, L. Matricryptins Derived from Collagens and Proteoglycans. *Front Biosci (Landmark Ed)* **2011**, *16*, 674–697, doi:10.2741/3712.

Ricard-Blum, S.; Ruggiero, F. The Collagen Superfamily: From the Extracellular Matrix to the Cell Membrane. *Pathologie Biologie* **2005**, *53*, 430–442, doi:10.1016/j.patbio.2004.12.024.

Ricard-Blum, S.; Salza, R. Matricryptins and Matrikines: Biologically Active Fragments of the Extracellular Matrix. *Experimental Dermatology* **2014**, *23*, 457–463, doi:10.1111/exd.12435.

Riesgo, A.; Farrar, N.; Windsor, P.J.; Giribet, G.; Leys, S.P. The Analysis of Eight Transcriptomes from All Poriferan Classes Reveals Surprising Genetic Complexity in Sponges. *Molecular Biology and Evolution* **2014**, *31*, 1102–1120, doi:10.1093/molbev/msu057.

Riesgo, A.; Maldonado, M. Differences in Reproductive Timing among Sponges Sharing Habitat and Thermal Regime. *Invertebrate Biology* **2008**, *127*, 357–367, doi: 10.1093/mol10.1111/j.1744-7410.2008.00128.x.

Rossert, J.; de Crombrugghe, B. Chapter 12 - Type I Collagen: Structure, Synthesis, and Regulation. In *Principles of Bone Biology (Second Edition)*; Bilezikian, J.P., Raisz, L.G., Rodan, G.A., Eds.; Academic Press: San Diego, 2002; pp. 189–XVIII ISBN 978-0-12-098652-1.

Roveta, C.; Annibaldi, A.; Calcinai, B.; Girolametti, F.; Illuminati, S.; Pulido Mantas, T.; Truzzi, C.; Puce, S. Distribution of Mercury inside the Mediterranean Sponge Chondrosia Reniformis: A Study Case from the Tuscan Archipelago National Park (Tyrrhenian Sea). *Journal of Sea Research* **2022**, *184*, 102206, doi:10.1016/j.seares.2022.102206.

Roveta, C.; Calcinai, B.; Girolametti, F.; Fernandes Couceiro, J.; Puce, S.; Annibaldi, A.; Costa, R. The Prokaryotic Community of Chondrosia Reniformis Nardo, 1847: From Diversity to Mercury Detection. *Zoology* **2023**, *158*, 126091, doi:10.1016/j.zool.2023.126091.

Runnegar, B. Collagen Gene Construction and Evolution. *J Mol Evol* **1985**, *22*, 141–149, doi:10.1007/BF02101692.

Rutschmann, C.; Baumann, S.; Cabalzar, J.; Luther, K.B.; Hennet, T. Recombinant Expression of Hydroxylated Human Collagen in Escherichia Coli. *Appl Microbiol Biotechnol* **2014**, *98*, 4445–4455, doi:10.1007/s00253-013-5447-z.

Sae-leaw, T.; O'Callaghan, Y.C.; Benjakul, S.; O'Brien, N.M. Antioxidant, Immunomodulatory and Antiproliferative Effects of Gelatin Hydrolysates from Seabass (Lates Calcarifer) Skins. *International Journal of Food Science & Technology* **2016**, *51*, 1545–1551, doi:10.1111/ijfs.13123.

Sato, K.; Tanahashi-Shiina, T.; Jun, F.; Watanabe-Kawamura, A.; Ichinomiya, M.; Minegishi, Y.; Tsukamasa, Y.; Nakamura, Y.; Kawabata, M.; Ohtsuki, K. Simple and Rapid Chromatographic Purification of Type V Collagen from a Pepsin Digest of Porcine Intestinal Connective Tissue, an Unmanageable Starting Material for Conventional Column Chromatography. *Journal of Chromatography B* **2003**, *790*, 277–283, doi:10.1016/S1570-0232(03)00083-7.

Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. Database resources of the national centre for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-D26. doi: 10.1093/nar/gkab1112. PMID: 34850941; PMCID: PMC8728269.

Scarfì, S.; Pozzolini, M.; Oliveri, C.; Mirata, S.; Salis, A.; Damonte, G.; Fenoglio, D.; Altosole, T.; Ilan, M.; Bertolino, M.; et al. Identification, Purification and Molecular Characterization of Chondrosin, a New Protein with Anti-Tumoral Activity from the Marine Sponge Chondrosia Reniformis Nardo 1847. *Marine Drugs* **2020**, *18*, 409, doi:10.3390/md18080409.

Schierwater, B.; DeSalle, R. *Invertebrate Zoology: A Tree of Life Approach*; CRC Press, 2021; ISBN 978-1-4822-3582-1.

Schmidt, M.M.; Dornelles, R.C.P.; Mello, R.O.; Kubota, E.H.; Mazutti, M.A.; Kempka, A.P.; Demiate, I.M. Collagen Extraction Process. *International Food Research Journal* **2016**, *23*, 913–922.

Senadheera, T.R.L.; Dave, D.; Shahidi, F. Sea Cucumber Derived Type I Collagen: A Comprehensive Review. *Marine Drugs* **2020**, *18*, 471, doi:10.3390/md18090471.

Shekhter, A.B.; Fayzullin, A.L.; Vukolova, M.N.; Rudenko, T.G.; Osipycheva, V.D.; Litvitsky, P.F. Medical Applications of Collagen and Collagen-Based Materials. *Current Medicinal Chemistry* **2019**, *26*, 506–516, doi:10.2174/0929867325666171205170339.

Shen, C.; Zhang, G.; Wang, Q.; Meng, Q. Fabrication of Collagen Gel Hollow Fibers by Covalent Cross-Linking for Construction of Bioengineering Renal Tubules. *ACS Appl. Mater. Interfaces* **2015**, *7*, 19789–19797, doi:10.1021/acsami.5b05809.

Shimizu, K., and Morse, D.E. Chapter Fourteen - Silicatein: A Unique Silica-Synthesizing Catalytic Triad Hydrolase from Marine Sponge Skeletons and Its Multiple Applications *Methods in Enzymology*, vol. 605, B. S. Moore, A c. di, in Marine Enzymes and Specialized Metabolism - Part B, vol. 605, Academic Press, 2018, pp. 429–455. doi: 10.1016/bs.mie.2018.02.025.

Shoulders, M.D.; Raines, R.T. Collagen Structure and Stability. *Annual Review of Biochemistry* **2009**, *78*, 929–958, doi:10.1146/annurev.biochem.77.032207.120833.

Sicot, F.-X.; Mesnage, M.; Masselot, M.; Exposito, J.-Y.; Garrone, R.; Deutsch, J.; Gaill, F. Molecular Adaptation to an Extreme Environment: Origin of the Thermal Stability of the Pompeii Worm Collagen1. *Journal of Molecular Biology* **2000**, *302*, 811–820, doi:10.1006/jmbi.2000.4505.

Silva, J.C.; Barros, A.A.; Aroso, I.M.; Fassini, D.; Silva, T.H.; Reis, R.L.; Duarte, A.R.C. Extraction of Collagen/Gelatin from the Marine Demosponge *Chondrosia reniformis* (Nardo, 1847) Using Water Acidified with Carbon Dioxide – Process Optimization. *Ind. Eng. Chem. Res.* **2016**, *55*, 6922–6930, doi:10.1021/acs.iecr.6b00523.

Silva, T.H.; Moreira-Silva, J.; Marques, A.L.P.; Domingues, A.; Bayon, Y.; Reis, R.L. Marine Origin Collagens and Its Potential Applications. *Marine Drugs* **2014**, *12*, 5881–5901, doi:10.3390/md12125881.

Simpson, T.L. *The Cell Biology of Sponges*; Springer Science & Business Media, 2012; ISBN 978-1-4612-5214-6.

Sionkowska, A.; Adamiak, K.; Musiał, K.; Gadomska, M. Collagen Based Materials in Cosmetic Applications: A Review. *Materials* **2020**, *13*, 4217, doi:10.3390/ma13194217.

Sivan, S.S.; Tsitron, E.; Wachtel, E.; Roughley, P.; Sakkee, N.; van der Ham, F.; Degroot, J.; Maroudas, A. Age-Related Accumulation of Pentosidine in Aggrecan and Collagen from Normal and Degenerate Human Intervertebral Discs. *Biochemical Journal* **2006**, *399*, 29–35, doi:10.1042/BJ20060579.

Skočilas, J.; Žitný, R.; Štancl, J.; Dostál, M.; Landfeld, A.; Houška, M. Rheological Properties of Collagen Matter Predicted Using an Extrusion Rheometer. *Journal of Texture Studies* **2016**, *47*, 514–522, doi:10.1111/jtxs.12194.

Söderhäll, C.; Marenholz, I.; Kerscher, T.; Rüschendorf, F.; Esparza-Gordillo, J.; Worm, M.; Gruber, C.; Mayr, G.; Albrecht, M.; Rohde, K.; et al. Variants in a Novel Epidermal Collagen Gene (COL29A1) Are Associated with Atopic Dermatitis. *PLOS Biology* **2007**, *5*, e242, doi:10.1371/journal.pbio.0050242.

Song, E.; Yeon Kim, S.; Chun, T.; Byun, H.-J.; Lee, Y.M. Collagen Scaffolds Derived from a Marine Source and Their Biocompatibility. *Biomaterials* **2006**, *27*, 2951–2961, doi:10.1016/j.biomaterials.2006.01.015.

Stein, H.; Wilensky, M.; Tsafrir, Y.; Rosenthal, M.; Amir, R.; Avraham, T.; Ofir, K.; Dgany, O.; Yayon, A.; Shoseyov, O. Production of Bioactive, Post-Translationally Modified, Heterotrimeric, Human Recombinant Type-I Collagen in Transgenic Tobacco. *Biomacromolecules* **2009**, *10*, 2640–2645, doi:10.1021/bm900571b.

Stolfi, A.; Brown, F.D. Tunicata. In *Evolutionary Developmental Biology of Invertebrates 6: Deuterostomia*; Wanninger, A., Ed.; Springer: Vienna, 2015; pp. 135–204 ISBN 978-3-7091-1856-6.

Subhan, F.; Hussain, Z.; Tauseef, I.; Shehzad, A.; Wahid, F. A Review on Recent Advances and Applications of Fish Collagen. *Critical Reviews in Food Science and Nutrition* **2021**, *61*, 1027–1037, doi:10.1080/10408398.2020.1751585.

Sun, S.; Genovese, F.; Karsdal, M.A. Chapter 6 - Type VI Collagen. In *Biochemistry of Collagens, Laminins and Elastin (Third Edition)*; Karsdal, M.A., Ed.; Academic Press, 2024; pp. 61–73 ISBN 978-0-443-15617-5.

Suo-Lian, W.; Huai-Bin, K.; Dong-Jiao, L. Technology for extracting effective components from fish scale. *J. Food Sci.* **2017**, *7*(10.17265), 2159-5828.

Swatschek, D.; Schatton, W.; Kellermann, J.; Müller, W.E.G.; Kreuter, J. Marine Sponge Collagen: Isolation, Characterization and Effects on the Skin Parameters Surface-pH, Moisture and Sebum. *European Journal of Pharmaceutics and Biopharmaceutics* **2002**, *53*, 107–113, doi:10.1016/S0939-6411(01)00192-8.

Tapia-Vasquez, A.E.; Ezquerra-Brauer, J.M.; Martínez-Cruz, O.; Márquez-Ríos, E.; Ramírez-Guerra, H.E.; Minjarez-Osorio, C.; Torres-Arreola, W. Relationship between Muscle Texture and the

Crosslinking Degree of Collagen Fibers from Octopus (*Octopus vulgaris*), Guitarfish (*Rhinobatos productus*), and Cazon (*Mustelus lunulatus*). *Journal of Food Processing and Preservation* **2019**, *43*, e14241, doi:10.1111/jfpp.14241.

Tassara, E.; Orel, B.; Ilan, M.; Cavallo, D.; Dodero, A.; Castellano, M.; Vicini, S.; Giovine, M.; Pozzolini, M. Seasonal Molecular Difference in Fibrillar Collagen Extracts Derived from the Marine Sponge Chondrosia Reniformis (Nardo, 1847) and Their Impact on Its Derived Biomaterials. *Marine Drugs* **2023**, *21*, 210, doi:10.3390/md21040210.

Tassara, E.; Oliveri, C.; Vezzulli, L.; Cerrano, C.; Xiao, L.; Giovine, M.; Pozzolini, M. 2D Collagen Membranes from Marine Demosponge *Chondrosia reniformis* (Nardo, 1847) for Skin-Regenerative Medicine Applications: An *In Vitro* Evaluation. *Marine Drugs* **2023**, *21*, 428, doi:10.3390/md21080428.

Teuscher, A.C.; Jongsma, E.; Davis, M.N.; Statzer, C.; Gebauer, J.M.; Naba, A.; Ewald, C.Y. The *in-Silico* Characterization of the *Caenorhabditis Elegans* Matrisome and Proposal of a Novel Collagen Classification. *Matrix Biology Plus* **2019**, *1*, 100001, doi:10.1016/j.mbplus.2018.11.001.

Tinghino, a. Setting optimal condition for rearing the collagen-rich sponge *Chondrosia reniformis* Nardo, 1847 in aquarium for production of sponge biomass.

Tomita, M.; Munetsuna, H.; Sato, T.; Adachi, T.; Hino, R.; Hayashi, M.; Shimizu, K.; Nakamura, N.; Tamura, T.; Yoshizato, K. Transgenic Silkworms Produce Recombinant Human Type III Procollagen in Cocoons. *Nat Biotechnol* **2003**, *21*, 52–56, doi:10.1038/nbt771.

Torre-Blanco, A.; Toledo, I. The Isolation, Purification, and Characterization of the Collagen of *Cysticercus cellulosae*. *Journal of Biological Chemistry* **1981**, *256*, 5926–5930, doi:10.1016/S0021-9258(19)69298-7.

Tucker, R.P.; Adams, J.C. Molecular Evolution of the Thrombospondin Superfamily. *Seminars in Cell & Developmental Biology* **2024**, *155*, 12–21, doi:10.1016/j.semcdb.2023.05.004.

Uriarte-Montoya, M.H.; Arias-Moscoso, J.L.; Plascencia-Jatomea, M.; Santacruz-Ortega, H.; Rouzaud-Sández, O.; Cardenas-Lopez, J.L.; Marquez-Rios, E.; Ezquerra-Brauer, J.M. Jumbo Squid (*Dosidicus gigas*) Mantle Collagen: Extraction, Characterization, and Potential Application in the Preparation of Chitosan–Collagen Biofilms. *Bioresource Technology* **2010**, *101*, 4212–4219, doi:10.1016/j.biortech.2010.01.008.

Van B. Robertson, W. Metabolism of Collagen in Mammalian Tissues. *Biophysical Journal* **1964**, *4*, 93–106, doi:10.1016/S0006-3495(64)86930-7.

Van Der Rest, M.; Garrone, R. Collagen Family of Proteins. *The FASEB Journal* **1991**, *5*, 2814–2823, doi:10.1096/fasebj.5.13.1916105.

Van Wart, H.E.; Steinbrink, D.R. A Continuous Spectrophotometric Assay for *Clostridium Histolyticum* Collagenase. *Analytical Biochemistry* **1981**, *113*, 356–365, doi:10.1016/0003-2697(81)90089-0.

Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; De Paepe, A.; Speleman, F. Accurate Normalization of Real-Time Quantitative RT-PCR Data by Geometric Averaging of Multiple Internal Control Genes. *Genome Biol* **2002**, *3*, research0034.1, doi:10.1186/gb-2002-3-7-research0034.

Verrecchia, F.; Mauviel, A. Transforming Growth Factor-β and Fibrosis. *World Journal of Gastroenterology: WJG* **2007**, *13*, 3056, doi:10.3748/wjg.v13.i22.3056.

Vuorio, E.; Crombrugghe, B. de THE FAMILY OF COLLAGEN GENES. *Annual Review of Biochemistry* **1990**, *59*, 837–872, doi:10.1146/annurev.bi.59.070190.004201.

Wess, T.J. Collagen Fibril Form and Function. In *Advances in Protein Chemistry*; Fibrous Proteins: Coiled-Coils, Collagen and Elastomers; Academic Press, 2005; Vol. 70, pp. 341–374.

Whittaker, C.A.; Hynes, R.O. Distribution and Evolution of von Willebrand/Integrin A Domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. *MBoC* **2002**, *13*, 3369–3387, doi:10.1091/mbc.e02-05-0259.

Wilkie, I. C. Mutable collagenous tissue: overview and biotechnological perspective. *Echinodermata*, 2005; pp. 221-250.

Wilkie, I.C.; Parma, L.; Bonasoro, F.; Bavestrello, G.; Cerrano, C.; Carnevali, M.D.C. Mechanical Adaptability of a Sponge Extracellular Matrix: Evidence for Cellular Control of Mesohyl Stiffness in *Chondrosia Reniformis* Nardo. *Journal of Experimental Biology* **2006**, *209*, 4436–4443, doi:10.1242/jeb.02527.

Wilkie, I.C.; Sugni, M.; Gupta, H.S.; Carnevali, M.D.C.; Elphick, M.R. The Mutable Collagenous Tissue of Echinoderms: From Biology to Biomedical Applications. **2021**, doi:10.1039/9781839161124-00001.

Wilkinson, C.R.; Vacelet, J. Transplantation of Marine Sponges to Different Conditions of Light and Current. *Journal of Experimental Marine Biology and Ecology* **1979**, *37*, 91–104, doi:10.1016/0022-0981(79)90028-5.

Woodley, D.T.; Keene, D.R.; Atha, T.; Huang, Y.; Lipman, K.; Li, W.; Chen, M. Injection of Recombinant Human Type VII Collagen Restores Collagen Function in Dystrophic Epidermolysis Bullosa. *Nat Med* **2004**, *10*, 693–695, doi:10.1038/nm1063.

Yang, H., & Shu, Z. The extraction of collagen protein from pigskin. *J. Chem. Pharm. Res* **2014**, *6*(2), 683-687.

Yoneda, C.; Ahsan, Md.N.; Nakaya, M.; Matsubara, Y.; Ebihara, T.; Irie, S.; Uno, Y.; Hatae, K.; Watabe, S. Abalone Collagens: Immunological Properties and Seasonal Changes of Their mRNA Levels. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2000**, *126*, 59–68, doi:10.1016/S0305-0491(00)00173-5.

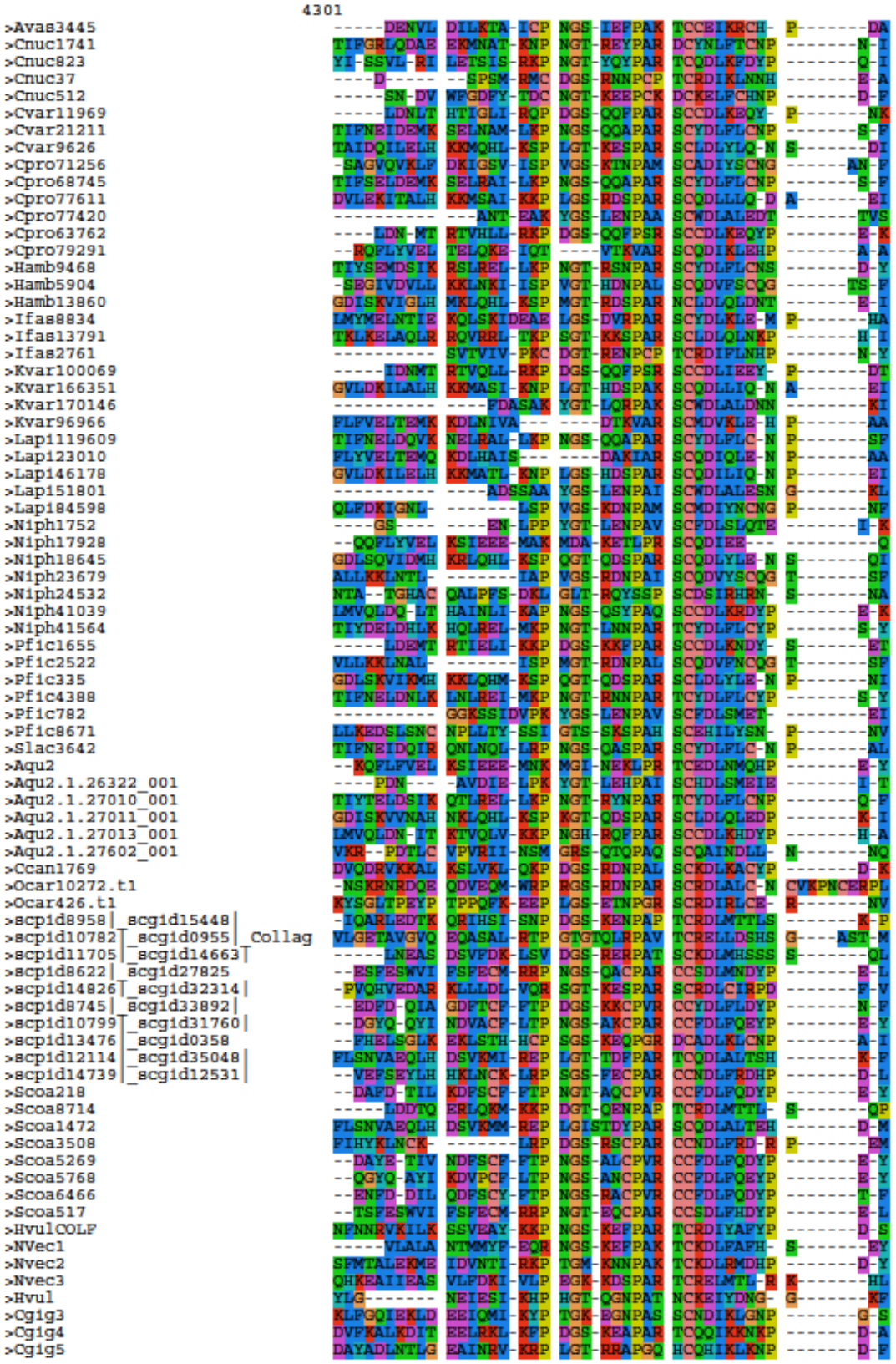
Yuan, F.-L.; Sun, Z.-L.; Feng, Y.; Liu, S.-Y.; Du, Y.; Yu, S.; Yang, M.-L.; Lv, G.-Z. Epithelial–Mesenchymal Transition in the Formation of Hypertrophic Scars and Keloids. *Journal of Cellular Physiology* **2019**, *234*, 21662–21669, doi:10.1002/jcp.28830.

Yuan, J.; Liu, M.; Yang, L.; Tu, G.; Zhu, Q.; Chen, M.; Cheng, H.; Luo, H.; Fu, W.; Li, Z.; et al. Acquisition of Epithelial-Mesenchymal Transition Phenotype in the Tamoxifen-Resistant Breast Cancer Cell: A New Role for G Protein-Coupled Estrogen Receptor in Mediating Tamoxifen Resistance through Cancer-Associated Fibroblast-Derived Fibronectin and B1-Integrin Signaling Pathway in Tumor Cells. *Breast Cancer Res* **2015**, *17*, 69, doi:10.1186/s13058-015-0579-y.

Zhang, M.; Chen, Y.; Li, G.; Du, Z. Rheological properties of fish skin collagen solution: Effects of temperature and concentration. *Korea-Aust. Rheol. J.* **2010**, *22*, 119–127.

Zhou, T.; Sui, B.; Mo, X.; Sun, J. Multifunctional and Biomimetic Fish Collagen/Bioactive Glass Nanofibers: Fabrication, Antibacterial Activity and Inducing Skin Regeneration in Vitro and in Vivo. *International Journal of Nanomedicine* **2017**, *12*, 3495–3507, doi:10.2147/IJN.S132459.

Supplementary 1 – COLFI alignments



Supplementary 2 – SMAD binding sequences in the upstream regions of the genes

>G1
TGACTGCAACTATCTTAGCCTAACCTGTCCGGTCAGGGTAAAAACCTGGATCTTTCTACCGAGAAACGGCAAAAACTTCTCCAGGAAGA
TATCGTCAAGATGGCGGACCTAGTCTAGCTGCAAGAACTCGAGCTAAGCACTAAACAATAGTCTTTGGACCGTTTGAGCACTTCCTTTG
CCGGCGCCCGCATCTAAACAATGAGTTGTAAAAACATTAGCGGCCAAGACAAAAATGCCAAATACTCGCTTGAGTATGTCAGGAATGTA
ATTATTTGCTTGCTAAAATCCGGATATAACTTATTCCGGAAGCATTTAGATTTAAGAACTATTGAACGTAAAATAATTTTGCAACTTAA
ATTCTCTCGACGAACACCGCAAGGATAAAACTCACCACTGTTAGCTAAAAATTTGCTAAAAATTTGTGAAAACTCTGTGAAGAGAACAT
TCAAAATAATTCCAATCGGCTATTGCAGCCAAACATTTTTTGCATTCTTTCAGACCCACATGATATGTAGCCATGAATATCATTACAT
AATTAGCACGCACGCCGCAAACAATAGACCGTCTCCTTTTGTAATCTCACAAAGAATACTCCAATTCCATTGTTAGGCTACTTTTGGGG
CGGCCATTGTGGTACATGGCTGCAACCAGGATTTTCCTGTAACAGAAGTGATCCCGTAGTTGTATAGGGTTCGAATAGCAGCCCATCTT
CTAGCTAGGTTGCTGTCGTTCTCGTTATCCAGCTACGAGCAGGAGTACCTGCTTCACCAAACA

>G2
TGTTTGGTGAAGCAGGTACTCCTGCTCGTAGCTGGATAACGAGAACGACAGCAACCTAGCTAGAAGATGGGCTGCTATTCGAACCCTAT
ACAACTACGGGATCACTTCTGTTACAGGAAAATCCTGGTTGCAGCCATGTACCACAATGGCCGCCCCAAAAGTAGCCTAACAATGGAAT
TGGAGTATTCTTTGTGAGATTACAAAAGGAGACGGTCTATTGTTTGCGGCGTGCGTGCTAATTATGTAATGATATTCATGGCTACATAT
CATGTGGGTCTGAAAGAATGCAAAAAATGTTTGGCTGCAATAAGCCGATTGGAATTATTTTGAATGTTCTCTTCACAGAGTTTTCACAA
ATTTTTAGCAAATTTTTAGCTAACAGTGGTGAGTTTTATCCTTGCGGTGTTCGTCGAGAGAATTTAAGTTGCAAAATTATTTTACGTTC
AATAGTTCTTAAATCTAAATGCTTCCGGAATAAGTTATATCCGGATTTTAGCAAGCAAATAATTACATTCCTGACATACTCAAGCGAGT
ATTTGGCATTTTTGTCTTGGCCGCTAATGTTTTTACAACTCATTGTTTAGATGCGGGCGCCGGCAAAGGAAGTGCTCAAACGGTCCAAA
GACTATTGTTTAGTGCTTAGCTCGAGTTCTTGCAGCTAGACTAGGTCCGCCATCTTGACGATATCTTCCTGGAGAAGTTTTTGCCGTTT
CTCGGTAGAAAGATCCAGGTTTTTACCCTGACCGGACAGGTTAGGCTAAGATAGTTGCAGTCA

>G3
TTACTTGTGAGGCTGTTCAAAATGTAATTCGTACTTCAAATACGTATACGTACTAACTGACATATGACTGTGTAGTCTTTACTTTATTT
GCTGTATATTCATCTATATATAGCTATTTACAGAAGAATTGTTATGGAAATCAAAAGTCAGTTACCTAGCAACATGTTCAATTAAGAAA
GAAGGACTGAGTGGTAGGGGATTATTAAAATTATTGGCTTAACCCATTCAGGTGGGAAAAGATAGTGAAGGAAAGTATTGGATTAAGGG
AAATTTGCATGGAATCAGCTGTGAGCAACTGCAACAGTAACTGTAAAGGAACTGAACTACATTTGTTCAATTGGACTGCGAGATTTAGC
ATGTCTACCAGTACCTAGTAACTAGTAATTACCATAGTTCATTTAGGGAGGTGTGCAGTACATGCTTAAAAGGTTCTCTATCCTATAGA
TCAGTTTACCTTGCTATTTGAATGTGACACTTTCAATCCCTTTGAGCAAGTTTATTGACATTAGTTAGTAGATAAATGGTAAAACTGTT
CAATTTTTCTTGTTGCAGTTGCTAATGGGATAATTACTCTTCATTCATTATTATATTCATTGTAATTAGATAAAGGAATACATTAAATT
TGTCCAACAGGTGTACCTGATTTAGGGTCCATTCCAGCTCCCAGTAGCTATGGCTACACTCATTAATTAAGTAGTAGAGTGCTTACAGT
GCTGGCTCATCTGAGGTTTAAATAAAGCAATGAACATTAAGTTATGTCTAAATGTGCTGTTATATGTGTGTGTAATGTAATACACATGT
ATATTCAGCCTTTAACAGGAACAGTACACTTTCTATATTGCCCTGCCTGGACTGGTAATGTACATAACTTCTGTGACTACGTATGTATC
TGTCTATGTACGTATGTTTACCATAGCAACTGACTTTAGGAGGATATTCCTCTTCATTAATTTTCGTGTGATGTTTTTGTTATTACTAA
GCTAGTAAATGGGTTCATGTAGCCTATTGTCTGAATTCTCTCAATGGCTGTCAACTTACAACTATAGAATAATGGATGTTCTTAAGATC
ACTTGCAAAACAATTTATTCATATTGCCATGGTAACATCTCATTACCAAATTACAAATGACAGCAAAAAGTGTTTAATTACCTAAAGTT
GGATGATAAATTAAATATTCATCAATTATACATATAAATCTGACAGTAGATTTATCTTAGAGTCATAACTAAAATGTACATTTTGTTTT
CTCTTCCCACAGCCCAGTTTCATCTCCAACAAATTTAAATCAAATGAAAGTTTCCATAGCAACCATTTCGATGACAACAAATCCAATAA
GTTTGCCATTAGTCTAGTAAAGTAAGGTTAATACTTTATGAGACACCTAGATATAGCTCATGGACATGTGCACCCTGTTTAAAGGAATT
CAGACTATTTGGACTTGCCTTAAACTCTTCACCATATATTATATCCTACGTATATAAGCATTAGAGGTACATTTTACAGTATATATATG
TATTTACATAACTCTTGTGATTTACATATTGCATTAATTAAGGTGTTGGCCTTATTTCAGGCATATATATATGCCTTTGTTAGCAACAA
CCTCAACTTTGAAGCAAATTAATTAGTTCTGAGCTTGAGGGGATTAAAATTGGGCCCTTTTCAGACATCATTAATTAGCATGGTAAATT
TTATGTTTATTAAGAGCTCTTGCTTATATGTGTACTAGTGCTCTTCCAATTACTATTAAGACTGTAATAATGTGTAGGGTTCAATATTC
TATTTATTTGAACTATATACTTATATTTTCAAGACATATATATACGTAGGTTCATACCTTCCTCTGCAAACCACCTTCTAATATACTTA
TTTTTCCATTTCAAGTGGTTTATTGAACTTTCCATACTATCATAAATTAAATTAGATGCACAAAATTTTTATAAAATTAATTACATTGT
TATGAAGAAATAAAAACTTCAGACATTATAATGAATATTAATT

>G4
TGGTCCTTCTCTGATACTCCCTCCAGTTCTCTGATGGCCTCTGGTAGCTGCTCTACACACAACAAAAACAATAAACATCAACAAACTTA
AGGACAAAGGTGATTCTACTTACACACAACAAAAAATAAACATCAACAAACTTAAGGACAAAGGTGATTCTACTTACACACAACAAAA
ACAATAAACATCAACAAACTTAGGGACAAAGGTGATTCTACTTACACACAACAAAAACAATAAACATCAACAAACTTAGGGACAAAGGT
GATTCTACTTACACACAACAAAAACAATAAACATCAACAAACTTAGGGACAAAGGTGATTCTACTTACACACAACAAAAACAATAAACA
TCAACAAACTTAGGGACAAAGGTGATTCTACTTACACACAACAAACCTACCTATCTAAGAATACTGAAATACTATCAATCAATCATTCA
CTACTTCTTTAACTGGGAGCCCCTTATACAAGCTACTTAACATACGTACTAATGGATGTCCTCTCAAATATATGAACTTTCCCATTGAA
CATTTCCTACTTACTTTCCAATCATGTTTTTTATATGTTTTTTTTATGTATTTTTTACAATGATTTACAATATAATAATTGAATACTAT
ATATAAATAAGGGAATTTTCAATCATATGACATTGTGTTGTATTTACTGACAGATTGATACATAGTACACATTATTTATAGGGGGTAAA
ACCCTTAGCCAATTTACAGTGGGCGGGGCCAGGTCAGGTACTTGATAACCACTGCACTATACTACAGATTAAAAATCAAACCACATAGA
GTGTCTACAATTATAAAGTGTTTCTAACCTTCCAGAACAAGAGCAAACGCCTTCCAAAATACTAGCACCGGATCCGATCCATACTTTTC
GATAAACTGCAACACGGCTGTTTGGGCGTGCCGGTACAGACCCTCCCGAAGATAGTGCGAGATGACTCCCTACACACACGAGTACGAGG

Acknowledgments

This work was possible thanks to the Genoa University and the Centro del Mare research funding. I would like to express my gratitude to my supervisors prof. Marina Pozzolini and prof. Marco Giovine for their guidance, and to everyone that made this possible throughout the years. A special thanks goes to the Ana Riesgo Lab, in which a fundamental part of this thesis was realized.