Title: Ctenophore mesoglea is composed of mucins and novel polysaccharides

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

Affiliations:

[1]Department of Molecular Biophysics & Biochemistry, Yale University; New Haven, CT, USA.

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

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

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

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

[6]Department of Physics, Yale University; New Haven, CT, USA.

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

Abstract: Ctenophora are one of the earliest, if not the earliest branching animal, making them crucial for insight into the early events of metazoan evolution. Ctenophores are largely composed of an extracellular mesoglea, but they lack homologs for the fibrous collagens that form a typical metazoan extracellular matrix (ECM). Therefore, the nature of this body-dominating material and its physiological role remains unknown. We found that ctenophore mesoglea is dominated by mucin-related proteins and novel sulfated polysaccharides. Mucus-like material that functions as an internal skeleton rather than as an external secretion indicates fundamental differences in epithelial function and ECM physiology between ctenophores and other Metazoa, suggesting an early period of biochemical and biomechanical diversity prior to the innovation of elastic connective tissue.

Main Text: Ctenophores are one of the earliest extant (possibly the earliest) branching animal taxa (Fig. 1A) and live in nearly every environment of the world's oceans, across wide temperature, pressure, and salinity ranges (*1–5*). Despite their key position for understanding animal evolution and ubiquity in marine environments, many fundamental aspects of their physiology are unexplored.

The ctenophore body consists of two thin epithelial layers, the epidermis and the gastrodermis, separated by a voluminous layer of gelatinous mesoglea, illustrated in Fig. 1C and 1D (*6–8*). The mesoglea contains sparse, isolated cells interspersed with bulk extracellular hydrogel material of largely unknown composition. This gelatinous mesoglea forms ~99% of the organism's volume, yet few studies have reported on its composition or physiological role. Franc (1985) provided the first, and only, focused description of the biochemical composition of the mesoglea. Franc hypothesized that the mesoglea is formed by a loose three-dimensional network of glycoproteins and sulfated polysaccharides (*7*). This study and others have also observed sparse "bundle" structures containing polymers with superficial similarity to collagens (*7, 9, 10*). However, ctenophores lack the genes coding for most collagens, and specifically lack homologs to the fibrous, gel-forming collagens that make up the volume-spanning extracellular tissues of other animals (*11, 12*). Ctenophores do have homologs to type IV collagen, which contributes to the sheet-like extracellular basement membrane of metazoan epithelia rather than expanded extracellular tissues (*11–13*). Similarly, ctenophores appear to lack fibronectin or tenascin, further differentiating the expanded extracellular material of ctenophores from other metazoans (*12*). Imaging studies, such as those from Norekian & Moroz (2018), have revealed that ctenophore mesoglea supports neuromuscular fibers that extend inward from epithelial layers (*14, 15*). Because the mesoglea dominates the body volume, prior compositional measurements of whole animals presumably reveal some information about its composition. Ctenophores' organic material has been previously assumed to be mostly protein, with little lipid or carbohydrate (*16–24*). Most lipids in ctenophores are membrane phospholipids, so are unlikely to participate in the mesoglea structure (*25*). However, elemental analyses show significant non-protein nitrogen, suggesting the presence of uncharacterized glycans, though these have not been carefully assayed (*18*).

Despite ctenophores' key evolutionary position, there has been no modern description of the body-dominating mesoglea. Thus, we present the first thorough biochemical characterization of ctenophore mesoglea. We focus on *Mnemiopsis leidyi* and *Pleurobrachia pileus*, two species present in the Long Island Sound that represent different clades within ctenophores, Nuda and Tentaculata, respectively (Fig. 1B,C).

Unexpectedly, we found that the mesoglea appears more similar to metazoan mucus than the fibrous-collagen-dominated extracellular matrices (ECMs) of other metazoans. For this study, we define mucus to be an extracellular aqueous solution rich in mucin proteins and complex polysaccharides. Since animal mucuses are typically secreted to epithelial surfaces that are topologically external to the organism, such as skin and gut linings, this finding is surprising and potentially marks a novel mucus function as a volume-filling skeleton. We also report a surprising lack of evidence for covalent linkage between novel mesoglea-constituent sulfated polysaccharides and the mucin protein homologs present in mesoglea, which if confirmed with more specialized experiments, would mark a major chemical structural difference between ctenophore mesogleal mucins and other metazoan mucins.

Ctenophore mesoglea contains mucin homologs

We identified extracellular proteins in mesoglea using MS-based proteomics with predicted proteins from the chromosome-level northern Atlantic *M. leidyi* genome and the *H. californensis* genome for

P. pileus (*4, 26*). For each MS hit, we used BLASTp to annotate predicted biological function based on sequence identity with known proteins. Of the extracellular proteins identified, mucus-related proteins were the most represented family (Fig. 2A,B). The mucin domain has high polymorphism (*27*), so protein domain architecture must also be analyzed to determine possible mucin homology.

The consensus definition of a functional gel-forming mucin is the serial arrangement of several protein domains: repeated VWD-C8-TIL domains with a C-terminal PTS domain (*27*). We analyzed proteins in our dataset with mucin-like homology using InterProScan and PTSPred search (*28, 29*). While all predicted mucin-like transcripts in these ctenophores had at least a partial set of these domains, we located the full arrangement in one set of proteins in a tandem array in the genome of both species (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 also contains all the necessary mucin components; thus, if expressed together, these genes may build a larger mucin protein complex. This possibility was also noted by Lang and colleagues (2016), who identified mucin proteins in both *M. leidyi* and *P. bachei*. For *P. pileus*, we identified an MS hit (Hcv1.av93.c7.g869.i1) using the *H. californensis* genome that contains all the mucin components in the definitional arrangement but with a very short PTS domain. Similar to *M. leidyi*, 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 definitional arrangement, Hcv1.av93.c7.g868.i1. If expressed together, these regions potentially encode two mucin proteins. While we do not find all of these proteins in our proteomics results, we do find at least one of them expressed in both species.

The VWD, TIL, C8, plus PTS domain architecture is a conserved feature of metazoan mucins, but the complete functions of these domains are not yet known, so it is difficult to extrapolate insights to ctenophore mucin structure and function. One motif with established function in mucins is the use of cysteine-rich domains CysD, cysteine knots (CK), and C8, for cross-linking polymerization via disulfide bonds (*30–35*). In human mucins, CysD domains are interspersed within the PTS regions and act as additional locations for interstrand disulfide bonding (*31*). In contrast, 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 (*30*). Interestingly, the presence of multiple CysD domains increases the stiffness of mucins and mucus-related molecules (*32*), so the absence of CysD in ctenophore mucins is possibly related to the liquid-like quality of ctenophore mesoglea. Similarly, CK motifs play roles in C-terminal dimerization, the first step of the mucin assembly process (*33*). The C-terminus of ctenophore mucins lacks the CK motif, though there are multiple CK sites throughout the VWD-TIL-C8 repeats along the N-terminus (separate from the C8 domains). These N-terminal cysteines may also contribute to dimerization, both via disulfide bonds and hydrophobic interactions (*31, 35*). We infer from these observations that ctenophore mucins are likely capable of further polymerization, but the differences in the regions responsible for multimerization are large enough that the resulting assemblies are likely to be quite structurally different from other metazoan mucin complexes.

The other mucin-related proteins we identified (Fig. 2A,B) do not all have typical mucin domains, but may complex with other proteins to perform other mucus-related functions. For example, mucus often has a high concentration of actins, myosins, and other cytoskeletal proteins, which play both structural and immune-related roles in these materials (*36, 37*). We similarly identified numerous actins, myosins, and tubulins in ctenophore mesoglea (data S1–S4), which

may participate in unknown internal mucus functions or could also be debris of sheared neuromuscular fibers (*14, 15*). However, the relative abundance of these cytoskeletal proteins makes it seem unlikely that cellular debris is the sole source of these proteins in our samples. We also identified many immune-related proteins (data S1 and S3) that often act in concert with mucin proteins (*36, 38*). Significant further research will be required to understand the functions of these proteins in the ctenophores' internal milieu.

Novel glycans in ctenophore mesoglea

Glycoproteins, proteoglycans, glycosylaminoglycans (GAGs), and free polysaccharides are constituent components of characterized metazoan connective tissue and mesoglea (*18, 39–46*). We searched for these macromolecules in mesoglea via Nuclear Magnetic Resonance (NMR) Spectroscopy. Various NMR experiments were performed with two protease-treated *M. leidyi* mesoglea biological replicates; spectra that produced solvable structures are presented here. In Figure 3A, a representative ^{1}H-^{13}C HSQC NMR spectrum for one of the *M. leidyi* samples demonstrates two sulfated-galactose-containing polymers (other replicates and DEAE fractionation confirm residues shown in fig. S1A,B; major peaks are described in table S1). The two solved 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 structures are consistent with our findings of high galactose content in gas chromatography-mass spectroscopy (GC-MS) (chromatograms shown in fig. S2A, quantified in table S2). Gal-Glc-Hyl without sulfation is common in collagens (*47–50*), but to our knowledge a sulfated Gal-Glc-Hyl residue is a novel structure.

In *P. pileus*, we found several sulfated mannan structures, as well as low levels of the same sulfated Gal-Glc-Hyl glycopeptide found in *M. leidyi*. ^{1}H-^{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, showing different degrees of sulfation and of homo- or hetero-polymerization. Six additional minor mannose-containing residues were also identified (labeled P, Q, R, Mr, Np, and Oq), although we could resolve only a limited number of these signals, so chemical shift assignments 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). GC-MS also showed some GalNAc and GlcNAc structures that we did not observe via NMR. These differences between techniques may result from low abundance or complex linkages to peptides. Notably, significant protein signals were observed in *P. pileus* via ^{1}H NMR despite extensive protease treatment (fig. S3), suggesting the presence of complex interactions among macromolecules that are difficult to parse.

GalNAc and GlcNAc are common modifications to mucins (*41, 51, 52*). However, neither NMR nor glycoproteomics revealed protein modifications, linkages, or glycosyltransferases consistent with the glycosylation of mucins or other proteins in *M. leidyi* or *P. pileus* mesoglea. Sample digestion did produce a moderate amount of insoluble material that could not be analyzed via NMR, such that glycosylated moieties may have been missed. Similarly, O-glycosylation common to mucins can be difficult to identify without specialized sample treatments (such as mucinases), and glycosylation biochemistry is understudied (there have been no reports of glycosyltransferases in ctenophores, despite an assumption that all metazoans have glycoproteins), so more specialized studies of mucins in ctenophores are still required for a full accounting of

mesoglea composition (*53*). If further studies support our preliminary finding that ctenophore mucins are unglycosylated, this may be correlated with their short PTS domains of only 20-80 residues, compared to 100+ in typical metazoan mucins (*54*). We hypothesize that ctenophore mucin proteins may instead complex with glycans via non-covalent interactions to build a mucus structure, a known property of other mucus-like materials such as biofilm complexes (*41, 55, 56*). Given the extraordinary diversity of both mucin sequences and glycans (*57*), clarifying structure-function relationships of ctenophore mucin proteins and their associated glycans will be crucial for understanding mesoglea evolution and architecture.

Major protein constituents of ctenophore mesoglea

Via proteomics and NMR, we also observed evidence of type IV collagen. Our data further confirm the presence of only the type IV isoform, which is not known to form fibrils but instead makes a thin network in the epithelial basement membrane (Fig. 2A,B, data S1, and data S3) (*12*). The Gal-Glc-Hyl structure we identified in both species (Fig. 3B) is common in collagens, including collagen type IV (*50*), and as we did not identify any other collagen types, this structure is presumably part of ctenophores' type IV collagen. However, total amino acid analyses indicated low relative hydroxylysine concentration in *M. leidyi*, so collagen IV is not likely to be a major constituent of the mesoglea and could come from epithelial debris (table S4). Additionally, we observed only moderate levels of glycine and proline in the mesoglea for both species; the amino acid content of collagenous tissues is typically dominated by glycine and proline, further evidence that it is unlikely that ctenophore mesoglea is primarily collagen-based (*58–60*).

Interestingly, one of the top ten protein MS hits in both *M. leidyi* and *P. pileus* had significant identity with the protein mesocentin (Fig. 2A,B, data S1, and data S3), a giant, extracellular glycoprotein for which there is only one functional description from *Caenorhabditis elegans* (*61*). Mesocentin contains multiple filamin domains interspersed with large disordered domains. In *C. elegans*, it maintains epithelial polarity and orientation of the pharynx with respect to the body wall (*61*). Given the sizable and largely acellular separation between the ctenophore epithelium and gut, it is intriguing to hypothesize similar functions for this protein in ctenophores. As is typical for MS-based proteomics, many unidentified spectra remain, so additional proteins participating in mesoglea structure remain unidentified (*62*).

Development of a mucus skeleton

We find that ctenophore mesoglea is largely a colloidal composite of mucin and novel sulfated polysaccharides that supports sparsely distributed cells of varying types (Fig. 4A). Atomic Force Microscopy (AFM) of mesoglea showed regular 100-nm-scale colloid-like particles in the absence of an obvious network structure (figs. S4 and S5), suggesting the sulfated carbohydrates and structural proteins we documented may form colloidal aggregates in a weakly bonded colloidal gel (*40*), unlike the strong collagenous networks more typical of metazoan connective tissue. Surprisingly, synovial fluid may be a more analogous material to ctenophore mesoglea than the elastic, fibrous-collagen-based mesoglea of cnidarians. This fluid is soft yet resilient due to the presence of mucopolysaccharides and lubricin, a protein that contains mucin domains along with other non-mucin motifs (*63–65*). This finding has broad functional and evolutionary implications about the use of colloidal mucin-based materials (*66*). Mucus can have diverse functions, including physical protection, chemical filtration, and lubrication. However, animal mucus is typically

secreted from the apical surface of a polarized epithelium in a thin layer external to the organism, whereas ctenophore mesoglea is a voluminous internal tissue.

Given its structure and composition, the nature of mesoglea development is a puzzle. In metazoans, mucus is typically produced by secretory goblet-type cells within an epithelium, and mucus-containing vesicles are released to the apical surface of the epithelium (*67, 68*). Goblet cells have been observed in ctenophore tentacles but are not well described elsewhere in the body (*55, 69–71*). Mucus is presumably produced in a similar secretory manner in ctenophores, but given the orientation of the two epithelia in the animal (the apical side of the dermal epithelium faces the external environment, and the apical side of the gut epithelium faces the interior of the gut), non-canonical secretion polarization would seem to be required for significant mucus to be deposited between the two basal surfaces of these epithelia. While unusual, alterations to epithelial polarity leading to altered geometry of mucus secretion have been documented. In certain human cancers, epithelial cell polarity inversion can occur due to defects in polarity-related gene expression, causing mucus to be secreted internally rather than externally (*72, 73*). The possibility of altered polarity of secretion is also consistent with the unusual basement membrane structure associated with epithelial tissues in some ctenophores (*12*), and with the presence of mesocentin, which alters epithelial polarity when mutated (*61,74,75*). Similarly, ctenophores lack two of nine major proteins responsible for the production and maintenance of metazoan epithelial polarity, further suggesting a less-strict regulation of the polarity of secretion in this group relative to later-branching metazoan taxa (*71, 76*).

Our results suggest two broad evolutionary findings. Firstly, while ctenophore epithelia have a basement membrane with type IV collagen, the interstitial mesoglea is a colloidal composite of mucins and novel sulfated polysaccharides and is therefore not capable of bearing loads or elastic storage of mechanical energy, in keeping with the organisms' locomotion via macrocilia (Fig. 4B, fig. S6). Secondly, ctenophore mucin proteins appear to be minimally or possibly non-glycosylated, in sharp contrast to other metazoan mucins. This may represent a transition in animals from undecorated mucins to decorated mucins (Fig. 4C). In *M. leidyi* and *P. pileus*, the protein composition and the presence of a sulfo-glycosylated hydroxylysine associated with collagen assembly are similar. In contrast, the composition of free polysaccharides differs, perhaps reflecting different physiological requirements or other differences between the two clades within ctenophores.

In summary, we find that ctenophore mesoglea features short mucin polymers in a milieu of novel, complex polysaccharides dominated by sulfated polygalactose in *M. leidyi*, and sulfated polymannans in *P. pileus*. These macromolecules likely interact via non-covalent counterionic bonding, leading to a highly hydrated and sparse colloidal gel interspersed with actins and myosins similar to mucus in other metazoans (*36, 37*). The resulting material has a liquid-like elastic modulus (fig. S6), making it functionally distinct from the energy-storing mesoglea and connective tissues of other metazoa, in particular those of cnidarian jellyfish and sponges, whose canonical function is to store and release mechanical energy (*77–79*).

Ctenophores are among the least understood animals, and correlatedly, there is a long history of underestimating their complexity and misinterpreting their biology (*80*). Because they are round and gelatinous, they were long grouped with Cnidaria (which includes medusoid jellyfish) in a clade called Coelenterata (*9, 81*). This led to surprise when phylogenetic analyses found that they diverged from other animals earlier than cnidarians or poriferans (*3, 4*). Our findings, in the context of the more resolved placement of ctenophores as sister to other metazoans, suggest there

was extensive evolutionary experimentation with material properties, macromolecular constituents, biomechanical strategies for being larger than a single cell, and the function of the clade-defining epithelium of multicellular animal bodies that occurred prior to the innovation of storage of elastic energy within a network of fibrillar collagen.

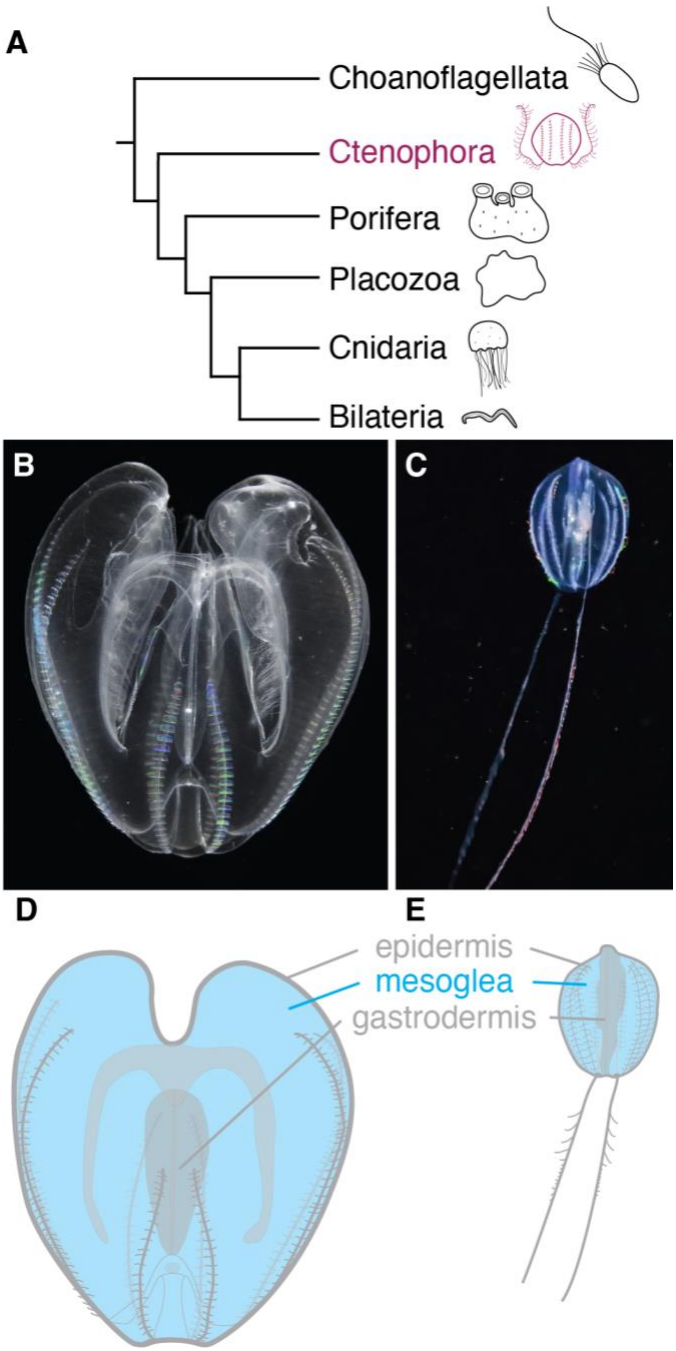

Figure 1: **Ctenophores have bulk gelatinous mesogleal tissue.** (**A**) Ctenophores are one of our most distant animal relatives. (**B**) *Mnemiopsis leidyi* is a lobate ctenophore, photo taken by SERC/MarineGEO Chesapeake BioBlitz Team. (**C**) *Pleurobrachia pileus* is a cydippid ctenophore. Image cropped from ”Sea Gooseberry (Pleurobrachia pileus) in Norway” by Vsevolod Rudyi, 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.

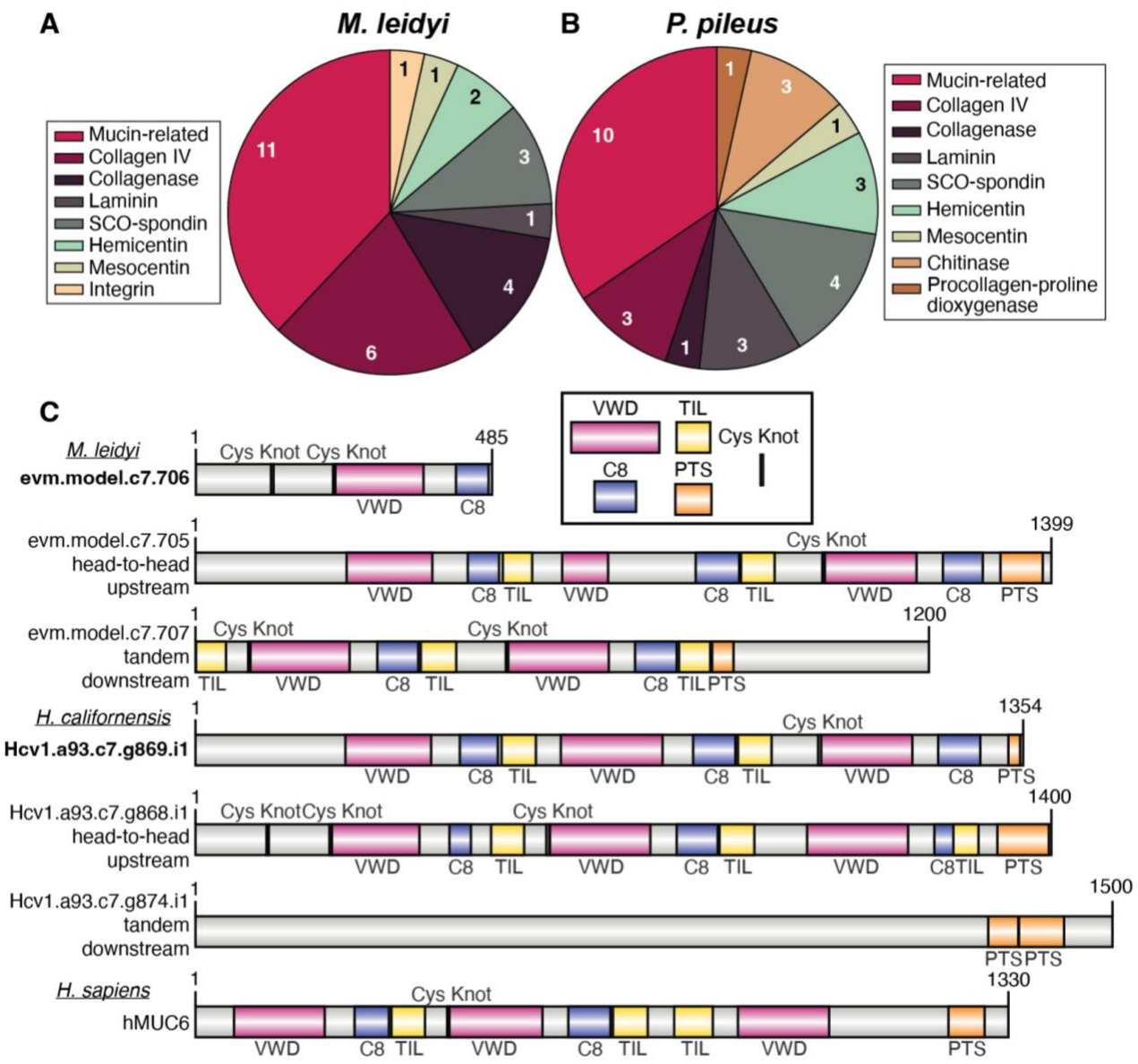

Figure 2: **Mucin-related proteins dominate mesoglea protein content.** (**A**) The major extracellular proteins identified in *M. leidyi* from LC-MS/MS proteomics. (**B**) The major extracellular proteins identified in *P. pileus* from LC-MS/MS proteomics using an *H. californensis* database. (**C**) Domain architecture of *M. leidyi* and *H. californensis* proteins with homology to metazoan mucins. Proteins that were observed in the proteomics results are noted in bold, and proteins with mucin-related domains that are either in tandem array or head-to-head orientation with the expressed proteins are also shown. hMUC6 is a human mucin protein with the characteristic organization of mucin functional domains. *Homo sapiens* mucin-6 isoform X3 was used (NCBI: XP 054185958.1).

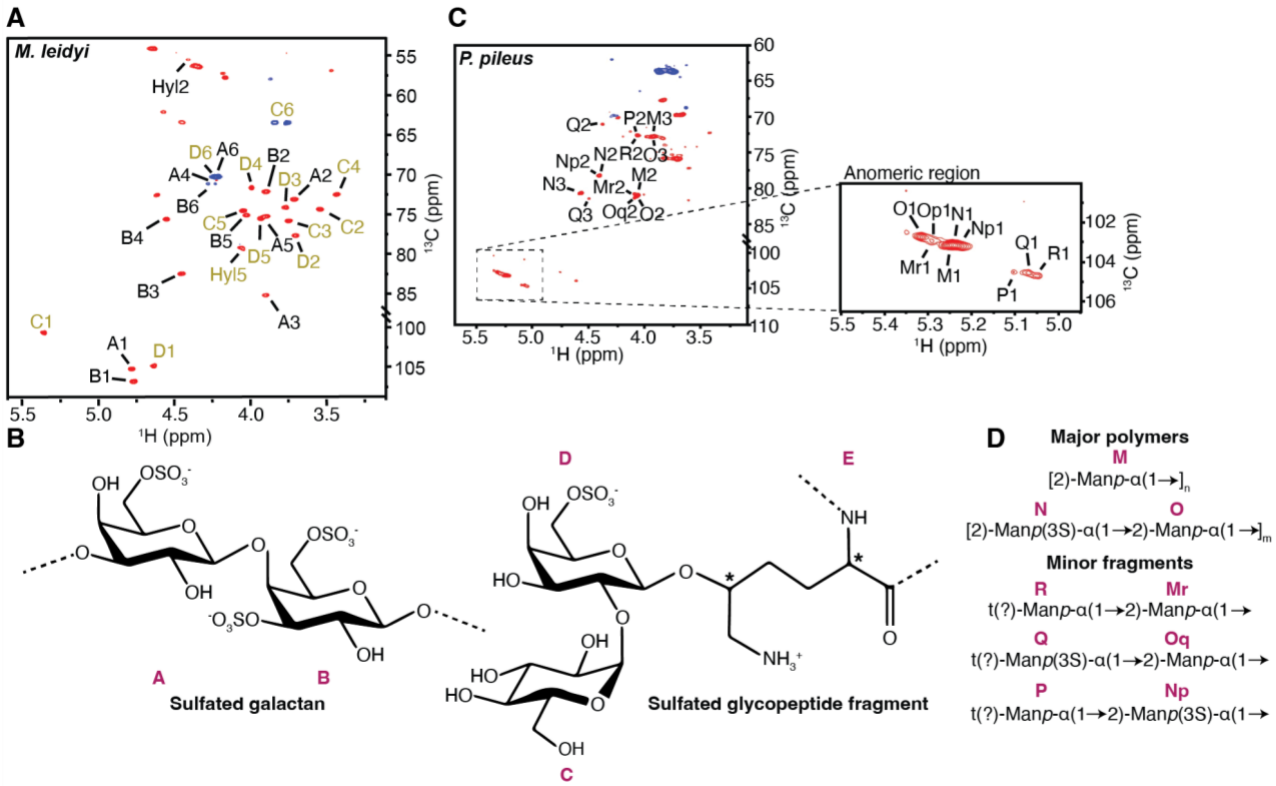

Figure 3: **Ctenophore mesoglea contains novel sulfated polysaccharides.** (**A**)^{1}H-^{13}C HSQC NMR spectra of an *M. leidyi* mesoglea sample, containing signals of both a sulfated galactan and a sulfated glycopeptide fragment. Anomeric and ring residue signals of the sulfated galactan disaccharide (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.

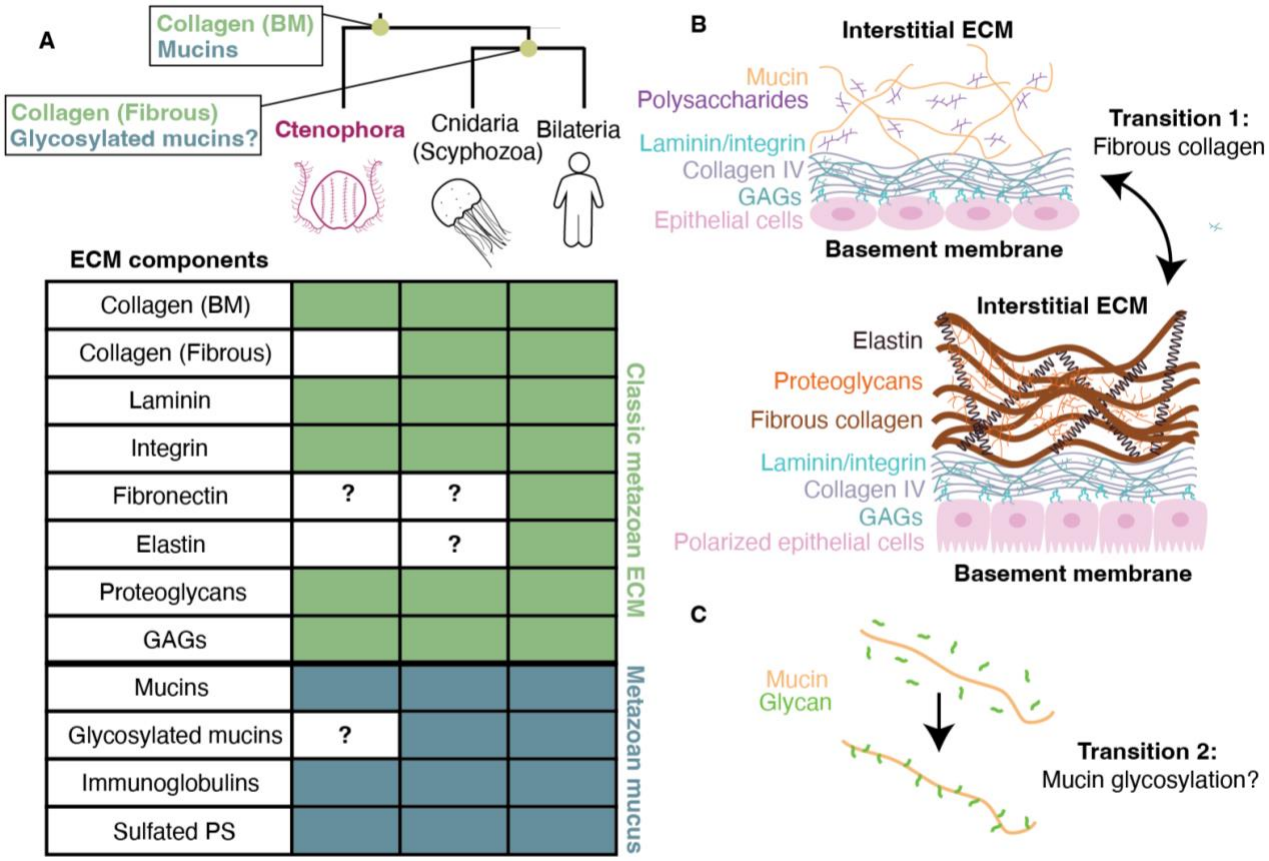

Figure 4: **Early animal evolution shows material experimentation. (A)** A simplified metazoan cladogram highlighting that the last common ancestor of all metazoans likely possessed mucins and basement membrane (BM) collagens. After ctenophores diverged, fibrous collagens and glycosylated mucins likely evolved, giving rise to two major extracellular material classes: the fibrous collagenous ECM (classic metazoan ECM) and mucus. Bilaterians retain the full complement of both, whereas ctenophores lack key components, leading to functional differences. Components that need further verification have a question mark. **(B)** In ctenophores, the ECM consists of mucin and polysaccharide (PS) molecules, and their epithelial cells are less organized, allowing mucus secretion into the mesoglea (top). In most metazoans, the basement membrane connects to polarized epithelial cells through integrins and laminins, creating a thin sheet of collagen type IV and glycosaminoglycans (GAGs). Their interstitial ECM is a bulk network of fibrous collagen, elastins, and proteoglycans (bottom). **(C)** A second potential transition of extracellular materials is the development of mucin glycosylation. Ctenophores do not appear to have glycosylated mucins, so their mucin proteins (orange) are likely bare and undecorated with no covalently linked glycans (green), compared to the highly decorated mucins of other metazoans.

References and Notes

1. D. T. Schultz, W. R. Francis, J. D. McBroome, L. M. Christianson, S. H. D. Haddock, R. E. Green, A chromosome-scale genome assembly and karyotype of the ctenophore Hormiphora californensis. *G3: Genes|Genomes|Genetics* **11** (11), jkab302 (2021).

2. J. F. Ryan, K. Pang, C. E. Schnitzler, A-D. Nguyen, R. T. Moreland, D. K. Simmons, B. J. Koch, W. R. Francis, P. Havlak, NISC Comparative Sequencing Program, S. A. Smith, N. H. Putnam, S. H. D. Haddock, C. W. Dunn, T. G. Wolfsberg, J. C. Mullikin, M. Q. Martindale, A. D. Baxevanis, The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science* **342** (6164), 1242592 (2013).

3. C. W. Dunn, A. Hejnol, D. Q. Matus, K. Pang, W. E. Browne, S. A. Smith, E. Seaver, G. W. Rouse, M. Obst, G. D. Edgecombe, M. V. Sørensen, S. H. D. Haddock, A. Schmidt-Rhaesa, A. Okusu, R. M. Kristensen, W. C. Wheeler, M. Q. Martindale, G. Giribet, Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452** (7188), 745–749 (2008).

4. D. T. Schultz, S. H. D. Haddock, J. V. Bredeson, R. E. Green, O. Simakov, D. S. Rokhsar, Ancient gene linkages support ctenophores as sister to other animals. *Nature* **618**, 110–117 (2023).

5. A. J. Jamieson, D. J. Lindsay, H. Kitazato, Maximum depth extensions for Hydrozoa, Tunicata and Ctenophora. *Marine Biology* **170** (3), 33 (2023).

6. S. Franc, J. M. Franc, R. Garrone, Fine Structure and Cellular Origin of Collagenous Matrices in Primitive Animals: Porifera, Cnidaria and Ctenophora. *Burkitt lymphoma, hemostasis and intercellular matrix. Frontiers of matrix biology* **3**, 143–156 (1976).

7. J-M. Franc, thesis, Claude Bernard Lyon 1 University (1985).

8. J. Ramon-Mateu, S. T. Ellison, T. E. Angelini, M. Q. Martindale, Regeneration in the ctenophore Mnemiopsis leidyi occurs in the absence of a blastema, requires cell division, and is temporally separable from wound healing. *BMC Biology* **17** (1), 80 (2019).

9. 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).

10. 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).

11. B. Barzansky, H. M. Lenhoff, On the Chemical Composition and Developmental

Role of the Mesoglea of Hydra. *American Zoologist* **14** (2), 575–581 (1974).

12. A. L. Fidler, C. E. Darris, S. V. Chetyrkin, V. K. Pedchenko, S. P. Boudko, K. L. Brown, W. Gray Jerome, J. K. Hudson, A. Rokas, B. G. Hudson, Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. *eLife* **6**, e24176 (2017).
13. 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).
14. T. P. Norekian, L. L. Moroz, Neural system and receptor diversity in the ctenophore. *Journal of Comparative Neurology* **527** (12), 1986–2008 (2018).
15. P. Burkhardt, J. Colgren, A. Medhus, L. Digel, B. Naumann, J. J. Soto-Angel, E-L. Nordmann, M. Y. Sachkova, M. Kittelmann, Syncytial nerve net in a ctenophore adds insights on the evolution of nervous systems. *Science* **380** (6642), 293–297 (2023).
16. 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).
17. G. Schneider, Zur chemischen Zusammensetzung der Ctenophore Pleurobrachia pileus in der Kieler Bucht. *Helgolander Meeresunters* **43**, 67–76 (1989).
18. 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).
19. T. G. Bailey, M. J. Youngbluth, G. P. 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).
20. T. G. Bailey, M. J. Youngbluth, G. P. 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).
21. J. A. Percy, F. J. Fife, The Biochemical Composition and Energy Content of Arctic Marine Macrozooplankton. *ARCTIC* **34** (4), 307–313 (1981).
22. U. Hoeger, Biochemical composition of ctenophores. *Journal of Experimental Marine Biology and Ecology* **72** (3), 251–261 (1983).
23. M. Yousefian, A. Erkan Kideys, Biochemical composition of Mnemiopsis leidyi in the southern Caspian Sea. *Fish Physiology and Biochemistry* **29** (2), 127–131 (2003).
24. B. E. Anninsky, G. A. Finenko, G. I. Abolmasova, E. S. Hubareva, L. S.

Svetlichny, L. Bat, A. E. Kideys, 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).

25. R. F. Lee, Lipids of Zooplankton from Bute Inlet, British Columbia. *Journal of the Fisheries Research Board of Canada* **31** (10), 1577–1582 (1974).

26. V. Koutsouveli, M. Torres-Oliva, T. Bayer, J. Fuß, N. Grossschmidt, A. M. Marulanda-Gomez, N. Jensen, D. Gill, R. A. Schmitz, L. Pita, T. B. H. Reusch, 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).

27. T. Lang, S. Klasson, E. Larsson, M. E. V. Johansson, G. C. Hansson, T. Samuelsson, Searching the Evolutionary Origin of Epithelial Mucus Protein Components—Mucins and FCGBP. *Molecular Biology and Evolution* **33** (8), 1921–1936 (2016).

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

29. 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).L. Khmelnitsky, A. Milo, O. Dym, D. Fass, Diversity of CysD domains in gel-forming mucins. *The FEBS Journal* **290** (21), 5196–5203 (2023).

30. L. Khmelnitsky, A. Milo, O. Dym, D. Fass, Diversity of CysD domains in gel-forming mucins. *The FEBS Journal* **290** (21) 5196–5203 (2023).

31. D. Fass, D. J. Thornton, Mucin networks: Dynamic structural assemblies controlling mucus function. *Current Opinion in Structural Biology* **79**, 102524 (2023).

32. B. Demouveaux, V. Gouyer, C. Robbe-Masselot, F. Gottrand, T. Narita, J-L. Desseyn, Mucin CYS domain stiffens the mucus gel hindering bacteria and spermatozoa. *Scientific Reports* **9** (16993) (2019).

33. 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).

34. L. E. Morgan, A. M. Jaramillo, S. K. Shenoy, D. Raclawska, N. A. Emezienna, V. L. Richardson, N. Hara, A. Q. Harder, J. C. NeeDell, C. E. Hennessy, H. M. El-Batal, C. M. Magin, D. E. Grove Villalon, G. Duncan, J. S. Hanes, J. S. Suk, D. J.

Thornton, F. Holguin, W. J. Janssen, W. R. Thelin, C. M. Evans, Disulfide disruption reverses mucus dysfunction in allergic airway disease. *Nature Communications* **12** (1), 249 (2021).

35. R. Bansil, B. S. Turner, Mucin structure, aggregation, physiological functions and biomedical applications. *Current Opinion in Colloid & Interface Science* **11** (2), 164–170 (2006).

36. J. Jurado, C. A. Fuentes-Almagro, F. A. Guardiola, A. Cuesta, M. Esteban, M-J. Prieto-Alamo, Proteomic profile of the skin mucus of farmed gilthead seabream (*Sparus aurata*). *Journal of Proteomics* **120**, 21–34 (2015).

37. 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).

38. I. R. Dubbelboer, V. Barmpatsalou, A. Rodler, E. Karlsson, S. F. Nunes, J. Holmberg, J. Häggström, C. A. S. Bergström, Gastrointestinal mucus in dog: Physiological characteristics, composition, and structural properties. *European Journal of Pharmaceutics and Biopharmaceutics* **173**, 92–102 (2022).

39. J. E. Scott, Structure and function in extracellular matrices depend on interactions between anionic glycosaminoglycans. *Pathologie Biologie* **49** (4), 284–289 (2001).

40. X. Zhang, K. Xia, L. Lin, F. Zhang, Y. Yu, K. St. Ange, X. Han, E. Edsinger, J. Sohn, R. J. Linhardt, Structural and Functional Components of the Skate Sensory Organ Ampullae of Lorenzini. *ACS Chemical Biology* **13** (6), 1677–1685 (2018).

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

42. E. B. Ledesma, C. Holland, Analysis, classification and identification of gastropod locomotive mucus by Fourier transform infrared spectroscopy. *Journal of Molluscan Studies* **90** (1) (2024).

43. K. Karamanou, D. C. R. Espinosa, A. Fortuna-Costa, M. S. G. Pavão, Biological function of unique sulfated glycosaminoglycans in primitive chordates. *Glycoconjugate Journal* **34** (3), 277–283 (2017).

44. J. A. Santos, B. Mulloy, P. A. S. Mourão, Structural diversity among sulfated α-L-galactans from ascidians (tunicates). *European Journal of Biochemistry* **204** (2), 669–677 (1992).

45. C. J. Jackson, I. Giles, A. Knop, A. Nethery, L. Schrieber, Sulfated Polysaccharides Are Required for Collagen-Induced Vascular Tube Formation. *Experimental Cell Research* **215** (2), 294–302 (1994).

46. B. Moussian, The role of GlcNAc in formation and function of extracellular matrices. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **149** (2), 215–226 (2008).

47. R. G. Spiro, Characterization and Quantitative Determination of the Hydroxylysine-linked Carbohydrate Units of Several Collagens. *Journal of Biological Chemistry* **244** (3), 602–612 (1969).

48. T. Hennet, Collagen glycosylation. *Current Opinion in Structural Biology* **56**, 131–138 (2019).

49. T. Basak, L. Vega-Montoto, L. J. Zimmerman, D. L. Tabb, B. G. Hudson, R. M. Vanacore, Comprehensive Characterization of Glycosylation and Hydroxylation of Basement Membrane Collagen IV by High-Resolution Mass Spectrometry. *Journal of Proteome Research* **15** (1), 245–258 (2016).

50. 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).

51. S. T. M. Lee, S. K. Davy, S-L. Tang, P. S. Kench, Mucus Sugar Content Shapes the Bacterial Community Structure in Thermally Stressed Acropora muricata. *Frontiers in Microbiology* **7**, 371 (2016).

52. D. Mentino, C. De Blasi, D. Semeraro, M. Mastrodonato, M. V. Guglielmi, Bivalves and Gastropods: Models for the Study of Mucomics. *Journal of Marine Science and Engineering* **13** (3), 566 (2025).

53. P. H. Jensen, D. Kolarich, N. H. Packer, Mucin-type O-glycosylation – putting the pieces together. *The FEBS Journal* **277**, 81–94 (2009).

54. 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).

55. A. McShane, J. Bath, A. M. Jaramillo, C. Ridley, A. A. Walsh, C. M. Evans, D. J. Thornton, K. Ribbeck, Mucus. *Current Biology* **31** (15), R938–R945 (2021).

56. H-C. Flemming, J. Wingender, The biofilm matrix. *Nature Reviews Microbiology* **8** (9), 623–633 (2010).

57. G. Raba, A. Luis, Mucin utilization by gut microbiota: recent advances on characterization of key enzymes. *Essays in Biochemistry* **67** (3), 345–353 (2023).

58. N. M. Khong, F. M. Yusoff, B. Jamilah, M. Basri, I. Maznah, K. W. Chan, J. Nishikawa, Nutritional composition and total collagen content of three commercially important edible jellyfish. *Food Chemistry* **196**, 953–960 (2016).

59. 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).

60. M. Gauza-Wlodarczyk, L. Kubisz, D. Wlodarczyk, Amino acid composition in determination of collagen origin and assessment of physical factors effects. *International Journal of Biological Macromolecules* **104**, 987–991 (2017).

61. R. B. Proenca, thesis, The Johns Hopkins University (2002).

62. K. Ning, D. Fermin, A. I. Nesvizhskii, Computational Analysis of Unassigned High Quality MS/MS Spectra in Proteomic Datasets. *Proteomics* **10** (14), 2712–2718 (2010).

63. H. Fam, J. Bryant, M. Kontopoulou, Rheological properties of synovial fluids. *Biorheology: The Official Journal of the International Society of Biorheology* **44** (2), 59–74 (2007).

64. B. Zappone, G. W. Greene, E. Oroudjev, G. D. Jay, J. N. Israelachvili, Molecular Aspects of Boundary Lubrication by Human Lubricin: Effect of Disulfide Bonds and Enzymatic Digestion. *Langmuir* **24** (4), 1495–1508 (2008).

65. B. N. Feng, A. J. Marks, F. Kim, M. Naik, T. A. Schmidt, K. A. Elsaid, G. Jay, B. Rubenstein, Lubricin's Mucin Domain Has Strong Polyproline Type-II Helical Character. *Computational and Structural Biotechnology Journal* **35** (1), 0021 (2026).

66. M. J. M. Wells, H. A. Stretz, Supramolecular architectures of natural organic matter. *Science of The Total Environment* **671**, 1125–1133 (2019).

67. R. D. Specian, M. R. Neutra, Mechanism of rapid mucus secretion in goblet cells stimulated by acetylcholine. *Journal of Cell Biology* **85** (3), 626–640 (1980).

68. G. M. H. Birchenough, M. E. V. Johansson, J. K. Gustafsson, J. H. Bergström, G. C. Hansson, New developments in goblet cell mucus secretion and function. *Mucosal Immunology* **8** (4), 712–719 (2015).

69. I. Eeckhaut, P. Flammang, C. Lo Bue, M. Jangoux, Functional morphology of the tentacles and tentilla of Coeloplana bannworthi (Ctenophora, Platyctenida), an ectosymbiont of Diadema setosum (Echinodermata, Echinoida). *Zoomorphology* **117** (3), 165–174 (1997).

70. J-M. Franc, ORGANIZATION AND FUNCTION OF CTENOPHORE COLLOBLASTS: AN ULTRASTRUCTURAL STUDY. *The Biological Bulletin* **155** (3), 527–541 (1978).

71. T. Brunet, D. S. Booth," Cell polarity in the protist-to-animal transition" in *Current Topics in Developmental Biology* (Elsevier, 2023), vol. 154, pp. 1–36.

72. V. N. Adsay, K. Merati, H. Nassar, J. Shia, F. Sarkar, C. R. Pierson, J. D. Cheng, D. W.

Visscher, R. H. Hruban, D. S. Klimstra, Pathogenesis of Colloid (Pure Mucinous) Carcinoma of Exocrine Organs: Coupling of Gel-Forming Mucin (MUC2) Production With Altered Cell Polarity and Abnormal Cell-Stroma Interaction May Be the Key Factor in the Morphogenesis and Indolent Behavior of Colloid Carcinoma in the Breast and Pancreas. *The American Journal of Surgical Pathology* **27** (5), 571 (2003).

73. C. Canet-Jourdan, D-L. Pagès, C. Nguyen-Vigouroux, J. Cartry, O. Zajac, C. Desterke, J.-B. Lopez, N. Signolle, J. Adam, J. Raingeaud, M. Polrot, P. Gonin, J. R. R. Mathieu, S. Souquere, G. Pierron, M. Gelli, P. Dartigues, M. Ducreux, V. Barresi, F. Jaulin, Patient-derived organoids identify an apico-basolateral polarity switch associated with survival in colorectal cancer. *Journal of Cell Science* **135** (14), jcs259256 (2022).

74. M. K. Chong, E. R. Cebul, K. Mizeracka, M. G. Heiman, Loss of the Extracellular Matrix Protein DIG-1 Causes Glial Fragmentation, Dendrite Breakage, and Dendrite Extension Defects. *Journal of Developmental Biology* **9** (4) (2021).

75. C. Y. Bénard, A. Boyanov, D. H. Hall, O. Hobert, DIG-1, a novel giant protein, non-autonomously mediates maintenance of nervous system architecture. *Development* **133** (17), 3329–3340 (2006).

76. H. Belahbib, E. Renard, S. Santini, C. Jourda, J-M. Claverie, C. Borchiellini, A. Le Bivic, New genomic data and analyses challenge the traditional vision of animal epithelium evolution. *BMC Genomics* **19** (1), 393 (2018).

77. W. Gladfelter, Structure and function of the locomotory system of Polyorchis montereyensis (Cnidaria, Hydrozoa). *Helgolander wiss. Meeresunters* **23**, 38–79 (1972).

78. C. Gambini, B. Abou, A. Ponton, A. J. M. Cornelissen, Micro- and Macrorheology of Jellyfish Extracellular Matrix. *Biophysical Journal* **102** (1), 1–9 (2012).

79. E. A. Kraus, L. E. Mellenthin, S. A. Siwiecki, D. Song, J. Yan, P. A. Janmey, A. M. Sweeney, Rheology of marine sponges reveals anisotropic mechanics and tuned dynamics. *Journal of The Royal Society Interface* **19** (195), 20220476 (2022).

80. 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).

81. G. Scholtz, Coelenterata versus Acrosomata – zur Position der Rippenquallen (Ctenophora) im phylogenetischen System der Metazoa, in *Kontroversen in der Phylogenetischen Systematik*, Richter, S., Sudhaus, W., Eds. (Sber. Ges. Naturf. Freunde zu Berlin (N.F.), Berlin), vol. 43, pp. 15–33 (2004).

82. I. M. Black, I. E. Ndukwe, J. Vlach, J. Backe, B. R. Urbanowicz, C. Heiss, P. Azadi, 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).

83. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59** (4), 687–696 (2005).

84. S. P. Skinner, R. H. Fogh, W. Boucher, T. J. Ragan, L. G. Mureddu, G. W. Vuister, CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *Journal of Biomolecular NMR* **66** (2), 111–124 (2016).

85. T. Masuko, A. Minami, N. Iwasaki, T. Majima, S-I. Nishimura, Y. C. Lee, Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Analytical Biochemistry* **339** (1), 69–72 (2005).

86. 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).

87. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Zˇ emgulyte˙, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Zidek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630** (8016), 493–500 (2024).

88. M. Van Kempen, S. S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C. L. M. Gilchrist, J. So¨ding, M. Steinegger, Fast and accurate protein structure search with Foldseek. *Nature Biotechnology* **42** (2), 243–246 (2024).

89. C. J. A. Sigrist, E. de Castro, L. Cerutti, B. A. Cuche, N. Hulo, A. Bridge, L. Bougueleret, .I. Xenarios, New and continuing developments at PROSITE. *Nucleic Acids Research* **41** (D1), D344–D347 (2013).

90. 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), 681–694 (2001).

91. Y. Xie, H. Li, X. Luo, H. Li, Q. Gao, L. Zhang, Y. Teng, Q. Zhao, Z. Zuo, J. Ren, IBS 2.0: an upgraded illustrator for the visualization of biological sequences. *Nucleic Acids Research* **50** (W1), W420–W426 (2022).

92. S. K. Lai, Y-Y. Wang, D. Wirtz, J. Hanes, Micro- and macrorheology of mucus.

Advanced Drug Delivery Reviews **61** (2), 86–100 (2009).

Acknowledgments:
We thank the Yale Center for Research Computing for the computational resources on the Grace cluster for BLASTp analyses, the Yale Institute for Nanoscience and Quantum Engineering for AFM usage, and Jing Yan for rheometer usage. We also thank AltaBioscience for performing the total amino acid analysis. Lauren Mellenthin, Mary Beth Decker, and other members of the Dunn Lab provided valuable 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 #DE-SC0015662 to PA.

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 PA.

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 to PA.

This work was also supported by an award to AS and CD from the Yale Natural Lands Committee.

A Graduate Research Fellowship from the National Science Foundation to SS also supported general research.

Author contributions:
Conceptualization: SS, CD AS
Methodology: SS, SAH, IB, IN, JV, PA, CD, AS
Investigation: SS, SAH, IB, IN, JV
Visualization: SS, IB, JV, CD, AS
Funding acquisition: SS, PA, CD, AS
Supervision: PA, AS
Writing – original draft: SS, AS
Writing – review & editing: SS, JV, CD, AS

Competing interests: Authors declare that they have no competing interests.

Data, code, and materials availability: All raw data on Dryad, doi will be provided upon publication.

Supplementary Materials
Materials and Methods
Figs. S1 to S6

Tables S1 to S4
References (*82–92*)
Data S1 to S4

Supplementary Materials for

Ctenophore mesoglea is composed of mucins and novel polysaccharides

Sara Siwiecki, Stephanie Archer-Hartmann, Ian Black, Ikenna Ndukwe, Jiri Vlach, Parastoo Azadi, Casey Dunn, Alison Sweeney
Corresponding author: alison.sweeney@yale.edu

The PDF file includes:

Materials and Methods
Figs. S1 to S6
Tables S1 to S4
Captions for Data S1-S4
References *(82-92)*

Other Supplementary Materials for this manuscript include the following:

Data S1 to S4

Materials and Methods

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 #0323011 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.

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 the 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.

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 (*82*). 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 42 µl was 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 (*83*) or 3.2.0 (*84*). We rendered figures in MestReNova 14.2.3 (Mesterelab Research). We referenced ^{1}H and ^{13}C chemical shifts 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 (*85*). 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 of the DEAE separation contained carbohydrates and was then run on the NMR according to the method above.

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 (*26*). 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 by Schultz et al. (2023) and accessible on GenBank under GCA 020137815.1 (*4*). 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 (*86*). 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 (*28*). 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. Additionally, structural searches were conducted 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/ (*87*). 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 (*88*). The e-values of all hits with structural homology are reported in our supplementary data (data S2 and 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 Lang and colleagues (2004), to identify PTS domains (*29*). For some mucin-like proteins, we also performed visual inspections to identify PTS-rich sequences 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 (*89, 90*). Illustrator for Biological Sequences 2.0 was used for protein domain architecture visualization (*91*).

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.

Atomic Force Microscopy (AFM)

We used 15 mm muscovite mica discs (Electron Microscopy Sciences, P/N: 71856-03) as a surface for sample testing, freshly cleaving a mica slice for each sample. We then deposited 5 µl of each mesoglea sample onto the mica disc. We tried several sample layering techniques to identify the most successful protocol, including wet samples, simple air dry, and using a spin coater to level and thin the sample and then air drying. The most successful protocol was depositing a sample onto the disc then immediately applying the disc to a spin coater. After a few minutes of spinning, the disc was removed, producing a thin layer of dried sample. Other sample preparation methods did not produce clear images. We used a few mesoglea samples from different specimens of *M. leidyi* and *P. pileus* for testing and verified the protocol using κ-carrageenan, successfully visualizing the gel network.

A Bruker Dimension FastScan Atomic Force Microscope was used to understand the structural morphology of ctenophore mesoglea. Mica discs with prepared samples were placed onto the granite platform to begin scanning. We tried multiple tip types and found SCANASYST-AIR and FASTSCAN-B tips (Bruker) to pair best with our soft mesoglea samples. For each experiment, a new tip was mounted onto the cantilever. The AFM was used in Scan Assist Mode (peak force mode) with the Scan Assist in Air setting. The laser was calibrated and aligned at the start of each experiment. The following parameters were used for each experiment: starting scan size of 15 µm, aspect ratio of 1, no x/y offset, scan rate of 4 Hz or 1 Hz, samples/line at 256 or 512, ScanAsyst Auto Control on, peak force amplitude at 150 nm, and lift height of 300 nm. Each scan produced a Height Sensor Image and a Peak Force Error Image. For each experiment, clear images were saved and exported for interpretation.

Rheology

All rheology experiments were performed on a shear rheometer (Anton Paar Physica MCR502WESP) in strain-control modes. Various measuring geometries, such as cone-and-plate and parallel plate, were initially used, but yielded noisy results. Therefore, a double gap system was used to maximize contact area. The inner and outer cylinders sandwich the rotational cup, which provides a large surface area for the sample to contact the measuring system. The double gap system used was compatible with the MCR equipment (Anton Paar DIN 54453; measuring cylinder: B-DG26.7Q1, measuring cup C-DG26.7T200SS). A Peltier Temperature Control Device (Anton Paar C-PTD200) designed for the double gap geometry was used to maintain constant temperature during all experiments. The temperature was set at 17 ◦C, consistent with the temperature of the ctenophore habitat.

Only *M. leidyi* results are reported here due to low sample size of *P. pileus* being available at the time of rheometry experiments. Approximately 5 ml of ctenophore mesoglea was deposited

via pipette between the inner and outer cylinders. Controls of soybean oil, gelatin, and water were used to ensure the system was operating properly.

We used amplitude sweeps to identify the linear viscoelastic strain range, then we measured the viscoelastic properties, specifically elastic and viscous moduli, using frequency sweeps. Mesoglea from a total of three *M. leidyi* specimens were used for amplitude sweeps. Here, we reported elastic and viscous moduli extracted from the viscoelastic region of frequency sweeps, specifically values resulting from shear strain applied at 0.01 Hz and 1% strain. All relevant amplitude and frequency sweep experiments are included in the raw data.

In brief, the elastic and viscous moduli are calculated as follows. Upon applying shear strain, the rheometer recorded the time (*seconds*), deflection (*radians*), torque (*N-m*), applied shear strain (*%*), and shear stress response (*Pa*). For this study, shear strain (Eq. S1) is defined as γ (unitless), shear deformation, s (*meters*), and shear gap, h (*meters*). Since we are using double gap geometry, h is a fixed value (often defined as an instrument-specific factor) and s is equal to the radius of the cylinder at the liquid interface r times the angular displacement, θ, in radians.

$$Equation\ S1:\ \gamma = \frac{s}{h} = \frac{r \cdot \theta}{h}$$

Shear stress is defined by Equation S2 with τ as shear stress, F for force (N), and A for area (m^2).

$$Equation\ S2:\ \tau = \frac{F}{A}$$

A basic shear modulus can be calculated to describe a material's properties at specific time points (Equation S3).

$$Equation\ S3:\ G = \frac{\tau}{\gamma}$$

$$Equation\ S4: \gamma(t) = \gamma_o sin(\omega t)$$

$$Equation\ S5:\ \tau(t) = \tau_o cos(\omega t)$$

The strain is applied as an oscillatory shear by the rheometer and the stress response results in a waveform, noted in Equations S4 and S5, where γ_o is the strain amplitude, τ_o is the stress amplitude, ω is the angular frequency, and δ is the phase lag between the stress and strain waveforms. To ensure that the signal of the stress response is higher than instrument noise, we evaluated the stress response waveforms in relation to the applied strain. If a given strain produced a sinusoidal stress response, the strain was considered valid, and Microsoft Excel was used to record a database of (elastic) modulus, G', and loss (viscous) modulus, G'', from the RheoCompass software as a function of γ. G' is defined in Equation S6 and G'' is defined in Equation S7.

$$Equation\ S6:\ G' = \frac{\tau_o}{\gamma_o} cos\delta$$

$$Equation\ S7:\ G'' = \frac{\tau_o}{\gamma_o} sin\delta$$

The viscoelastic moduli were visualized as a function of γ for the valid strain range using MATLAB 2023a. The linear viscoelastic (LVE) region was evaluated via visual inspection by identifying regions where *G'* and *G''* exhibit plateaus across strains. The LVE region is defined as the strain amplitude range where viscoelastic moduli are independent of strain. Once we identified the strain values that produced viscoelastic moduli within the LVE region, we performed frequency sweeps. A frequency range, such as 0.01-5 Hz, starting with the highest frequency and decreasing to the slowest frequency until the sweep was complete. We then performed the same analysis described above for the amplitude sweeps to evaluate the raw waveform data and ensure we only continued studying samples with high signal and low noise.

Then, we created a database of the frequencies that produced data with a sinusoidal stress response. Using the understanding of the LVE region, we extracted the calculated values of the storage (elastic) modulus, *G'*, and loss (viscous) modulus, *G''*, from the RheoCompass software as a function of ν for sweeps performed with strains in the LVE region, such as 1%. The viscoelastic moduli were visualized compared to other tissues using MATLAB 2023a.

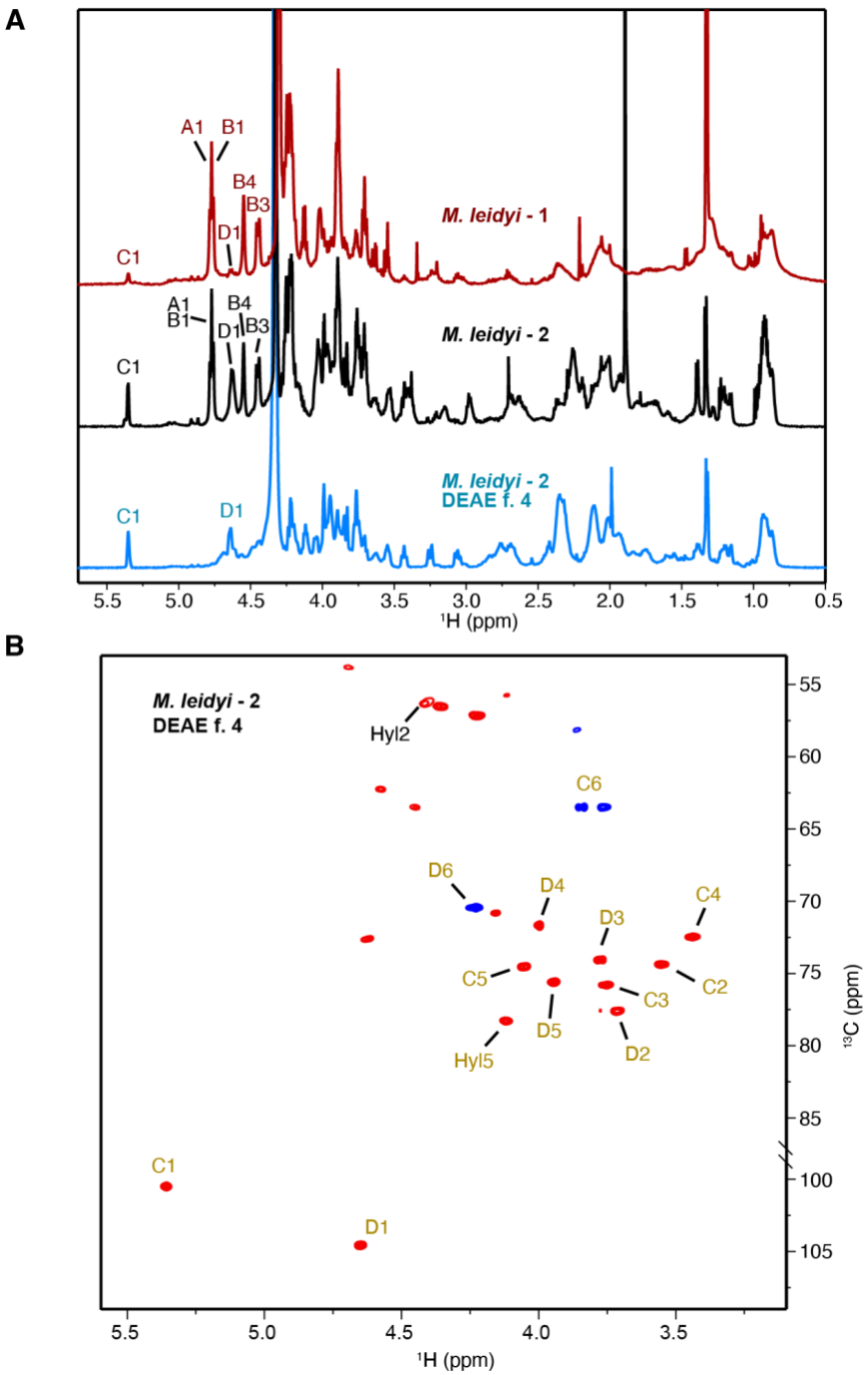

Fig. S1. Unique sulfated galactans are present in *M. leidyi* mesoglea. (A) ^{1}H NMR spectra of two *M. leidyi* mesoglea samples. The bottom sample is DEAE fraction 4 of the previous *M. leidyi* sample. (B) ^{1}H-^{13}C HSQC NMR spectra of the DEAE fraction 4 of the second *M. leidyi* mesoglea sample, where only signals of sulfated glycopeptide fragment were present, confirming that sulfated galactan and sulfated glycoprotein were not covalently bound. The sulfated glycan fragment (gold) and the amino acid fragment (black) are indicated. Unlabeled signals arose from residual peptides.

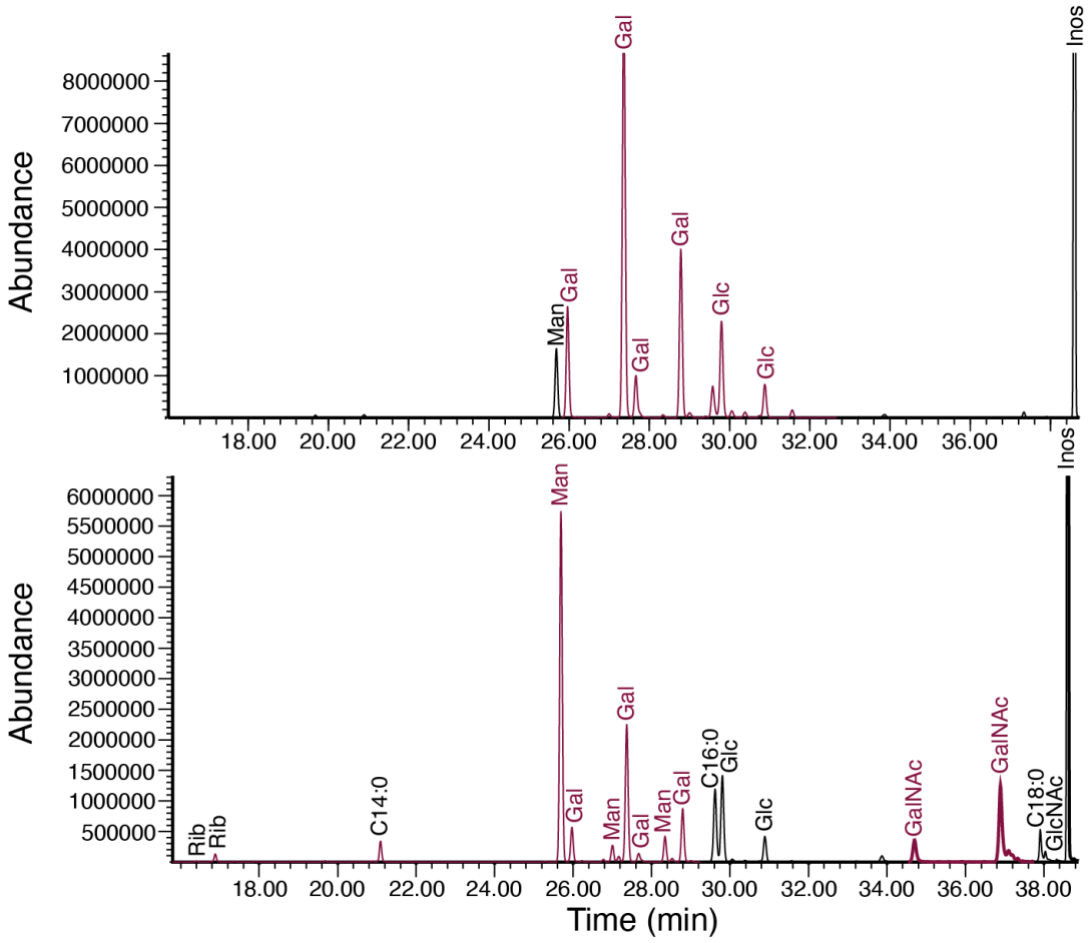

Fig. S2. Galactose-based monosaccharides are abundant in ctenophore mesoglea. (A) Total ion chromatograms from GC-MS of *M. leidyi*, highlighting high galactose content. (B) Total ion chromatograms from GC-MS of *P. pileus*, highlighting high mannose content. Most abundant peaks are highlighted in red. Chromatograms were taken after derivatization of the sample that underwent protease treatment and dialysis. n=1 for both species, and inositol is present in both plots as a reference.

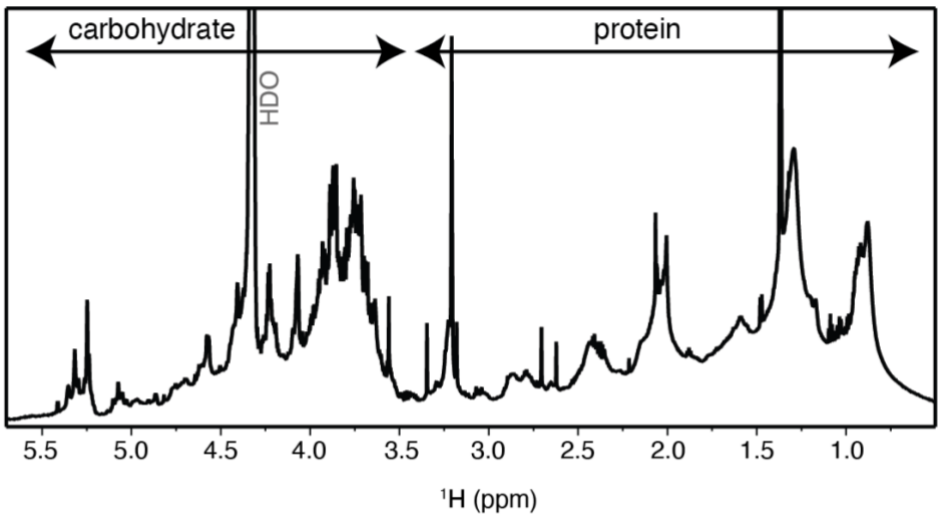

Fig. S3. Sulfated mannan structures and residual protein found in *P. pileus*. ^{1}H NMR spectra of one *P. pileus* mesoglea sample, highlighting significant protein and carbohydrate signal.

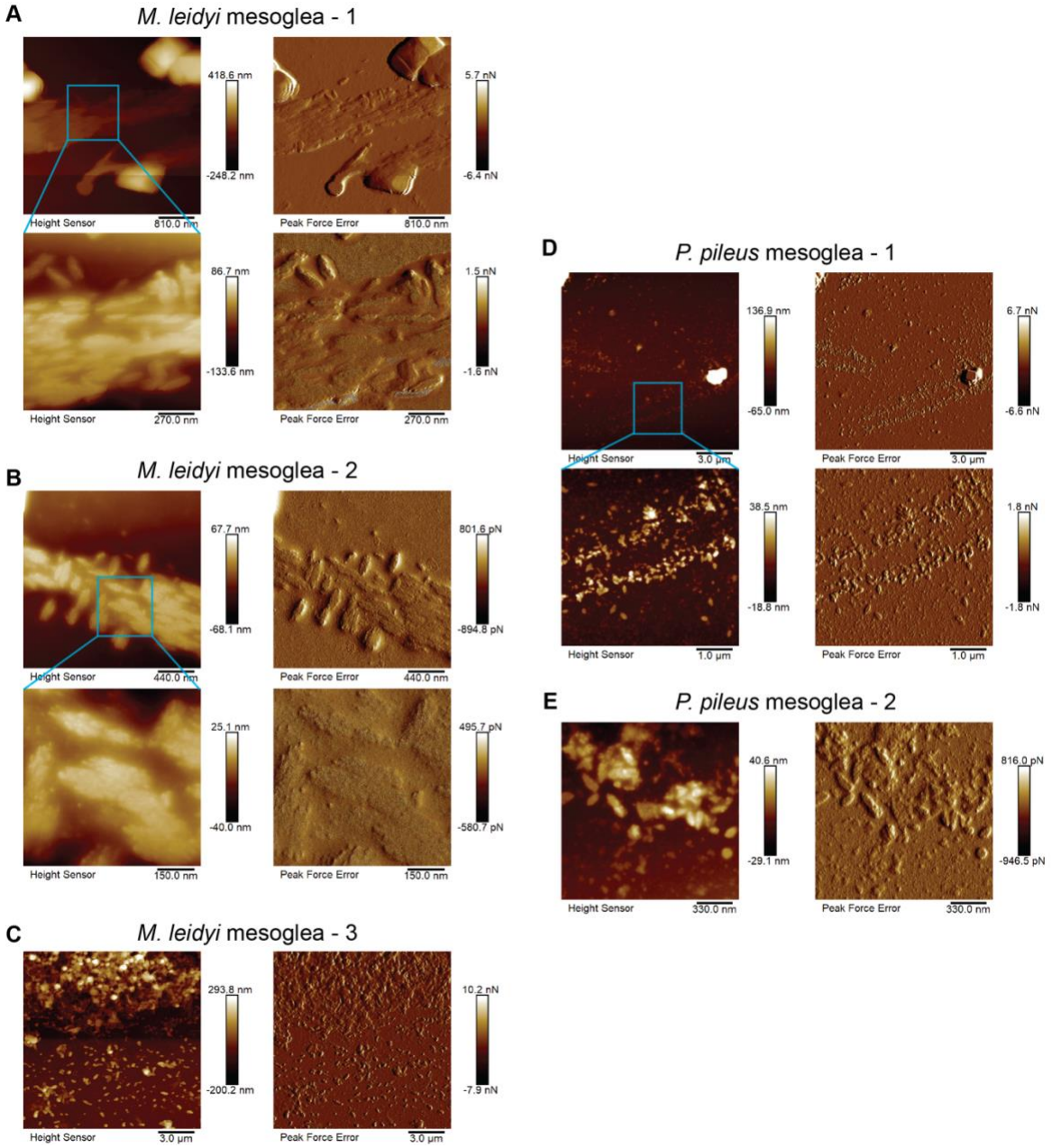

Fig. S4: AFM height sensor and peak force error images representing organized colloidal structures in ctenophore mesoglea. (A): An example of *M. leidyi* mesoglea, zooming in on an organized structure featuring rod-like morphology at 800 nm and 270 nm. The left image in all AFM results is the height sensor image and the right image is the peak force error image. (B): Another *M. leidyi* mesoglea sample showing organized structures with rod-like morphology at 400 nm and 150 nm. (C): An additional *M. leidyi* mesoglea sample showing a dense, organized structure, with many aggregates in a gradient at 3 µm. (D): A representative sample of *P. pileus* mesoglea showing two aggregated, organized structures at two levels, 3 µm and 1 µm, both also highlighting a speckled background. (E): Another P. pileus sample showing an aggregate of rod-like structures with speckled background at 330 nm.

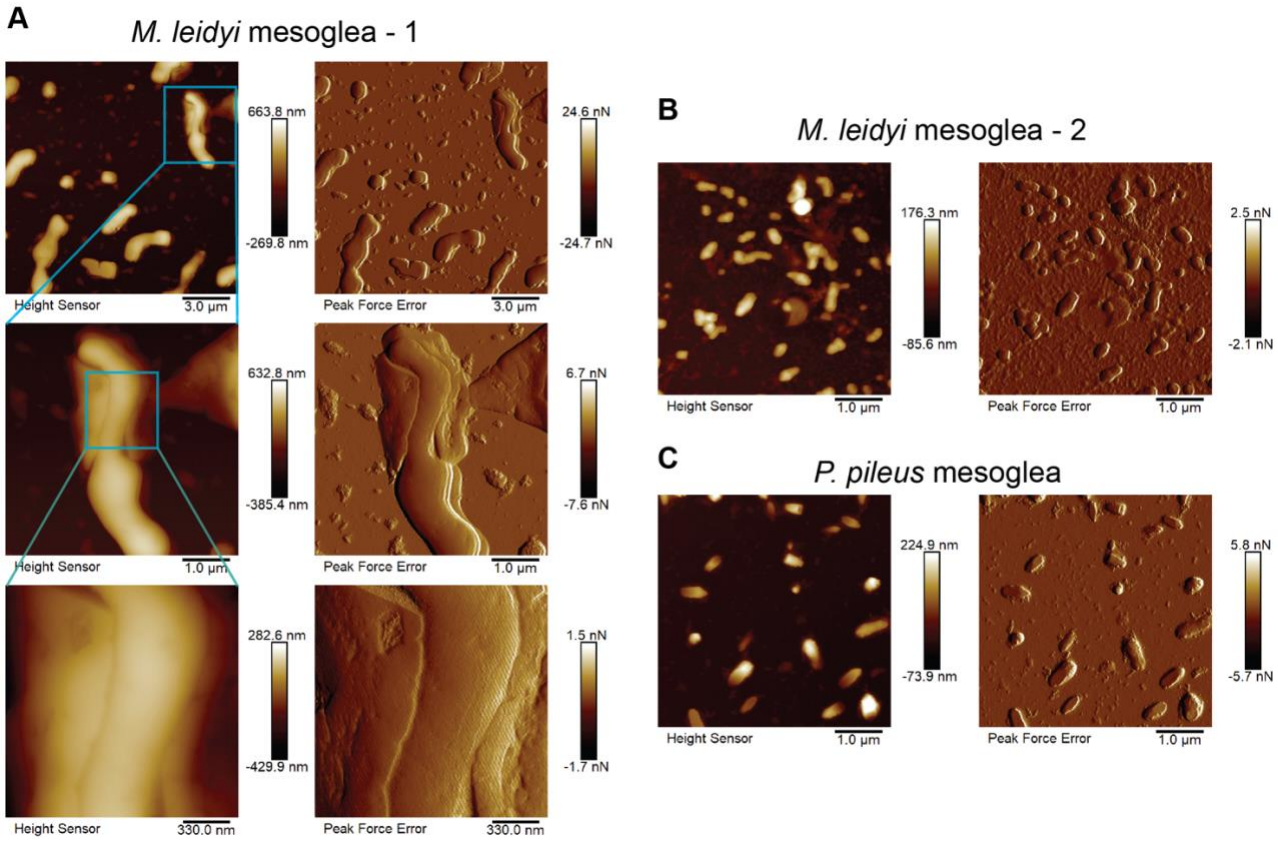

Fig. S5: **AFM height sensor and peak force error images representing amorphous blob colloidal structures in ctenophore mesoglea.** (A): An example of *M. leidyi* mesoglea, zooming in on a particular aggregate structure. (B): Another *M. leidyi* mesoglea sample showing a variety of blob structures at 1 µm. (C): A representative *P. pileus* mesoglea sample exhibiting blob structure at 1 µm.

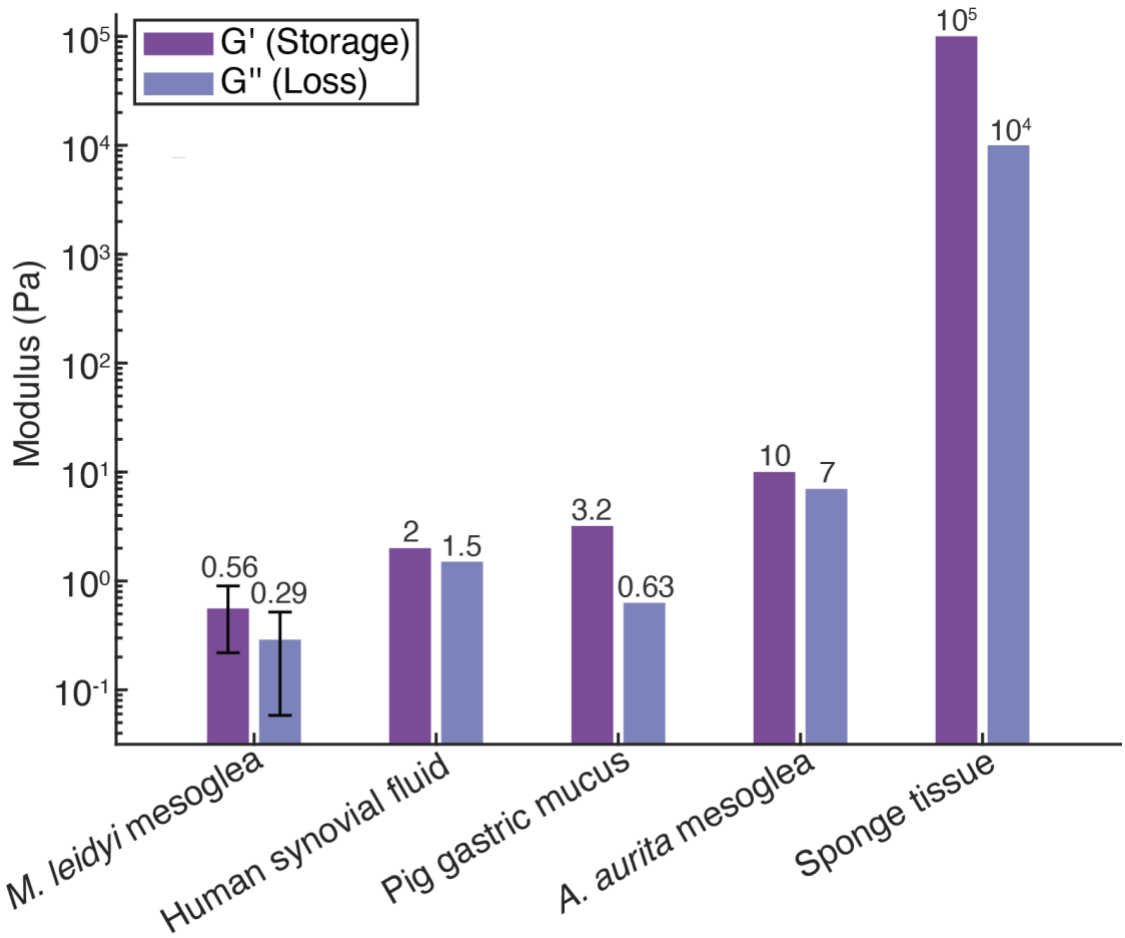

Figure S6: **Rheological properties of ctenophore mesoglea are more similar to fluids than typical viscoelastic tissues in metazoans.** Average viscoelastic moduli of *M. leidyi* mesoglea compared to other metazoan tissues (*63, 78, 79, 92*). *G'* and *G''* for *M. leidyi* mesoglea are extracted from the viscoelastic region of a frequency sweep and represent viscoelastic properties from shearing applied at 0.01 Hz and 1% strain. n=3 for *M. leidyi* mesoglea and error bars represent standard error.

Table S1. ^{1}H-^{13}C NMR chemical shifts of galactan residues found in *M. leidyi* (Fig. 3A,B and fig. S1).

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

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

Table S2. GC-MS (fig. S2) highlights high galactose in ctenophore mesoglea. (n.d. = no data)

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**	

Table S3. ^{1}H-^{13}C NMR chemical shifts of sulfated mannan residues found in *P. pileus* (Fig. 3C,D and fig. S3).

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

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%**

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

Data S1. (separate file) 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.

Data S2. (separate file) 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.

Data S3. (separate file) 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.

Data S4. (separate file) 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.