Ctenophore mesoglea: building a mucus-like body

Sara Siwiecki[1,2,3], Stephanie Archer-Hartmann[4], Ian Black[4],
Jiri Vlach[4], Parastoo Azadi[4], Casey Dunn[6], Alison Sweeney[2,5,6*]

[1]Molecular Biophysics & Biochemistry, Yale University, New Haven, CT 06520, USA.

[2]Quantitative Biology Institute, Yale University, New Haven, CT 06520, USA.

[3]Integrated Graduate Program in Physical and Engineering Biology, Yale University, New Haven, CT 06520, USA.

[4]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA.

[5]Department of Physics, Yale University, New Haven, CT 06511, USA.

[6]Department of Ecology & Evolutionary Biology, Yale University, New Haven, CT 06520, USA.

*Corresponding author. Email: alison.sweeney@yale.edu

Ctenophores are one of our most distant animal relatives and highly abundant and widespread marine predators, yet much of their biology remains undiscovered. Their bodies are dominated by a gelatinous, largely extracellular mesoglea of unknown composition. Ctenophores lack homologs for fibrous collagens that form the typical metazoan extracellular matrix (ECM), so the composition of their extracellular material is unknown. Using spectroscopic analyses of *Mnemiopsis leidyi* and *Pleurobrachia pileus* mesoglea, we found abundant mucus-related proteins and sulfated polysaccharides. The mucins and glycans appear unlinked, unlike the typical heavily glycosylated mucins. Our results suggest that ctenophores have a mucus-like mesoglea, a marked contrast from standard collagenous ECMs and mucus. This study indicates that ctenophores are even more different from other animals than previously appreciated, and expands our understanding of soft body composition and biophysics in animals.

Ctenophores are one of our most distant animal relatives Fig. 1A and have an impressive ability to live in nearly every environment of the world's oceans, ranging across wide temperatures, pressures, and salinities (*1, 2*). Surviving these environmental stressors throughout their history requires specific adaptations to their biochemistry and biophysical structures. However, despite their importance to evolutionary biology and ubiquity in marine environments, many fundamental aspects of their structure and physiology are unexplored.

The ctenophore body consists of two thin epithelial layers, the epidermis and the gastrodermis, with a voluminous gelatinous layer of mesoglea between the two, acting as an internal "skeleton" (Fig. 1C and 1D). The mesoglea is a hydrogel containing sparse, isolated cells interspersed with bulk extracellular material of indeterminate composition. The gelatinous mesoglea forms ~99% of the organism's volume, yet few studies have reported on its composition or role despite its clear importance to the animal's body.

Franc (1985) provided the first, and only, focused description of the biochemical composition of the mesoglea (*3*). Franc hypothesized that the mesoglea is formed by a loose three-dimensional network of glycoproteins and sulfated polysaccharides. Franc (1985) and others have also observed sparse "bundle" structures containing collagen-like polymers, the canonical extracellular matrix (ECM) protein family (*3–5*). However, ctenophores lack the genes coding for most known collagens, including fibrous collagens that make up the mesoglea of other gelatinous animals (*6, 7*). Ctenophores only have isoforms of collagen type IV, which is typically enriched only in the basal lamina (epithelia) of metazoans, so there is no direct structural evidence for a collagen type IV network in the mesoglea (*6–8*). Additionally, ctenophores appear to lack fibronectin or tenascin, making the expanded extracellular material of the mesoglea markedly different from other metazoan (*7*). Imaging studies, such as those from Norekian & Moroz (2018), have revealed that the inner mesoglea is not completely acellular and contains sparse neural and muscle cell fibers that extend from epithelial layers into the mesoglea (*9*). Lastly, whole ctenophore body measurements reveal some information on the mesoglea since the mesoglea dominates the body composition. The organic material has been previously assumed to be mostly protein, with relatively little lipid or carbohydrate (*10–18*). Most lipids in ctenophores are membrane phospholipids, so are unlikely to participate in the mesoglea structure (*19*). Elemental analyses show significant non-protein nitrogen, suggesting unknown glycoproteins whose carbohydrate moieties remain uncharacterized (*12*).

Therefore, the mesoglea likely contains sparse cells plus glycoprotein, neuromuscular, sulfated polysaccharide, and collagen-like components, but its precise composition remains unknown.

Despite their key evolutionary position and extracellular-rich bodies, there has been no modern chemical characterization of their mesoglea, and such fundamental information about their physiology and extracellular architecture may be essential for understanding the transitions to multicellularity in animals. Thus, we present the first thorough biochemical characterization of ctenophore mesoglea. We focus on *Mnemiopsis leidyi* and *Pleurobrachia pileus*, two species present in Long Island Sound and representing different clades within ctenophores, Nuda and Tentaculata, respectively (Fig. 1B and 1C). Curiously, we found that the mesoglea appears more akin to mucus than the traditional collagenous ECM of metazoans. For this study, we defined mucus to be the canonical viscoelastic aqueous solution secreted in many metazoans, rich in mucins, but also considered similarities to other types of non-mucin mucus, such as mucilage and biofilms. Since mucus-based materials are typically secreted in layers outside of epithelial tissues, this finding was surprising and potentially marks an original mucus function as a space-filling skeleton. We observed several unique features of mesoglea compared to other metazoan soft bodies: an abundance of mucus-related proteins and novel sulfated polysaccharides, likely producing a highly hydrophilic body.

Ctenophore mesoglea contains mucins

Firstly, we searched for extracellular proteins in mesoglea via MS-based proteomics. We used predicted proteins from the chromosome-level northern Atlantic *M. leidyi* genome as our search database for *M. leidyi* (*20*) and predicted proteins from the *H. californensis* genome for *P. pileus* (*1*). For each MS hit, we used BLASTp to annotate the potential biological function based on sequence identity with known proteins. Out of the extracellular proteins identified, mucus-related proteins constituted the most represented family (Fig. 2A and 2B), among many expected ECM proteins, such as laminins and integrins. Mucins, the canonical mucus protein, have high sequence polymorphism in their namesake ”mucin” domain (*21*), so sections of sequence similarity alone cannot define mucins, and protein domain architecture must be analyzed to determine possible canonical mucin function.

The field's consensus definition of a functional gel-forming mucin is the specific arrangement of several protein domains: von Willebrand factor type D (VWD), C8 (cysteine-rich), trypsin inhibitor-like (TIL), and Pro-Thr-Ser (PTS) domains (*21*). A protein that contains repeated VWD-C8-TIL domains (in this order) with a PTS domain at the C-terminus is considered a mucin. We analyzed the functional domains of our proteins with mucin-like homology using InterProScan and PTSPred search. While all mucin-like proteins had some of these domains, only one protein (or set of proteins in a tandem array) in each species had the correct arrangement of mucin-related domains (Fig. 2C). For *M. leidyi*, evm.model.c7.706 contains a partial VWD-C8-TIL repeat and is in a tandem array with evm.model.c7.707, which contains the remaining parts of the repeat and a short PTS domain. Interestingly, these two tandem genes are in a head-to-head orientation with gene evm.model.c7.705, which contains all the necessary mucin components; thus these genes may build a larger mucin protein complex if expressed together. This is also hypothesized in Lang et al. (2016), who identified mucin proteins in both *M. leidyi* and *P. bachei* ((*21*)). For *P. pileus*, we had an MS hit (Hcv1.av93.c7.g869.i1) using the *H. californensis* genome that contains all the mucin components arranged properly with a very short PTS domain. Similarly, a nearby gene, Hcv1.av93.c7.g874.i1, on the opposite strand has multiple PTS domains. A tandem gene immediately upstream also contains all mucin-related domains in proper arrangement, Hcv1.av93.c7.g868.i1. These genes may be expressed together and form potentially two mucin proteins. While we do not find all of these adjacent proteins in our proteomics results, the presence of at least one of them in each species indicates that there is a mucin-like protein that can possibly perform structural mucin functions.

The VWD, TIL, C8, and PTS domain architecture is a conserved feature of mucins in metazoans, but the function of these domains is not yet known, so it is difficult to expand the understanding of mucin structure to function. One well-established structure in mucins is the use of cysteines for disulfide bonds to enable polymerization via cysteine-rich domains, such as CysD, cysteine knots (CK), and C8. In human mucins, CysD domains are interspersed within the PTS regions and act as additional "linkers" for disulfide bonding (*22*). We did not identify any CysD domains in the PTS regions of the ctenophore mucins (Fig. 2C). However, a few human mucins, such as MUC19 and MUC6, also lack CysD domains within the PTS region, so this domain is not an absolute requirement for polymerization (*23*). Interestingly, having multiple CysD domains increases the stiffness of mucins and mucus-related molecules (*24*), so the lack of CysD in ctenophore mucins

is possibly related to the softness and fragility of ctenophore mesoglea. Next, CK motifs play roles in C-terminal dimerization, which is known to be the first step of the typical metazoan mucin assembly process (*25*). The C-terminus of ctenophore mucins appears to lack the CK motif, but ctenophore mucins do contain multiple CK sites throughout their VWD-TIL-C8 repeats along the N-terminal (separate from the C8 domains). These N-terminal cysteines are also important for dimerization, both via disulfide bonds and hydrophobic interactions (*22, 26*). Thus, ctenophores may exhibit alternative mucin polymerization compared to other metazoans, which could present functional differences, such as being less stiff. Considering a mucus-like mesoglea hypothesis, the role of mucins and mucin-related proteins in ctenophores is poorly described and has only been hypothesized to be part of the digestion and immune responses (*21*), so closer analysis of these mucin proteins in ctenophores is necessary.

The other identified mucus-related proteins (Fig. 2A and 2B) do not have all typical mucin domains, but likely complex with other proteins to perform other mucus-related functions. Mucus often has an abundance of actins and myosins, among other cytoskeletal-related proteins, which play both structural and immune-related roles in the materials (*27, 28*). We identified numerous actins, myosins, and tubulins in ctenophore mesoglea (data S1 to S4), which may participate in unknown mucus functions, or may be debris of sheared neuromuscular fibers. We also identified many immune-related proteins, which likely act in concert with mucin proteins. Significant further research is required to understand the interplay between these proteins in a mucus structure.

Novel glycans likely participate in mesoglea structure

Glycoproteins, proteoglycans, glycosylaminoglycans (GAGs), and free polysaccharides are essential components of any soft biological tissue, so we searched for these in mesoglea via Nuclear Magnetic Resonance (NMR) Spectroscopy. Various NMR experiments were performed with two protease-treated *M. leidyi* mesoglea biological replicates, and spectra that produced solvable peaks are presented here. In Figure 3A, a representative ^{1}H,^{13}C-HSQC NMR spectrum for one of the *M. leidyi* samples highlights two sulfated galactose-based structures (other replicates and DEAE fractionation to confirm residues shown in fig. S1A and S1B; major peaks are described in table S1). The two structures are shown in Figure 3B: a sulfated galactan (residues A and B) and a sulfated Gal-Glc

disaccharide (residues C and D) bound to a hydroxylysine residue (Hyl, residue E). These groups are consistent with high galactose found via GC-MS (chromatograms shown in fig. S2A, quantified in table S2). Gal-Glc-Hyl without sulfation is common in collagens, but a sulfated Gal-Glc-Hyl residue has never been identified before. A discussion of how this structure may participate in the mesoglea is expanded upon more below.

In *P. pileus*, we found several sulfated mannan structures, as well as low levels of the same sulfated Gal-Glc-Hyl glycopeptide fragment. ^{1}H and ^{13}C-HSQC NMR spectra for a *P. pileus* sample are shown in Fig. 3C with mannan structures depicted in Fig. 3D. These structures are labeled M, N, and O in the HSQC spectrum and chemical shift assignment table S3, with different levels of sulfation and homo- or hetero-polymerization. Six additional, minor mannose-based residues were also identified (labeled P, Q, R, Mr, Np, and Oq), although we could resolve only a limited number of their signals, so chemical shift assignments for these residues are thus incomplete. These mannan structures are consistent with GC-MS results showing high mannose content (chromatograms shown in fig. S2B and quantified in table S2), although GC-MS identified some GalNAc and GlcNAc structures that we did not observe via NMR. These differences may result from low abundance being difficult to detect in different techniques or complex linkages to peptides. Notably, significant protein signals were observed via ^{1}H NMR despite extensive protease treatment (fig. S3), so this may indicate that there are complex conjugated macromolecules that are difficult to parse. Interestingly, GalNAc and GlcNAc are common modifications to mucins, so if ctenophore mesoglea contains a mucus structure, we would expect to find those residues.

Surprisingly, neither NMR nor glycoproteomics revealed protein modifications, linkages, or glycotransferases indicative of mucin glycosylation in *M. leidyi* or *P. pileus*. However, sample digestion produced a moderate amount of insoluble material that could not be analyzed via NMR, and glycosylated residues may therefore have been missed. O-glycosylation common to mucins can be difficult to identify without specialized sample treatments (such as mucinases), and glycosylation machinery is understudied (there have been no reports of glycosyltransferases in ctenophores, despite an assumption that all metazoans have glycoproteins), so targeted studies of mucins in ctenophores are still required (*29*). If ctenophore mucins are truly unglycosylated, short PTS domains (only 20-80 residues long, compared to repeats of 100+ amino acid long sequences in typical metazoan mucins) may explain the lack of linkage (*30*). The mucin proteins may instead

complex with glycans via non-covalent interactions to build a mucus structure, a known property of other mucus-based materials (*31–33*). Given the extraordinary diversity of both mucin sequences and glycans (*34*), clarifying ctenophore mucin glycosylation remains crucial for understanding mesoglea architecture

Mesoglea may have multiple organization levels

Via proteomics and NMR, we also observed evidence of the typical metazoan ECM scaffold collagens. We confirmed the presence of the only collagen type previously identified in ctenophores, basement membrane collagen (homologous to collagen type IV) (Fig. 2A and 2B and data S1 to S2). As mentioned, the Gal-Glc-Hyl structure found in *M. leidyi* (Fig. 3B) is common in collagens, including collagen type IV (*35*). We did not identify any other collagen types, so the Gal-Glc-Hyl structure likely originates from collagen type IV. Total amino acid analyses indicated low hydroxylysine concentration in *M. leidyi*, so collagen IV is not likely to be a major contributor to the mesoglea structure and probably comes from epithelial debris (table S4). Additionally, we observed only moderate levels of glycine and proline in the mesoglea for both species, and the amino acid content of collagenous tissues is typically dominated by glycine and proline, so it is unlikely that the mesoglea is primarily collagen-based (*36–38*).

Many other major structural proteins, including numerous neuromuscular proteins, were also identified via proteomics and may play roles in the mesoglea (data S1 to S4). The functions of many neuromuscular proteins, such as actins or myosins, are not well-described in the extracellular space, so it is not obvious how they complex with mucus. One hypothesis may be that the mesoglea is physically supported by a traditional neuromuscular network in an unorthodox way. Norekian & Moroz (2018) reported images showing muscle and neural fibers extending from ctene rows and other sections of epithelial layers far into the mesoglea (*9*). This neuromuscular structure is sparse compared to that of the epithelial layers but may provide a core network for mesoglea structure with additional support coming from other mesoglea components. Lastly, one of the top ten MS hits for both species had a significant identity with a unique extracellular protein known as mesocentin, which has only ever been described in the genus *Caenorhabditis*. Mesocentin is a large, extracellular glycoprotein containing multiple filamin domains and large disordered domains scattered in between

the filamin domains. The function of mesocentin is little described, but is believed to have adhesive properties (*39*), so it plausibly could play a role in the gelatinous mesoglea structure. Further investigations into mesocentin in ctenophores seem warranted to evaluate its abundance and role. Many MS spectra also went unidentified, so unidentified proteins could also play roles in the mesoglea structure. Functional structural proteins would need to be load-bearing and may have highly disordered regions or flexible protein domains, such as Von Willebrand factor domains to build the mesoglea.

Lastly, the sulfated carbohydrates and structural proteins could play non-network-forming structural roles by forming colloidal aggregates in a colloidal gel (*40,41*). Box-like extracellular structures adjacent to microtubules have been identified previously in the mesoglea of ctenophore tentacles (*5*), and Franc (1985) hypothesized that the mesoglea anatomy may consist of a loose 3D network aided by condensates. These structures have not been described since these reports, but they provide evidence of aggregates in mesoglea, and may be a critical avenue of future research to understand mesoglea organization.

Mucus as a ctenophore internal skeleton?

Because mucin was the only major protein identified that can form a gelatinous extracellular material, we propose that the mesoglea is a composite of mucin proteins and non-covalently bound polysaccharides (Fig. 4A), which has broad functional and evolutionary implications. Mucus can have diverse functions, including physical protection, chemical filtration, and lubrication. Some hypothesized functions of the mesoglea resemble mucus functions, including acting as a diffusive medium for chemical filtration and mechanical support by acting as a soft hydrostatic skeleton. However, animal mucus is typically secreted as a thin layer external to the organism, whereas mesoglea is a bulk, internal, voluminous tissue. Thus, mucus-like material in the mesoglea would represent a fundamentally distinct application of mucus. Voluminous, structural mucus is highly unusual, but one exception is the mucus "houses" of larvaceans, built with oikosins that only faintly resemble mucins (*25, 42*). Ctenophore mesoglea may be another example of atypical bulk mucus.

Our results suggest two evolutionary shifts. Firstly, while ctenophores have a basement membrane, they lack the collagen toolkit that build load-bearing interstitial ECM structures in other

metazoans (Fig. 4B). Fibrous collagen networks store elastic energy for muscle-driven locomotion, but ctenophores swim via cilia, so their body must dissipate, not store, energy. A loose carbohydrate-rich mucus gel meets this requirement, implying that interstitial collagen ECMs evolved after ctenophores diverged, or were lost in ctenophores. Secondly, ctenophore mucins appear less (or not at all) glycosylated than other metazoan mucins, yielding a more dissipative gel due to lack of covalent linkages, which is also advantageous for ciliary-based locomotion. This may represent a transition in animals from undecorated mucins to decorated mucins, either due to the trait only emerging after ctenophores diverged, or ctenophores losing this trait (Fig. 4C). In *M. leidyi* and *P. pileus*, the protein composition and sulfation patterns on protein-linked glycans are similar, but the composition of free glycans differs, perhaps reflecting ecological niches that future studies can investigate. Mucus is an adaptable biomaterial found in animals, plants (mucilage), and bacteria (biofilms), and may also be an early builder of metazoans, yet understudied compared to collagens. Further studies of ctenophore mucins will be enlightening both for ctenophore biology and for understanding the evolution of one of the most ancient and prevalent materials in biology.

Ctenophores are among the least understood animals. There is a long history of underestimating their complexity and misinterpreting their biology (*43*). They have nervous systems, which led zoologists to place them with other animals that have nervous systems (Cnidaria and Bilateria). More specifically, because they are gelatinous and fragile, they were placed with Cnidaria (which includes medusa jellyfish) in a clade called Coelenterata. This led to surprise when phylogenetic analyses indicated they diverged from other animals much earlier than did cnidaria (*44*). The more we learn about ctenophores, though, the more different we realize they are from other animals, explaining and resolving this surprise. Recent work, for example, has found that their nervous systems are radically different from those of other animals and include nerve nets with syncytial rather than synaptic connections (*45*). Our finding that ctenophore ECM is fundamentally different from the ECM of other animals continues this pattern.

Overall, the mesoglea macromolecules are likely organized as follows: mucins form short polymers that non-covalently complex with sulfated glycans via counterions from seawater to form a stable, loose gel. Actins, myosins, and collagen type IV may extend into the mesoglea from cellular layers to provide additional structural integrity. This loose composite may be colloidal-gel-like in nature. This results in a material that is more fluid-like than the jelly of cnidarian jellyfish or other

strong viscoelastic tissues in other animals, yet ideal for energy dissipation and environmental flexibility. Early divergence from other animals and exposures to extreme environments throughout history may have acted as physical drivers that pushed ctenophores to select for this unique form of extracellular material. These findings provide a new perspective on the evolution of the extracellular matrix near the origin of animal multicellularity and the emergence of diverse animal tissues.

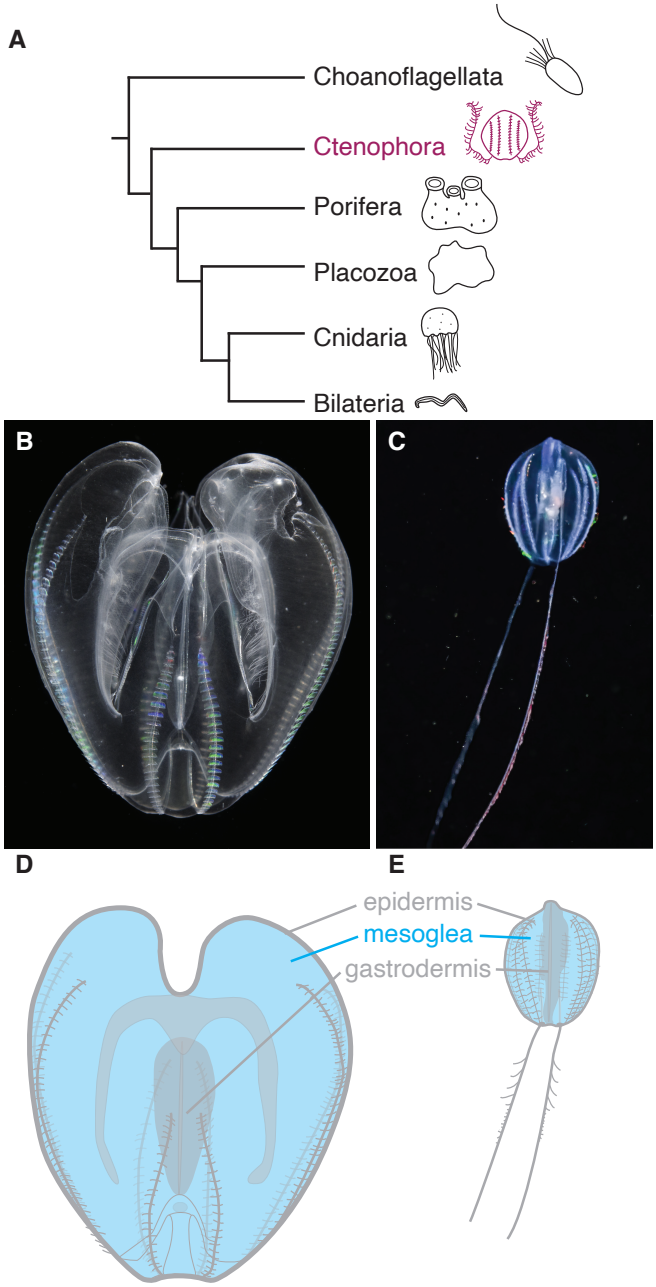

Figure 1: **Ctenophores have bulk gelatinous mesogleal tissue.** (**A**) Ctenophores are one of our most distant animal relatives. (**B**) *Mnemiopsis leidyi* is a lobate ctenophore, photo taken by SERC/MarineGEO Chesapeake BioBlitz Team. (**C**) *Pleurobrachia pileus* is a cydippid ctenophore. Image cropped from ”Sea Gooseberry (Pleurobrachia pileus) in Norway” by Vsevolod Rudyi is licensed under CC BY 4.0. (**D and E**) The simplified anatomy of *M. leidyi* and *P. pileus*, respectively, comprising gelatinous mesogleal tissue between an outer epidermal layer and an inner gastrodermal layer.

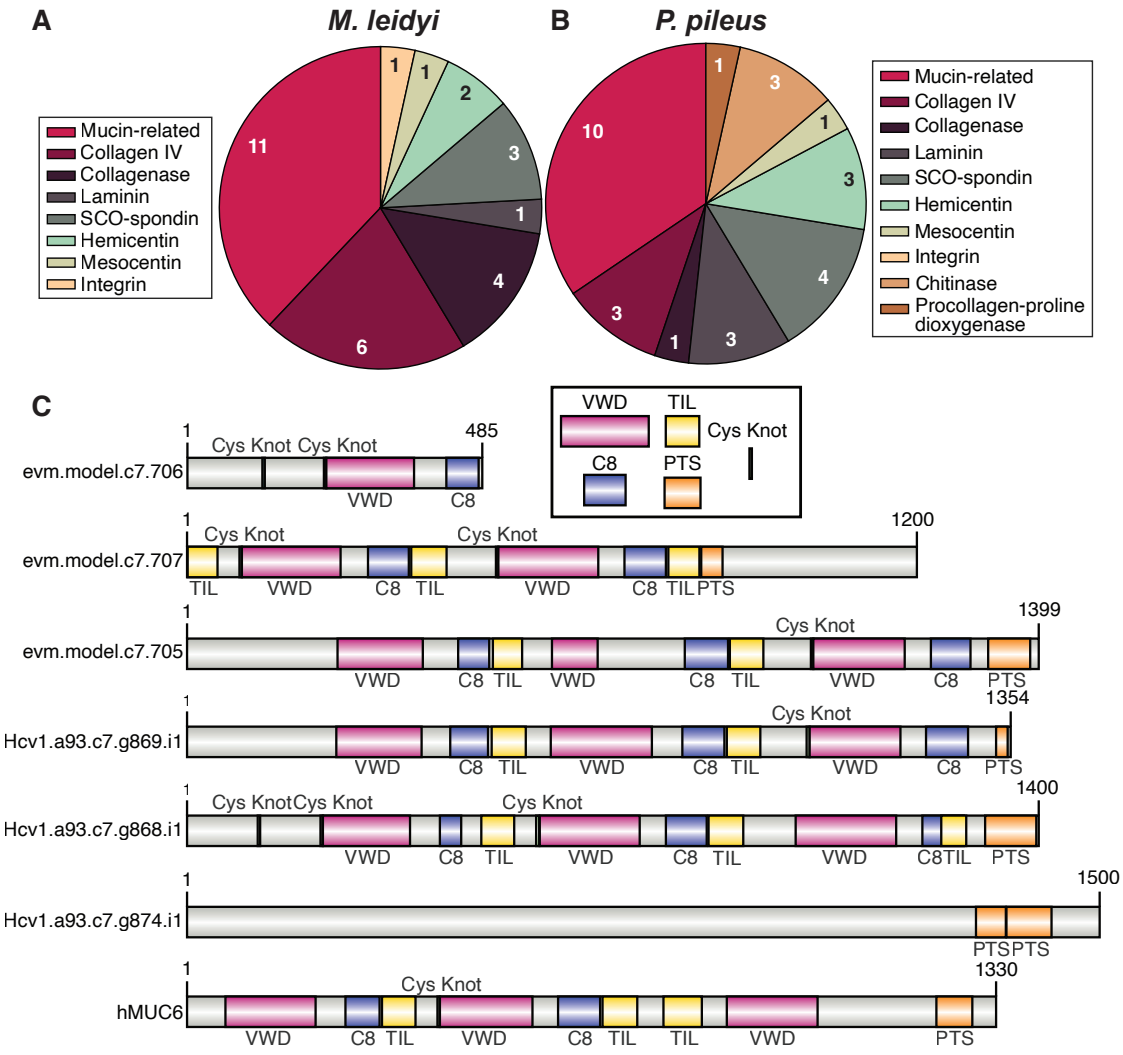

Figure 2: **Many mucus-related proteins are present in ctenophore mesoglea.** (**A**) The major extracellular proteins identified in *M. leidyi* from LC-MS/MS proteomics. (**B**) The major extracellular proteins identified in *P. pileus* from LC-MS/MS proteomics using an *H. californensis* database. (**C**) Domain architecture of *M. leidyi* and *H. californensis* proteins with homology to metazoan mucins. Proteins evm.model.c7.706 and Hcv1.av93.c7.g869.i1 were observed in the proteomics results for *M. leidyi* and *P. pileus*, respectively, and proteins either in tandem array or head-to-head orientation near evm.model.c7.706 and Hcv1.av93.c7.g869.i1 also contain mucin-related domains. hMUC6 is a human mucin protein with the characteristic organization of functional domains in mucins, namely repeats of VWD-C8-TIL repeats with interspersed Cys Knots, and a PTS domain on the C terminus. *Homo sapiens* mucin-6 isoform X3 was used (NCBI: XP_054185958.1).

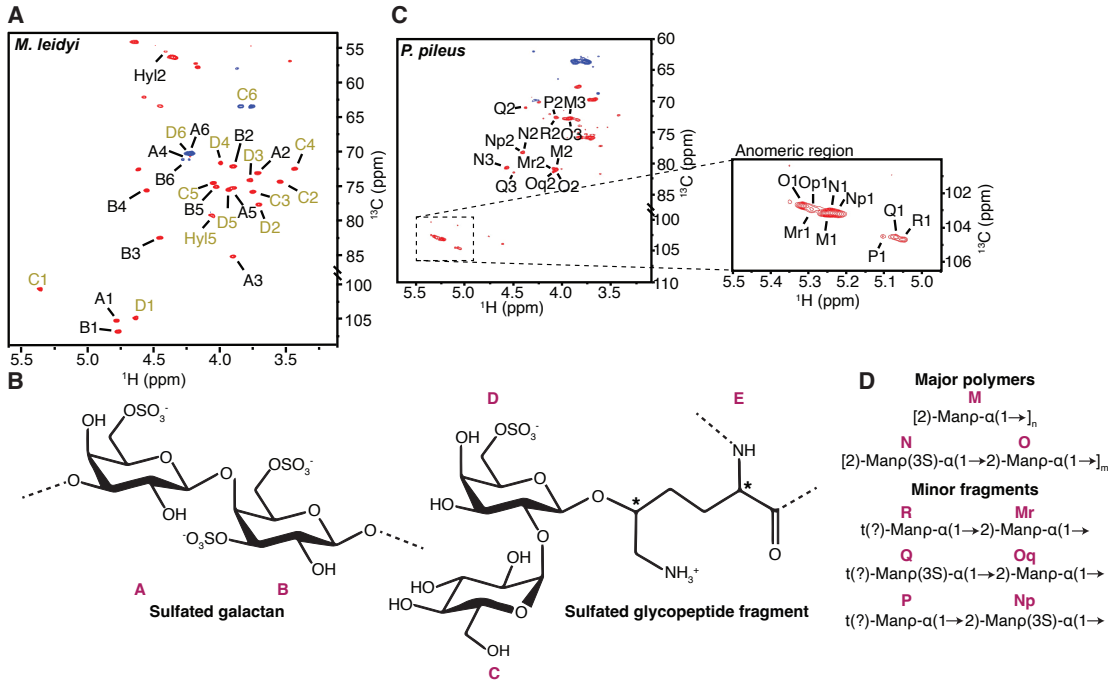

Figure 3: **Novel sulfated polysaccharides are present in ctenophore mesoglea.** (**A**) ^{1}H,^{13}C-HSQC NMR spectra of an *M. leidyi* mesoglea sample, containing signals of both a sulfated galactan and a sulfated glycopeptide fragment. Anomeric and ring residue signals of the sulfated galactan disaccharide (black) and sulfated glycoprotein fragment (gold) are indicated. Unlabeled signals arose from residual peptides. (**B**) Monomer structures of the sulfated galactan polymer (residues A and B) and the sulfated disaccharide linked to hydroxylysine (residues C, D and E) found in *M. leidyi*. (**C**) ^{1}H,^{13}C-HSQC NMR spectrum of a *P. pileus* mesoglea sample. Signals of ring C–H groups 2 and 3 are labeled in the spectrum shown on the left. On the right is an expanded anomeric region with signal assignments. (**D**) Structures of sulfated and non-sulfated mannan monomers observed in *P. pileus* with labels corresponding to the spectra.

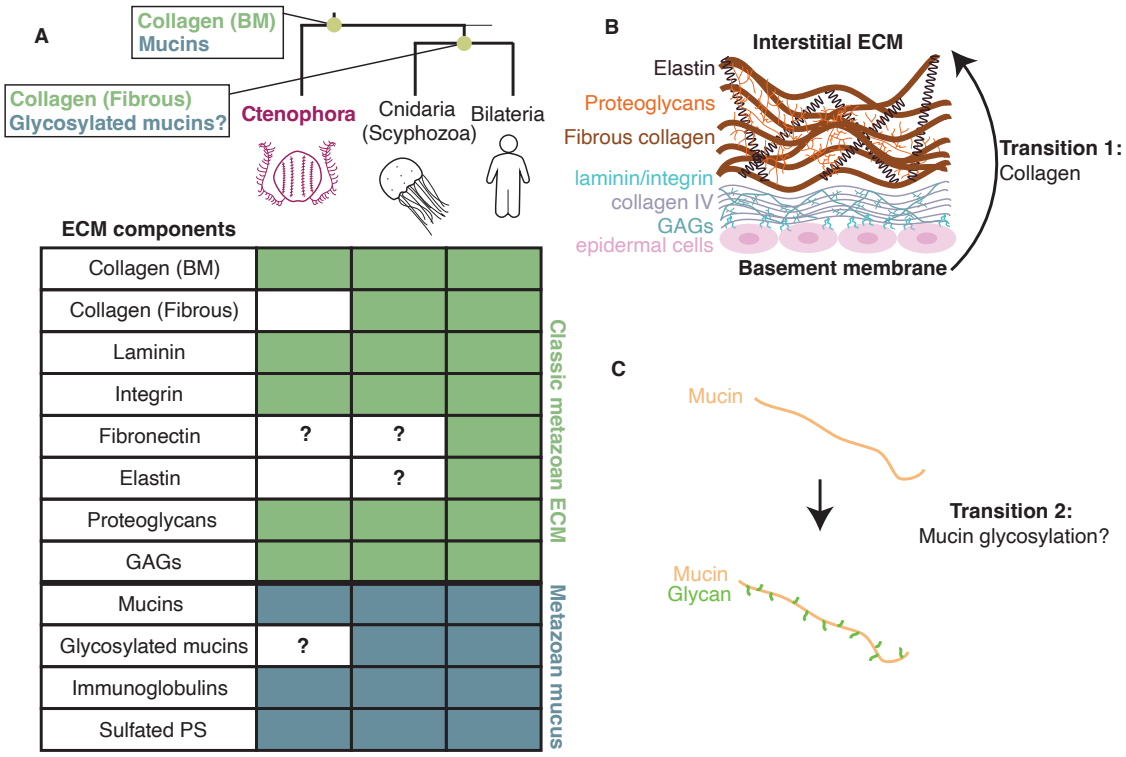

Figure 4: **Animal evolution has led to diverse extracellular materials.** (**A**) A simplified cladogram of metazoans, highlighting that the most recent common ancestor of all metazoans likely had mucin proteins and collagens capable of producing a basement membrane. After ctenophores diverged, fibrous collagens and glycosylated mucins may have developed. These components relate to two major extracellular materials in animals: the classic collagenous ECM and mucus. Bilateria contain all classic components of both, while ctenophora and cnidaria lack some components, which likely produce functional differences. (**B**) A major transition of extracellular materials in animals is the transition from the basement membrane to the interstitial ECM that arose from the development of fibrous collagens. The basement membrane is directly attached to epithelial cells via integrins and laminins, and a thin sheet is produced from a network of collagen type IV and GAGs. The interstitial ECM is a bulk network without cells produced from fibrous collagen, elastins, and proteoglycans. (**C**) A second major transition of extracellular materials is the potential development of mucin glycosylation. Ctenophores do not appear to have glycosylated mucins, so their mucin proteins (depicted as the orange polymer) are likely bare and undecorated, compared to the highly decorated mucins of other metazoans (glycans in green, mucin in orange).

References and Notes

1. D. T. Schultz, *et al.*, Ancient gene linkages support ctenophores as sister to other animals. *Nature* pp. 1–8 (2023), doi:10.1038/s41586-023-05936-6, https://www.nature.com/articles/s41586-023-05936-6.

2. A. J. Jamieson, D. J. Lindsay, H. Kitazato, Maximum depth extensions for Hydrozoa, Tunicata and Ctenophora. *Marine Biology* **170** (3), 33 (2023), doi:10.1007/s00227-023-04177-5, https://link.springer.com/10.1007/s00227-023-04177-5.

3. J.-M. Franc, *The mesoglea of ctenophores*, Ph.D. thesis, Claude Bernard Lyon 1 University (1985).

4. S. Franc, Collagen of Coelenterates, in *Biology of Invertebrate and Lower Vertebrate Collagens*, A. Bairati, R. Garrone, Eds. (Springer, Boston, MA), vol. 93 of *NATO ASI Series*, pp. 197–210 (1985), http://link.springer.com/10.1007/978-1-4684-7636-1_16.

5. G. O. Mackie, C. E. Mills, C. L. Singla, Structure and function of the prehensile tentilla of Euplokamis (Ctenophora, Cydippida). *Zoomorphology* **107** (6), 319–337 (1988), doi:10.1007/BF00312216, http://link.springer.com/10.1007/BF00312216.

6. B. Barzansky, H. Lenhoff, M., On the Chemical Composition and Developmental Role of the Mesoglea of Hydra. *American Zoologist* **14** (2), 575–581 (1974), doi:10.1093/icb/14.2.575, https://doi.org/10.1093/icb/14.2.575.

7. A. L. Fidler, *et al.*, Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. *eLife* **6**, e24176 (2017), doi:10.7554/eLife.24176, https://doi.org/10.7554/eLife.24176.

8. G. W. Draper, D. Shoemark, J. Adams, Modelling the early evolution of extracellular matrix from modern Ctenophores and Sponges. *Essays in Biochemistry* **63** (3), 389–405 (2019), doi:10.1042/EBC20180048, https://portlandpress.com/essaysbiochem/article/63/3/389/218813/Modelling-the-early-evolution-of-extracellular.

9. T. P. Norekian, L. L. Moroz, Neural system and receptor diversity in the ctenophore. *Journal of Comparative Neurology* **527** (12), 1986–2008 (2018), doi:10.1002/cne.24633, https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.24633.

10. M. J. Youngbluth, P. Kremer, T. G. Bailey, C. A. Jacoby, Chemical composition, metabolic rates and feeding behavior of the midwater ctenophore Bathocyroe fosteri. *Marine Biology* **98** (1), 87–94 (1988), doi:10.1007/BF00392662, http://link.springer.com/10.1007/BF00392662.

11. G. Schneider, Zur chemischen Zusammensetzung der Ctenophore PIeurobrachia pileus in der Kieler Bucht. *Helgolander Meeresunters* **43**, 67–76 (1989).

12. A. Clarke, L. J. Holmes, D. J. Gore, Proximate and elemental composition of gelatinous zooplankton from the Southern Ocean. *Journal of Experimental Marine Biology and Ecology* **155** (1), 55–68 (1992), doi:10.1016/0022-0981(92)90027-8, https://linkinghub.elsevier.com/retrieve/pii/0022098192900278.

13. T. Bailey, M. Youngbluth, G. Owen, Chemical composition and oxygen consumption rates of the ctenophore Bolinopsis infundibulum from the Gulf of Maine. *Journal of Plankton Research* **16** (6), 673–689 (1994), doi:10.1093/plankt/16.6.673, https://doi.org/10.1093/plankt/16.6.673.

14. T. Bailey, M. Youngbluth, G. Owen, Chemical composition and metabolic rates of gelatinous zooplankton from midwater and benthic boundary layer environments off Cape Hatteras, North Carolina, USA. *Marine Ecology Progress Series* **122**, 121–134 (1995), doi:10.3354/meps122121, http://www.int-res.com/abstracts/meps/v122/p121-134/.

15. J. Percy, F. Fife, The Biochemical Composition and Energy Content of Arctic Marine Macrozooplankton. *ARCTIC* **34** (4), 307–313 (1981), doi:10.14430/arctic2533, https://journalhosting.ucalgary.ca/index.php/arctic/article/view/65585.

16. U. Hoeger, Biochemical composition of ctenophores. *Journal of Experimental Marine Biology and Ecology* **72** (3), 251–261 (1983), doi:10.1016/0022-0981(83)90110-7, https://linkinghub.elsevier.com/retrieve/pii/0022098183901107.

17. M. Yousefian, A. Erkan Kideys, Biochemical composition of Mnemiopsis leidyi in the southern Caspian Sea. *Fish Physiology and Biochemistry* **29** (2), 127–131 (2003), doi: 10.1023/B:FISH.0000035921.70639.fb, http://link.springer.com/10.1023/B:FISH.0000035921.70639.fb.

18. B. E. Anninsky, *et al.*, Effect of starvation on the biochemical compositions and respiration rates of ctenophores Mnemiopsis leidyi and Beroe ovata in the Black Sea. *Journal of the Marine Biological Association of the United Kingdom* **85** (3), 549–561 (2005), doi:10.1017/S0025315405011471, https://www.cambridge.org/core/journals/journal-of-the-marine-biological-association-of-the-united-kingdom/article/effect-of-starvation-on-the-biochemical-compositions-and-respiration-rates04808BD83B4F72FC0D3A8AB88C51F930.

19. R. F. Lee, Lipids of Zooplankton from Bute Inlet, British Columbia. *Journal of the Fisheries Research Board of Canada* **31** (10), 1577–1582 (1974), doi:10.1139/f74-198, https://cdnsciencepub.com/doi/abs/10.1139/f74-198.

20. V. Koutsouveli, *et al.*, The Chromosome-level Genome of the Ctenophore Mnemiopsis leidyi A. Agassiz, 1865 Reveals a Unique Immune Gene Repertoire. *Genome Biology and Evolution* **17** (2), evaf006 (2025), doi:10.1093/gbe/evaf006, https://doi.org/10.1093/gbe/evaf006.

21. T. Lang, *et al.*, Searching the Evolutionary Origin of Epithelial Mucus Protein Components—Mucins and FCGBP. *Molecular Biology and Evolution* **33** (8), 1921–1936 (2016), doi:10.1093/molbev/msw066, https://doi.org/10.1093/molbev/msw066.

22. D. Fass, D. J. Thornton, Mucin networks: Dynamic structural assemblies controlling mucus function. *Current Opinion in Structural Biology* **79**, 102524 (2023), doi:10.1016/j.sbi.2022.102524, https://linkinghub.elsevier.com/retrieve/pii/S0959440X22002032.

23. L. Khmelnitsky, A. Milo, O. Dym, D. Fass, Diversity of CysD domains in gel-forming mucins. *The FEBS Journal* **290** (21), 5196–5203 (2023), _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1111/febs.16918, doi:10.1111/febs.16918, https://onlinelibrary.wiley.com/doi/abs/10.1111/febs.16918.

24. B. Demouveaux, *et al.*, Mucin CYS domain stiffens the mucus gel hindering bacteria and spermatozoa | Scientific Reports. *Scientific Reports* **9** (16993) (2019), doi:https://doi.org/10.1038/s41598-019-53547-x, https://www.nature.com/articles/s41598-019-53547-x.

25. J.-L. Desseyn, Mucin CYS domains are ancient and highly conserved modules that evolved in concert. *Molecular Phylogenetics and Evolution* **52** (2), 284–292 (2009), doi:10.1016/j.ympev.2009.03.035, https://www.sciencedirect.com/science/article/pii/S1055790309001237.

26. R. Bansil, B. S. Turner, Mucin structure, aggregation, physiological functions and biomedical applications. *Current Opinion in Colloid & Interface Science* **11** (2), 164–170 (2006), doi:10.1016/j.cocis.2005.11.001, https://www.sciencedirect.com/science/article/pii/S1359029405001184.

27. J. Jurado, *et al.*, Proteomic profile of the skin mucus of farmed gilthead seabream (*Sparus aurata*). *Journal of Proteomics* **120**, 21–34 (2015), doi:10.1016/j.jprot.2015.02.019, https://www.sciencedirect.com/science/article/pii/S1874391915000834.

28. I. Sanahuja, A. Ibarz, Skin mucus proteome of gilthead sea bream: A non-invasive method to screen for welfare indicators. *Fish & Shellfish Immunology* **46** (2), 426–435 (2015), doi:10.1016/j.fsi.2015.05.056, https://www.sciencedirect.com/science/article/pii/S1050464815002478.

29. P. H. Jensen, D. Kolarich, N. H. Packer, Mucin-type O-glycosylation – putting the pieces together. *The FEBS Journal* **277**, 81–94 (2009), https://febs.onlinelibrary.wiley.com/doi/10.1111/j.1742-4658.2009.07429.x.

30. T. Lang, G. C. Hansson, T. Samuelsson, Gel-forming mucins appeared early in metazoan evolution. *Proceedings of the National Academy of Sciences* **104** (41), 16209–16214 (2007), doi:10.1073/pnas.0705984104, https://www.pnas.org/doi/abs/10.1073/pnas.0705984104.

31. A. McShane, *et al.*, Mucus. *Current Biology* **31** (15), R938–R945 (2021), doi:10.1016/j.cub.2021.06.093, https://www.cell.com/current-biology/abstract/S0960-9822(21)00922-2.

32. H.-C. Flemming, J. Wingender, The biofilm matrix. *Nature Reviews Microbiology* **8** (9), 623–633 (2010), doi:10.1038/nrmicro2415, https://www.nature.com/articles/nrmicro2415.

33. H. Abo, H. Kawashima, Significance of Sulfated Glycans on Mucins in the Gut. *Trends in Glycoscience and Glycotechnology* **35** (207), E73–77 (2023).

34. G. Raba, A. Luis, Mucin utilization by gut microbiota: recent advances on characterization of key enzymes. *Essays in Biochemistry* **67** (3), 345–353 (2023), doi:10.1042/EBC20220121, https://doi.org/10.1042/EBC20220121.

35. Y. Ishikawa, R. Lennon, F. Forneris, J. Myllyharju, A. M. Salo, Collagen IV biosynthesis: Intracellular choreography of post-translational modifications. *Matrix Biology* **140**, 59–77 (2025), doi:10.1016/j.matbio.2025.07.002, https://www.sciencedirect.com/science/article/pii/S0945053X25000617.

36. N. M. Khong, *et al.*, Nutritional composition and total collagen content of three commercially important edible jellyfish. *Food Chemistry* **196**, 953–960 (2016), doi:10.1016/j.foodchem.2015.09.094, https://linkinghub.elsevier.com/retrieve/pii/S0308814615014375.

37. J. Toppe, S. Albrektsen, B. Hope, A. Aksnes, Chemical composition, mineral content and amino acid and lipid profiles in bones from various fish species. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **146** (3), 395–401 (2007), doi:10.1016/j.cbpb.2006.11.020, https://www.sciencedirect.com/science/article/pii/S1096495906004210.

38. M. Gauza-Wodarczyk, L. Kubisz, D. Wodarczyk, Amino acid composition in determination of collagen origin and assessment of physical factors effects. *International Journal of Biological Macromolecules* **104**, 987–991 (2017), doi:10.1016/j.ijbiomac.2017.07.013, https://linkinghub.elsevier.com/retrieve/pii/S0141813017315787.

39. R. B. Proenca, *Evolution and developmental functions of mesocentin: A novel Caenorhabditis elegans extracellular matrix protein*, Ph.D., The Johns Hopkins University,

United States – Maryland (2002), https://www.proquest.com/docview/305338793/abstract/A9221A7942B747D5PQ/1.

40. X. Zhang, *et al.*, Structural and Functional Components of the Skate Sensory Organ Ampullae of Lorenzini. *ACS Chemical Biology* **13** (6), 1677–1685 (2018), doi:10.1021/acschembio.8b00335, https://doi.org/10.1021/acschembio.8b00335.

41. M. J. M. Wells, H. A. Stretz, Supramolecular architectures of natural organic matter. *Science of The Total Environment* **671**, 1125–1133 (2019), doi:10.1016/j.scitotenv.2019.03.406, https://www.sciencedirect.com/science/article/pii/S0048969719314044.

42. A. R. Cerullo, *et al.*, Comparative Animal Mucomics: Inspiration for Functional Materials from Ubiquitous and Understudied Biopolymers. *ACS Biomaterials Science & Engineering* **6** (10), 5377–5398 (2020), doi:10.1021/acsbiomaterials.0c00713, https://doi.org/10.1021/acsbiomaterials.0c00713.

43. C. W. Dunn, S. P. Leys, S. H. Haddock, The hidden biology of sponges and ctenophores. *Trends in Ecology & Evolution* **30** (5), 282–291 (2015), doi:10.1016/j.tree.2015.03.003, https://linkinghub.elsevier.com/retrieve/pii/S0169534715000622.

44. C. W. Dunn, *et al.*, Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452** (7188), 745–749 (2008), doi:10.1038/nature06614.

45. P. Burkhardt, *et al.*, Syncytial nerve net in a ctenophore adds insights on the evolution of nervous systems. *Science* **380** (6642), 293–297 (2023), doi:10.1126/science.ade5645, https://www.science.org/doi/10.1126/science.ade5645.

46. I. M. Black, *et al.*, Acetylation in Ionic Liquids Dramatically Increases Yield in the Glycosyl Composition and Linkage Analysis of Insoluble and Acidic Polysaccharides. *Analytical Chemistry* **95** (34), 12851–12858 (2023), doi:10.1021/acs.analchem.3c02056, https://doi.org/10.1021/acs.analchem.3c02056.

47. W. F. Vranken, *et al.*, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59** (4), 687–696 (2005), doi:10.1002/prot.20449.

48. S. P. Skinner, *et al.*, CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *Journal of biomolecular NMR* **66** (2), 111–124 (2016), doi:10.1007/s10858-016-0060-y.

49. T. Masuko, *et al.*, Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Analytical Biochemistry* **339** (1), 69–72 (2005), doi:10.1016/j.ab.2004.12.001.

50. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *Journal of Molecular Biology* **215** (3), 403–410 (1990), doi:10.1016/S0022-2836(05)80360-2.

51. P. Jones, *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30** (9), 1236–1240 (2014), doi:10.1093/bioinformatics/btu031, https://doi.org/10.1093/bioinformatics/btu031.

52. J. Abramson, *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630** (8016), 493–500 (2024), doi:10.1038/s41586-024-07487-w, https://www.nature.com/articles/s41586-024-07487-w.

53. M. Van Kempen, *et al.*, Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* **42** (2), 243–246 (2024), doi:10.1038/s41587-023-01773-0, https://www.nature.com/articles/s41587-023-01773-0.

54. T. Lang, M. Alexandersson, G. C. Hansson, T. Samuelsson, Bioinformatic identification of polymerizing and transmembrane mucins in the puffer fish Fugu rubripes. *Glycobiology* **14** (6), 521–527 (2004), doi:10.1093/glycob/cwh066.

55. C. J. A. Sigrist, *et al.*, New and continuing developments at PROSITE. *Nucleic Acids Research* **41** (D1), D344–D347 (2013), doi:10.1093/nar/gks1067, https://doi.org/10.1093/nar/gks1067.

56. U. A. Vitt, S. Y. Hsu, A. J. W. Hsueh, Evolution and Classification of Cystine Knot-Containing Hormones and Related Extracellular Signaling Molecules. *Molecular Endocrinology* **15** (5) (2001).

57. Y. Xie, *et al.*, IBS 2.0: an upgraded illustrator for the visualization of biological sequences. *Nucleic Acids Research* **50** (W1), W420–W426 (2022), doi:10.1093/nar/gkac373, https://doi.org/10.1093/nar/gkac373.

Acknowledgments

We thank the Yale Center for Research Computing for the computational resources on the Grace cluster for BLASTp analyses and AltaBioscience who performed the total amino acid analysis. We also thank Lauren Mellenthin, Mary Beth Decker, and other members of the Dunn Lab for assistance with ctenophore care and collection. We are also grateful to members of the Sweeney Lab, Aurelia Moriyama-Gurish, Shuhan Zhang, and Yale Peabody Museum of Natural History staff Lourdes Rojas and Eric Lazo-Wasem for additional help with ctenophore collection.

Funding: This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, under award #ED-SC0015662. Glycoproteomics analysis was performed at the Complex Carbohydrate Research Center and was supported in part by the National Institutes of Health-funded grant R24GM137782 to Parastoo Azadi. The Eclipse mass spectrometer used in the glycopeptide analysis was supported by GlycoMIP, a National Science Foundation Materials Innovation Platform funded through Cooperative Agreement DMR-1933525.

Author contributions: Conceptualization: S.S., C.D., A.S. Data curation: S.S., S.A.H., I.B., I.K., J.V., C.H. Methodology: S.S., S.A.H., I.B., I.K., J.V., C.H., P.A., C.D., A.S. Funding acquisition: P.A., A.S. Supervision: A.S. Visualization: S.S., I.B., J.V., C.D., A.S. Writing - original draft: S.S., A.S. Writing - review and editing: S.S., J.V., C.H., C.D., A.S.

Competing interests: There are no competing interests to declare.

Data and materials availability: Raw data for all experiments will be available upon journal submission.

Supplementary materials

Materials and Methods

Figures S1-S3

Tables S1-S4

Data S1-S4

Reference *(57)*

Supplementary Materials for

Ctenophore mesoglea: building a mucus-like body

Sara Siwiecki, Stephanie Archer-Hartmann, Ian Black, Jiri Vlach,

Christian Heiss, Parastoo Azadi, Casey Dunn, Alison Sweeney*

*Corresponding author. Email: alison.sweeneye@yale.edu

This PDF file includes:

Materials and Methods

Figures S1-S3

Tables S1-S4

Captions for Data S1-S4

Reference *(57)*

Other Supplementary Materials for this manuscript:

Data S1-S4

Materials and Methods

0.0.1 Ctenophore Collection

Two species of ctenophores, *Mnemiopsis leidyi* and *Pleurobrachia pileus*, were collected via dip cups at the Connecticut shoreline of the Long Island Sound between 2020 to 2024. Ctenophores were collected under the State of Connecticut Department of Energy & Environmental Protection permits #0326011 and #0323011b. After collection, individuals were either maintained in a pseudokreisel at the salinity and temperature at which they were collected or immediately frozen at -80 °C.

0.0.2 Mesoglea Extraction

Whole individuals were vortexed for 1 minute to shear apart epithelial layers, then the whole organism was centrifuged at 8000 *g* for 15 minutes to pellet cells, nuclei, and other cellular debris. The resulting supernatant was a transparent, slightly viscous liquid. This supernatant sample was defined to be mesoglea for purposes of these studies. This sample was decanted from the resulting pellet and used for all experiments described below. Typically, a new mesoglea sample was prepared for each experiment. A few experiments used mesoglea samples from the same individual, and these are indicated below.

0.0.3 Total Amino Acid Analysis

Two biological replicates of *M. leidyi* and *P. pileus* mesoglea were sent to a contract service to determine their total amino acid composition (AltaBiosciences, Redditch, UK). Upon arrival, the mesoglea samples were degassed, sealed under vacuum, and hydrolyzed for 24 hours in 6 M HCl. After vacuum hydrolysis, the samples were dried to remove the concentrated HCl and diluted to the required volume using 0.1 M HCl buffer. The hydrolyzed samples were analyzed via ion-exchange chromatography with post-column ninhydrin detection for quantification of each amino acid (Biochrom 30+, Harvard Biosciences, Holliston MA). Hydroxyproline and hydroxylysine were separately identified using absorbance at 440 nm and 570 nm, respectively.

0.0.4 Glycosyl and glycopeptide analyses

Mesoglea samples (*M. leidyi*: n=2; *P. pileus*: n=1) were shipped to the Complex Carbohydrate Research Center at the University of Georgia for comprehensive glycosyl analyses. The mesoglea samples were preserved for two downstream experiments: (1) samples for glycoproteomics were treated with 0.02% sodium azide (ThermoFisher), 0.1 mM phenylmethylsulfonyl fluoride (PMSF, ThermoFisher), and 1X Halt Protease Inhibitor (ThermoFisher); (2) samples for all other glycosyl analyses were treated with 0.02% sodium azide and heated to 100 °C for 5 minutes.

Aliquots of these mesoglea samples were digested with proteases to remove proteins for carbohydrate composition experiments. First, samples underwent a double protease digestion with proteinase K (Sigma no.P4302-5G, 200 µg/10 mg sample) in 50 mM MgCl buffer (pH 7.5) overnight at 50 °C. This was followed by Pronase digestion (Sigma no.10165921001, 100 µg/10 mg sample) overnight at 50 °C. A 6 kDa MWCO dialysis membrane (Spectra Por) was used to remove digested proteins and any remaining salt from the sample. Samples were then dried using a centrifugal vacuum evaporator. These samples were used for GC-MS and nuclear magnetic resonance (NMR) experiments, described below.

GC-MS was used to perform glycosyl composition analysis, following the method outlined by Black and colleagues to produce per-O-trimethylsilyl (TMS) derivatives of the monosaccharide methyl glycosides from the digested samples (*46*). Briefly, samples are heated in 400 µl of 1 M methanolic HCl overnight at 80 °C and then dried with nitrogen. The samples are then re-N-acetylated by adding methanol:acetic anhydride:pyridine 2:1:1 for 20 minutes at room temperature and dried again. The sample was then derivatized with Tri-Sil® (ThermoFisher) at 80 °C for 30 minutes. The TMS methyl glycosides were then analyzed and visualized on the GC-MS using an Agilent 7890A GC interfaced to a 5975C MSD with a Supelco Equity-1 fused silica capillary column (30 m X 0.25 mm ID).

Dried mesoglea samples were exchanged twice from D_2O (99.9% D, Sigma), then dried again using a centrifugal vacuum evaporator. The samples were then suspended in 520 µl of D_2O with 50 nmol DSS-d_6 (99.9% D, CIL), vortexed, briefly sonicated, and warmed up (±60 °C), producing a slightly viscous gelatinous suspension. For *M. leidyi* samples, centrifugation was used to remove the insoluble material, and the volume of the solution was adjusted to ±540 µl with D_2O and

transferred to a 5 mm NMR tube. For the *P. pileus* sample, the sample would not fully pellet at 16,000 *g* and was opaque. To prepare this sample for an 800 MHz spectrometer, the sample was freeze-dried and resuspended in 200 µl D_2O (99.96% D, Sigma) and then transferred to a 1.7 mm NMR tube.

NMR data were collected using a Bruker Avance III at 70 °C with a 5 mm TCl cryoprobe (^{1}H, 600.13 MHz), or a Bruker NEO spectrometer with a 1.7 mm TCl cryoprobe (^{1}H, 799.17 MHz) where indicated. We used standard pulse programs in the spectrometer library to collect 1D ^{1}H and 2D ^{1}H, ^{1}H-COSY, ^{1}H, ^{1}H-TOCSY, ^{1}H, ^{1}H-NOESY, ^{1}H, ^{1}H-ROESY, ^{1}H, ^{13}C-HSQC, ^{1}H, ^{13}C-HMBC, and ^{1}H, ^{13}C-HSQC-TOCSY (*M. leidyi* only*)* experiments. A total recycling delay of 60 seconds was used while collecting quantitative ^{1}H NMR data. We used a mixing time of 60 to 120 ms for NOESY and 200 ms for ROESY experiments. The spin-lock time was 60 or 120 ms for TOCSY and 70 ms for HSQC-TOCSY. NMR data were processed in Topspin 4.0.5 (Bruker) and analyzed in CCPN Analysis 2.4 (*47*) or 3.2.0 (*48*). We rendered figures in MestReNova 14.2.3 (Mesterelab Research). We used ^{1}H and ^{13}C chemical shifts as references to the respective DSS signals at 0.00 ppm.

An aliquot of one of the digested *M. leidyi* mesoglea samples underwent anion exchange chromatography to further investigate the presence of potentially sulfated polysaccharides. A 5 ml DEAE HiTrap column (GE) was used to perform high-performance liquid chromatography analysis (HPLC). 250 µl of a 5 mg/ml sample (dissolved in water) was injected first and run for four minutes with 100% water. Then, the buffer was switched to a 2 M NH_4OAc solution over 31 minutes and held at 100% 2 M NH_4OAc for five minutes. 100% water was then re-injected into the column until equilibration for the last minute of the run. Sample fractions were collected and dried at one-minute intervals and reconstituted in 200 µl of DI water. Sulfated galactans were not observed after fractionation because they only eluted with high NaCl concentration.

Verification of carbohydrate presence in the fractions was detected using a phenol sulfuric acid assay. We used a modified version of the microscale phenol sulfuric acid assay of Masuko and colleagues (*49*). In short, 0.2 ml of concentrated sulfuric acid was added to wells containing 10 µl of resuspended sample fractions. 5% phenol was added to the wells, and plates were incubated at 50 °C for 20 minutes. Absorbance was read at 490 nm using a Molecular Devices Spectra Max plate reader. Fraction four was then run on the NMR according to the method above.

0.0.5 Glycoproteomics

All of the following reagents were purchased from Sigma Aldrich unless otherwise mentioned. Mass spectrometric data was performed on a Thermo Scientific LTQ Orbitrap Eclipse Tribid mass spectrometer attached with an Ultimate RSLCnano-LC system. Data analysis was performed using Xcalibur 3.0 and Byonic v4.10 software.

Proteomics of both tryptic digests from one sample of *M. leidyi* and one *P. pileus* sample, as well as their pre-digested samples, were performed using LC-MS/MS (Liquid Chromatography-MS/MS). The provided samples (pre-digested or not) were reduced, alkylated with iodoacetamide, and the protein was digested with (1:10) sequencing-grade trypsin (Promega) at 37 °C for 16 hours. The peptides were analyzed on an Orbitrap Eclipse or Orbitrap Ascend mass spectrometer equipped with a nanospray ion source. Pre-packed Nano-LC columns of 15 cm length with 75 µm internal diameter, filled with 3 µm C18 reverse-phase material, were used for chromatographic separation of the samples. The separation conditions were low to high acetonitrile in a solution containing 0.1% formic acid, and the separation time was 60 minutes. The precursor ion scan was acquired at 120,000 resolution in an Orbitrap analyzer, and precursors at a time frame of 3 seconds were selected for subsequent fragmentation using higher-energy collisional dissociation (HCD). The threshold for triggering an MS/MS event on the ion-trap was set to 500 counts. Charge state screening was enabled, and precursors with an unknown charge state or a charge state of +1 were excluded. Dynamic exclusion was enabled (exclusion duration of 30 s). The fragment ions were analyzed on an Orbitrap of HCD at 30,000 resolution.

The resulting data were either hand-processed or processed with Byonic and searched against protein models for each species (described below) and a catalog of basic N-glycans and O-glycans. Post-translational modifications, including oxidation and deamidation, were included in the variable search. Carboxyamidomethylation of the cysteine was included as a fixed modification. The precursor mass tolerance and fragment mass tolerances were set to 10 ppm and 20 ppm, respectively.

For *M. leidyi*, we used protein models from the recently published chromosome-level genome for the northern Atlantic *M. leidyi* population, accessed on GenBank under GCA_963919725.1 (*20*). For *P. pileus*, there are no genome or assembled protein models available so we used protein models from the closely related *Hormiphora californensis* chromosome-level genome, generated in (*1*) and

accessible on GenBank under GCA_020137815.1. We provided the functional annotation of each identified protein and glycoprotein through sequence and structural searches. For sequence searches, BLASTp, Pfam, and Gene Ontology terms were queried for both species. For both species, we used BLASTp with the NCBI nr database (April 2025 for *M. leidyi,* June 2025 for *H. californensis*) with an e-value cut-off of 1E-5 (*50*). For further functional annotation, we used InterProScan via its webpage (`https://www.ebi.ac.uk/interpro/search/sequence/`) to identify protein family domains and GO terms (*51*). We used the default parameters with all databases available. The MOTIF Search tool on `https://www.genome.jp/` was also used for domain architecture visualization.

Several proteins for both species did not contain any results via the sequence searches or only had identity to hypothetical proteins with no functional annotations, so structural searches were also performed for those particular proteins. Structural searches were also performed for proteins of particular interest, such as collagens and mucins, described further in this paper. We used AlphaFold3 to predict structures via the online server, `https://alphafoldserver.com/` (*52*). We searched for any homology and thus used structure predictions regardless of pLDDT accuracy to maximize the likelihood of finding any small structural similarities. We then used the top-ranked prediction for each protein and searched for structural similarity to any other proteins in the PDB or AlphaFold DB via FoldSeek at `https://search.foldseek.com/search` (*53*). The e-values of all hits with structural homology are reported in our supplementary data (data S3 to S4). Any given protein's particular structural details would need to be further investigated with confirmed high-confidence structure predictions to show true structural homology.

For proteins with potential mucin homology, we used PTSPred, a Perl-based PTS/mucin domain search tool developed by (*54*), to identify PTS domains. For some mucin-like proteins, we also performed visual inspections to identify PTS-rich sequnces that were too short to be detected with the search tool. Cysteine knots were also identified using the motif search on ScanProsite with the consensus motif C-x-G-x-C (*55, 56*). Illustrator for Biological Sequences 2.0 was used for protein domain architecture visualization (*57*).

Supplementary Data

Table S1: **^{1}H and ^{13}C NMR chemical shifts of galactan residues found in *M. leidyi*.**

ID	Residue	Chemical shift (ppm)						Links
		1	2	3(a,b)	4	5	6(a,b)	
A	→ 3)-Gal(6S)-β(1→	4.78	3.71	*3.90*	4.25	3.90	**4.23**	A1→B4
		105.4	73.5	*85.4*	71.6	75.6	**70.6**	
B	→ 4)-Gal(3S,6S)-β(1→	4.77	3.90	**4.45**	*4.56*	4.03	**4.26**	B1→A3
		107.0	72.5	**82.8**	*76.0*	75.5	**71.5**	
C	t-Glcρ-α(1→	5.36	3.55	3.75	3.44	4.05	3.84, 3.76	C1→D2
		100.5	74.4	75.8	72.5	74.5	63.5	
D	→ 2)-Gal(6S)-β(1→	4.65	*3.71*	3.78	4.00	3.95	**4.23**	D1→Hyl5
		104.6	*77.6*	74.0	71.7	75.6	**70.4**	
E	Hyl		4.36	2.02, 1.84	1.75	*4.12*	3.25, 3.07	
		176.2	56.5	29.1	31.2	*78.3*	45.5	

Italicized values indicate glycosyl linkage positions; **bold values** indicate sulfation.

Table S2: **GC-MS (Fig. S2) highlights galactose in ctenophore mesoglea.**

Glycosyl residue	*M. leidyi*		*P. pileus*	
	Mass (µg)	Mol %	Mass (µg)	Mol %
Ribose (Rib)	n.d.	–	0.3	0.8
Mannose (Man)	2.4	7.1	12.2	26.3
Galactose (Gal)	26.7	79.2	8.3	17.9
Glucose (Glc)	4.6	13.7	3.8	8.1
N-Acetyl Glucosamine (GlcNAc)	n.d.	–	0.9	2.4
N-Acetyl Galactosamine (GalNAc)	n.d.	–	20.9	44.4
Total	**33.7**		**46.4**	

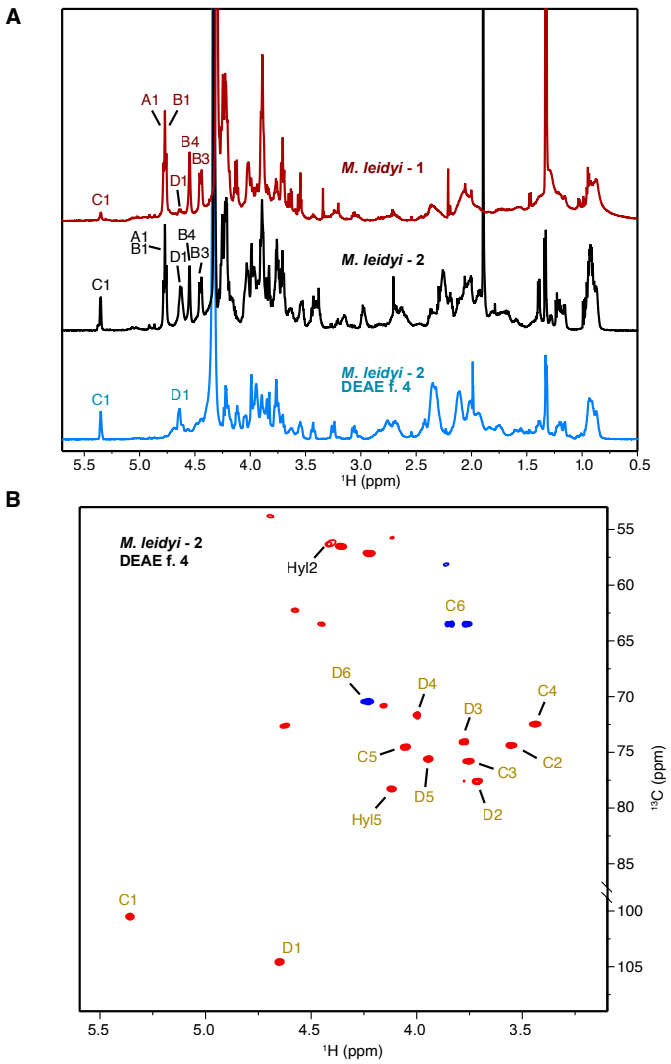

Figure S1: **Unique sulfated galactans are present in *M. leidyi* mesoglea.** (**A**) ^{1}H NMR spectra of two *M. leidyi* mesoglea samples. The bottom sample is DEAE fraction 4 of the previous *M. leidyi* sample. (**B**) ^{1}H,^{13}C-HSQC NMR spectra of the DEAE fraction 4 of the second *M. leidyi* mesoglea sample, where only signals of sulfated glycopeptide fragment were present. The sulfated glycan fragment (gold) and the amino acid fragment (black) are indicated. Unlabeled signals arose from residual peptides.

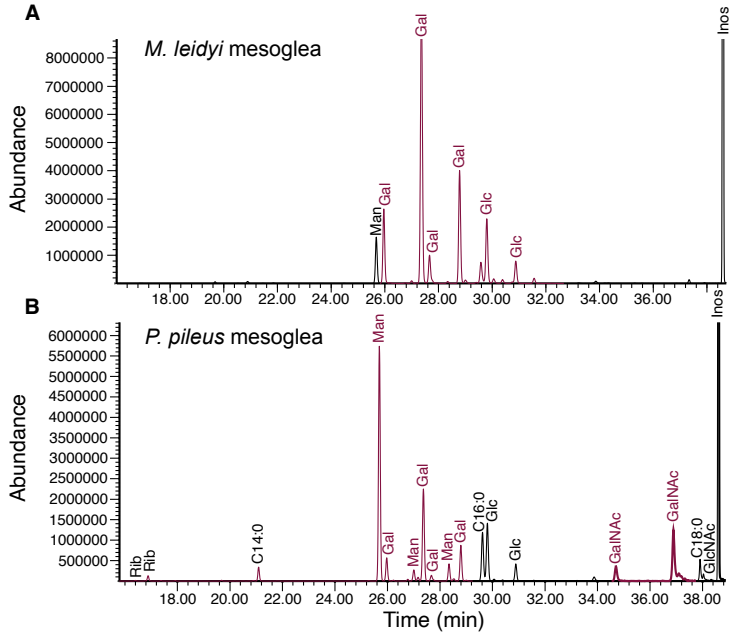

Figure S2: **Galactose-based monosaccharides are abundant in ctenophore mesoglea.** (**A**) Total ion chromatograms from GC-MS of *M. leidyi*. (**B**) Total ion chromatograms from GC-MS of *P. pileus*. Most abundant peaks are highlighted in red. Chromatograms were taken after derivatization of the sample that underwent protease-treatment and dialysis (n=1 for both species).

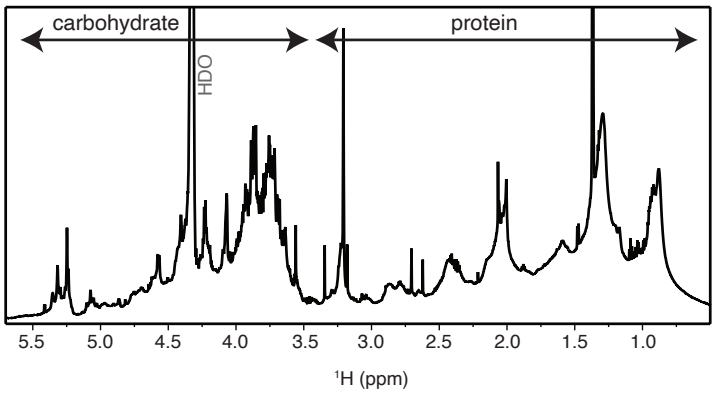

Figure S3: **Sulfated mannan structures found in *P. pileus*.** ^{1}H NMR spectra of one *P. pileus* mesoglea sample.

Table S3: ^{1}H and ^{13}C NMR chemical shifts of sulfated mannan residues found in *P. pileus*.

ID	Residue	Chemical shift (ppm)						Links
		1	2	3	4	5	6(a,b)	
M	→ 2)-Manρ-α(1→	5.25	*4.07*	3.92	3.68	3.72	3.87, 3.75	M1→M2
		103.1	*81.0*	72.8	69.8	76.0	63.8	
N	→ 2)-Manρ(3S)-α(1→	5.24	*4.41*	**4.57**	3.85	3.82	3.89, 3.78	N1→O2
		103.1	*78.3*	**80.8**	67.7	75.9	63.6	
O	→ 2)-Manρ-α(1→	5.32	*4.08*	3.94	3.72	3.71	3.86, 3.74	O1→N2
		102.7	*81.2*	72.9	69.8	76.1	63.8	
Mr	→ 2)-Manρ-α(1→	5.29	*4.09*	/	/	/	/	
		102.9	*80.9*	/	/	/	/	
Np	→ 2)-Manρ(3S)-α(1→	5.22	*4.43*	/	/	/	/	
		103.2	*78.4*	/	/	/	/	
Oq	→ 2)-Manρ-α(1→	5.29	*4.10*	/	/	/	/	
		102.9	*81.5*	/	/	/	/	
P	t(?)-Manρ-α(1→	5.10	4.07	/	/	/	75.9	P1→Np2
		104.5	72.6	~73.2	/	75.9	/	
Q	t(?)-Manρ(3S)-α(1→	5.07	4.38	**4.50**	3.83	/	/	Q1→Oq2
		104.5	71.1	**81.6**	67.6	75.9	/	
R	t(?)-Manρ-α(1→	5.05	4.06	/	/	/	/	R1→Mr2
		104.7	72.7	73.2	/	75.9	/	

/ = not determined. *Italicized values* indicate glycosyl linkage positions; **bold values** indicate sulfation.

Table S4: **Ctenophore mesoglea is rich in negatively charged amino acids.**

Amino Acid	***M. leidyi*** **(µg/ml)**	**% total**	***P. pileus*** **(µg/ml)**	**% total**
Glutamic acid	86 ± 18	18.9 ± 0.3	333 ± 63	14.2 ± 0.0
Aspartic acid	55.8 ± 10.7	12.4 ± 0.0	234 ± 28	10.1 ± 0.7
Arginine	24.5 ± 4.4	5.44 ± 0.07	217 ± 55	9.12 ± 0.59
Lysine	19.8 ± 3.3	4.42 ± 0.11	173 ± 30	7.38 ± 0.14
Leucine	19.2 ± 3.6	4.26 ± 0.01	166 ± 41	6.98 ± 0.40
Valine	24.1 ± 4.3	5.35 ± 0.07	150 ± 35	6.34 ± 0.29
Glycine	45.8 ± 10.2	10.1 ± 0.3	131 ± 12	5.69 ± 0.57
Threonine	24.4 ± 4.7	5.41 ± 0.02	126 ± 31	5.33 ± 0.29
Isoleucine	15.3 ± 3.1	3.38 ± 0.04	114 ± 26	4.83 ± 0.19
Phenylalanine	13.7 ± 2.6	3.03 ± 0.01	112 ± 24	4.75 ± 0.12
Proline	31.2 ± 4.4	6.99 ± 0.35	110 ± 12	4.78 ± 0.41
Alanine	22.5 ± 4.8	4.97 ± 0.12	102 ± 11	4.40 ± 0.35
Tyrosine	7.83 ± 0.45	1.82 ± 0.44	101 ± 27	4.25 ± 0.33
Serine	16.2 ± 2.5	3.62 ± 0.13	95 ± 17	4.05 ± 0.03
Methionine	7.31 ± 1.30	1.63 ± 0.02	58.1 ± 11.8	2.47 ± 0.03
Histidine	6.43 ± 1.12	1.43 ± 0.02	47.9 ± 8.5	2.05 ± 0.03
Cystine	9.7 ± 1.4	2.16 ± 0.09	47.8 ± 12.7	2.01 ± 0.16
Hydroxyproline	9.1 ± 3.3	1.95 ± 0.36	13.0 ± 2.9	0.549 ± 0.017
Hydroxylysine*	12.6 ± 3.0	2.78 ± 0.13	11.9 ± 0.4	0.527 ± 0.115
Totals	**451 ± 86**	**100%**	**2345 ± 445**	**100%**

*Percentage recovery not fully validated. n=2 per species; ±SE.

Caption for Data S1. BLASTp annotations of *M. leidyi* proteomics results highlight many structural and ECM proteins. Each MS proteomics hit that had identity with a known ECM or load-bearing protein is listed. For each protein, annotations of three major BLAST hit are included with accession numbers, e-values, and identity percentages. For the protein, the log probability from the LC-MS/MS results is also included.

Caption for Data S2. FoldSeek annotations of *M. leidyi* proteomics results highlight additional structural and ECM proteins. For the proteins that did not result in any functional information from BLAST searches, AlphaFold and FoldSeek were used to search for structural homology to any known proteins. Through this method, we identified a couple of additional proteins that may play roles in the ECM. For each protein, the top FoldSeek hit is listed with its accession number, e-value, identity percentage, the FoldSeek database that was used, and the log probability from the proteomics results.

Caption for Data S3. BLASTp annotations of *P. pileus* proteomics results highlight many structural and ECM proteins. The results for the *P. pileus* sample are listed in the same format as Data S1, using predicted proteins from the *H.californensis* genome.

Caption for Data S4. FoldSeek annotations of *P. pileus* proteomics results highlight additional structural and ECM proteins. The results for the *P. pileus* sample are listed in the same format as Data S2, using predicted proteins from the *H.californensis* genome.