Published in final edited form as:

J Proteome Res. 2013 October 4; 12(10): 4547–4555. doi:10.1021/pr4006734.

Anomalous N-glycan structures with an internal fucose branched to GlcA and GlcN residues isolated from a Mollusk shell-forming fluid

Hui Zhou^{(1),(3)}, Andrew J. Hanneman^{(1),(2)}, N. Dennis Chasteen⁽⁴⁾, and Vernon N. Reinhold^{(1),(2)}

⁽¹⁾Glycomics Center, 35 Colovos Road, University of New Hampshire, Durham, NH 03824

⁽²⁾Glycan Connections, LLC, 293 River Road, Lee, NH 03861

⁽⁴⁾Department of Chemistry, 23 Academic Way, University of New Hampshire, Durham, NH 03824

Abstract

This report describes the structural details of a unique N-linked valence epitope on the major protein within the extrapallial (EP) fluid of the mollusk, *Mytilus edulis*. Fluids from this area are considered to be responsible for shell expansion by a self-assembly process that provides an organic framework for the growth of CaCO₃ crystals. Previous reports from our laboratories have described the purification and amino acid sequence of this EP protein which was found to be a glycoprotein (EPG) of approximately 28 KDa with 14.3% carbohydrate on a single N-linked consensus site. Described herein is the *de novo* sequence of the major glycan and its glycomers. The sequence was determined by ion trap mass spectrometer (ITMSⁿ) resolving structure by tracking precursor-product relationships through successive rounds of collision induced disassociation (CID), thereby spatially resolving linkage and branching details within the confines of the ion trap. Three major glycomers were detected, each possessing a 6-linked fucosylated N-linked core. Two glycans possessed four and five identical antennae, while the third possessed four antennae, but with an additional methylfucose 2-linked to the glucuronic acid moiety, forming a pentasaccharide. The tetrasaccharide structure was: 4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc, while the pentasaccharide was shown to be: mono-*O*-methyl-Fuc(1-2)-4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc. Samples were differentially deuteriomethylated (CD₃/CH₃) to localize indigenous methylation, further analyzed by HRMS to confirm monomer compositions, and finally GC-MS to assign structural and stereoisomers. The interfacial shell surface location of this major extrapallial glycoprotein, its calcium heavy metal binding properties and unique structure suggests a probable role in shell formation and possibly metal ion detoxification. A closely related terminal tetrasaccharide structure has been reported in spermatozoan glycolipids of freshwater bivalves.

Correspondence to: Vernon N. Reinhold.

³Current affiliation, Department of Urology, Boston Children's Hospital and Harvard Medical School, 300 Longwood Ave, Boston, MA 02115.

Honoring N. Dennis Chasteen, Professor of Biophysical Chemistry, UNH upon his retirement

Supporting Information

Supporting information is supplied, including a schematic describing the modified PMAA procedure, GC-MS data, and additional MSⁿ spectra. The supporting information is available free of charge via the internet at <http://pubs.acs.org/>.

Keywords

Mytilus edulis; extrapallial fluid; *N*-glycans; branched fucose; GC-MS; HRMS; ITMSⁿ; sequential disassembly; 4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc; *O*-methyl-Fuc(1-2)-4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc; (+) MSⁿ; (–) MSⁿ

Introduction

Molluscan shell formation offers an ideal opportunity to study biomineralization, the process of converting inorganic ions into the hardened structures needed to carry out diverse biological functions. Calcification, acting in concert with organic macromolecules, including proteins and oligosaccharides, impacts a wide range of metabolic processes including ontogenesis, mineral ion storage, nutrition, ultraviolet protection, reproduction, and the perception of light and magnetic fields^{1, 2}. In spite of the major importance of biomineralization exact mechanisms describing the transformation of ions in solution to mineral biostructures are complex and elusive, but are of considerable interest to biologists and engineers seeking new approaches for various applications ranging from ecotoxicological monitoring³, to pearl cultivation⁴, and the fabrication of hybrid organic-inorganic materials for bone replacement therapy⁵.

Calcium carbonate biomineralization of the mollusk shell supports these soft-bodied organisms, protecting them from predation and desiccation. A number of studies have focused on the composition of the organic/inorganic shell matrix originating from the extrapallial (EP) fluid that fills the cavity between the inner shell surface and the mantle epithelium, forming an integrated system regulating protein interactions with biominerals⁶. Previous reports from our laboratories have described purification and characterization of the *Mytilus edulis* EP protein, accounting for 56% of the total^{7, 8}, the metal binding properties^{7–9}, carbohydrates and amino acids. The fluid is believed to be supersaturated with respect to shell-forming minerals^{7, 9–13}. The primary sequence of EP glycoprotein was characterized by RT-PCR and cDNA techniques and confirmed by peptide sequencing using mass spectrometry⁹. The extrapallial glycoprotein (EPG) is a homodimer of 28,340 Da, consisting of 213 amino acids and approximately 14.3% carbohydrate by weight with a single *N*-glycosylation site at asparagine 54. Ultracentrifugation and polyacrylamide gel electrophoresis demonstrated that this glycoprotein binds calcium and other divalent metals causing it to assemble into a series of higher order then assembles into a series of higher order multimeric species of increasing molecular mass (8) as might be expected of a shell matrix protein. Calcium binding is coupled with a reversible rearrangement of the protein's secondary structure, observed as a reduction of its beta-sheet content⁸. The amino acid sequence indicates it may be closely related to a heavy metal binding protein isolated from hemolymph⁹, and a histidine rich glycoprotein (HRG) from the plasma of the same species, *Mytilus edulis*¹⁴. These metal binding proteins may be present in various tissues^{15, 16}, suggesting diverse biological functions including shell formation, metal ion transportation, and detoxification^{3, 13, 17–19}. All such functions apparently originate from their most intriguing chemical feature: a strong affinity for metal ions, including Cu²⁺, Cd²⁺, Mg²⁺, Mn²⁺, Hg²⁺, Ni²⁺, Pd²⁺, Zn²⁺, and Ca²⁺^{8, 13, 16, 19}. In the case of HRG, the cadmium binding activity persisted even upon heat denaturation, and was associated exclusively with the glycoprotein moiety¹⁴. The physical location of EPG, its metal binding characteristics, and unique structure, suggest a role in shell formation, making it of considerable interest to also characterize its single *N*-linked glycan. Our initial failed attempts to lectin purify this glycoprotein or its glycans, even though positive evidence of PNGase F release were perplexing, but did lead to a detailed study of compositions by methanolysis and GC-MS of their TMS derivatives. Even here, divergent retention times and spectral results contributed

by some unique monomers with indigenous methylation demanded a more comprehensive and integrated MS-based strategy that was subsequently provided by sequential disassembly with ITMSⁿ.

Detailed characterization of carbohydrate structures brings unique challenges; replete with stereo and structural isomers and embedded in branching arrays, these biopolymers have exposed inordinate challenges well beyond the usual MS/MS and CID approaches. Adding to these complications are the variable bond energies, requiring unique strategies during disassembly²⁰. Compared to the well-studied *N*-glycans of mammalian species, deciphering those detected on EPG introduced new array challenges, unusual monomer components and their modifications^{21–23}. Thus, by necessity the MS-based approach described here was developed for a comprehensive *de novo* structural understanding of *N*-glycans. The strategy consists of high resolution mass measurements and GC-MS to determine the monosaccharide composition, (+) mode CID-MSⁿ of permethylated (CH₃) and perdeuteromethylated (CD₃) derivatives to elucidate the antennal sequences, and (–) CID-MSⁿ of the native glycans to position multiple antennae with respect to the chitobiose core. The result was identification of a typical 6-linked fucosylated *N*-glycan core: [Man₃GlcNAc(Fuc)GlcNAc] possessing four complex type antennae, with GlcNAc extended from the trimannosyl core as usual, but surprisingly, then leading to a fucose that is branched to both GlcNAc and 4-*O*-methyl glucuronic acid non-reducing end termini. The second most abundant glycan had an additional antenna (five in total), while the third structure indicated mono-*O*-methyl fucose linked to the C-2 position of the 4-*O*-methyl glucuronic acid. Sequencing the *N*-glycans by ion trap (IT) MSⁿ resolved these structural details and the inter-residue linkage and branching specifics by tracking precursor-product ion relationships through successive rounds of collision induced disassociation (CID) fragmentation. An additional sample aliquot was perdeuteromethylated to pinpoint the site of endogenous methylation by MSⁿ.

Materials and Methods

Materials

All materials were of the highest purity, and used as received. Ammonium bicarbonate, trifluoroacetic acid (TFA), sodium hydroxide (NaOH), dimethyl sulfoxide (DMSO), iodomethane (CH₃I), iodomethane-d₃ (CD₃I), sodium borohydride (NaBH₄), sodium borodeuteride (NaBD₄), borane-ammonia complex, toluene, cellulose (medium fibrous), HPLC grade ethanol, and 1-butanol were obtained from Sigma-Aldrich (St. Louis, MO). Monosaccharide standards for GC-MS were also from Sigma-Aldrich. Chloroform, acetonitrile, and methanol were from EMD chemicals (Gibbstown, NJ), and HPLC grade water from J. T. Baker (Phillipsburg, NJ). Sep-Pak C18 solid-phase extraction (SPE) columns were obtained from Waters (Milford, MA), and graphitized carbon (GC) from Fisher Scientific (Fairlawn, NJ). PNGase F (glycerol free) was purchased from New England Biolabs (Ipswich, MA).

Sample Preparation of *N*-glycans

Native blue mussels, *Mytilus edulis*, were collected from waters off the northeast coast of the United States at the University of New Hampshire Marine Research Facility, New Castle, NH. EP glycoprotein was purified as described previously^{8,9}. The *N*-glycans were released by PNGase F according to the vendor protocols, then separated from the deglycosylated protein by passing through a C18 SPE cartridge in a solution of 5% acetonitrile/95% water/0.1% TFA (v/v/v). After dry-down they were reduced to alditols by re-suspension in a 28% ammonia solution containing 10 mg/mL borane-ammonia complex²⁴. The residual borane was removed by repeated addition and evaporation of methanol with centrifugal vacuum

evaporation. Next, the reduced glycans were further purified by GC SPE. To carry out (–) ESI-MSⁿ of the native glycan sample an additional purification step was performed using a hand packed cellulose SPE column²⁵. The sample was taken up in 200 µL of ethanol: H₂O (1:1, v/v), mixed with 2 mL of an organic solvent mixture comprised of 1-butanol/ethanol/water (4/1/1, v/v/v), and then applied to the cellulose column. The column was washed with 6 mL of the organic solvent mixture and the glycans were eluted with 4 mL ethanol/H₂O (1:1, v/v). For permethylation, the reduced glycans were dissolved in DMSO and treated with powdered sodium hydroxide and iodomethane²⁶, or by using a spin column procedure²⁷. Deuterium methylation with iodomethane-d₃ involved the same protocols.

Mass Spectrometry

MSⁿ disassembly was carried out using an LTQ ion trap (ThermoFisher, San Jose, CA) equipped with chip-based infusion²⁸, (Triversa Nanomate, Advion, Ithaca, NY). MSⁿ employed the following values, collision energy: 35%, activation Q: 0.25, and activation time: 30 ms. High resolution MS was carried out for native glycans on the LTQ using the ultra zoom scan mode, and included an internal standard within the same *m/z* scan width for mass correction (A2 glycan: Neu5Ac₂ Gal₂GlcNAc₂Man₃GlcNAc₂, ProZyme, Hayward, CA). Following permethylation, accurate mass was obtained on an Orbitrap-XL mass spectrometer using the 60,000 resolution scan setting (ThermoFisher, San Jose, CA).

GC-MS Monosaccharide Composition

Monosaccharide composition analysis employed the methanolysis/TMS derivatization procedure²⁹. Additionally, considering hexuronic acids might be naturally modified by a methoxyl, a sample aliquot was processed according to the partially methylated alditol acetates (PMAA) linkage analysis method with one modification, as follows: the *N*-glycan sample was permethylated and mass confirmed by orbitrap-MS, the TFA hydrolysis was carried out as usual (4M, 100°C, 6hrs), followed by deuterium reduction (1M NaBD₄, 0.1M NaOH, RT, overnight) wherein the hexuronic acid standards (glucuronic and galacturonic acid) were introduced into the procedure. Finally, permethylation was used in place of the typical acetylation step (Supplementary Figure 1). This procedure was implemented in order to compare the sample against the commercial hexuronic acid standards because 4-*O*-methyl glucuronic acid was not readily available. GC-MS was carried out using a Polaris Q Trace GC Ultra (Thermo, San Jose, CA) equipped with a DB-5ms, 30 m × 0.25 mm × 25 µm capillary chromatography column (Agilent, Santa Clara CA). The temperature program was: 50°C for 0.5 minutes, ramping to 250°C at 8°C/min, with the injector and transfer line temperatures set to 260°C. Detection was by electron impact (EI) MS.

Results and Discussion

Fucose commonly exists as a chain terminating residue, however, evidence has accumulated of glycosyltransferases capable of adding sugars to fucose³⁰. The most fascinating feature about the *N*-glycans of *M. edulis* EP glycoprotein are their unique antennary sequences: 4-*O*-methyl-GlcA(1–4)[GlcNAc(1–3)]Fuc(1–4)GlcNAc, Glycans (a) and (b), and mono-*O*-methyl-Fuc(1–2)-4-*O*-methyl-GlcA(1–4)[GlcNAc(1–3)]Fuc(1–4)GlcNAc, Glycan (c), (described below) each including a branched internal fucose moiety. Interestingly, similar glycan structures, including internal fucose and *O*-methyl sugars, have been reported within glycosphingolipids isolated from the spermatozoa of the fresh water bivalve *Hyriopsis schlegelii*^{31–34}. This may implicate a unique biosynthetic pathway in mollusks responsible for branched fucosylation, supported by additional reports of the tetrasaccharide moiety: 3-*O*-methyl-Gal-[GalNAc]-Fuc-GlcNAc, identified in the *N*-glycan antennae of the copper-containing respiratory hemocyanin from the sea snail *Rapana thomasi*^{35, 36}. Although the metal binding characteristics of EP glycoprotein were previously ascribed to its amino

acid sequence⁹, based on these new glycan structures, it is reasonable to wonder whether they might also be involved in the ability of EP glycoprotein to bind metal ions. With flexible antennae terminating in acidic residues, and four or five antennae within a single sugar molecule, the affinity for metal ions is likely to be enhanced if binding were coordinated by two or more antennae, helping to explain how molluscan metal-binding glycoproteins might interact with metal ions and retain this affinity upon denaturation¹⁴.

Monosaccharide composition analysis

The three most abundant *N*-glycans detected by (–) ESI-MS, were: (a) 4027.4 Da, (b) 4769.7 Da, and (c) 4667.8 Da (Figure 1). These masses suggested unusual monosaccharides, or other modifications sometimes noted in lower organisms³⁷. Negative ion MS² of Glycan (a) suggested a mono-methoxyl modified hexuronic acid within the antennae, due to facile neutral loss of methanol (–32 Da, *m/z* 998) and mono-methoxyl hexuronic acid (–190 Da, *m/z* 1278), (Figure 5), fitting the composition: (mono-MeO-HexA)₄(HexNAc)₁₀(Hex)₃(DeoxyHex)₅ suggested by accurate mass measurements. The mass errors were: 21 ppm for the reduced native glycan using the LTQ ultrazoom scan mode, and 4.7 ppm for the reduced and permethylated glycan by orbitrap-MS. Monosaccharide composition analysis using classical methanolysis/TMS derivatization identified fucose, mannose, and *N*-acetyl glucosamine by matching against standards, and 4-*O*-methyl-glucuronic acid on the basis of its GC retention time on the DB-5 column and characteristic EI-MS ions³⁸. The modified PMAA procedure confirmed the identification of glucuronic acid by matching the retention time and EI-MS spectrum vs. the standard (Supplementary Figure 2). From these data Glycan (a) was defined as: (mono-*O*-methyl-GlcA)₄(GlcNAc)₁₀(Man)₃(Fuc)₅, with the presence of ten GlcNAc residues and three hexoses suggesting a complex-type *N*-glycan comprised of four antennae. The three structures in Figure 2 were further characterized by (+) and (–) MSⁿ methodologies.

Characterization of Glycan (a) antennary sequence by (+) MSⁿ

MS² of permethylated Glycan (a) produced the distinct antennal B-ion (*m/z* 919.4¹⁺) resulting from rupture of the labile glycosidic bond between antennal GlcNAc and the trimannosyl core (Figure 3A),^{28, 39}. Other major peaks in the MS/MS spectrum were multiply-charged ions corresponding to larger moieties (Supplementary Figure 3). MS³ of the antennal B-ion revealed a non-reducing end trisaccharide consisting of: 4-*O*-methyl-GlcA/GlcNAc/Fuc, linked to position C-4 of the GlcNAc attached to the trimannosyl core, as distinguished by a prominent^{4–6} A cleavage (*m/z* 762.5, Figure 3B),^{20, 40}. Surprisingly, the second GlcNAc and the 4-*O*-methyl-GlcA were each observed as chain terminating residues (*m/z* 687.3 and 660.3, Figure 3B), indicating fucose occupies an internal branching position. Further disassembly at MS⁴ and MS⁵ stages following the pathways: 1240.1⁴⁺ 919.4 660.3, and 1240.1⁴⁺ 919.4 660.3 433.2, confirmed this arrangement and provided diagnostic cross-ring cleavage ions (Figure 3C, 3D) indicating the 4-*O*-methyl-GlcA is linked to C-4, and GlcNAc to position C-3, of the fucose. The antennal sequence of Glycan (a) was characterized as: 4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc.

Location of endogenous methylation by MSⁿ

To confirm the location of endogenous methylation within the antenna, a separate sample aliquot was deuterium permethylated (CD₃). MS² of perdeuteromethylated Glycan (a) distinguished a deuterium methylated antennal B-ion possessing a single natural methoxyl group (*m/z* 952.6). MS³ of the B-ion confirmed a methoxyl on the terminal 4-*O*-methyl-GlcA (*m/z* 711.5, 681.4, Figure 4A). Neutral loss of methanol (CH₃OH, *m/z* 920.5) from C-4 was observed to be favored over any losses of deuterated methanol (CD₃OH, *m/z* 917.5, Figure 4A) suggesting the product ion, upon methanol loss, yields a conjugated double bond

with the C-6 carboxyl. Further MS⁴ of m/z 920.5 indicated Retro Diels-Alder fragmentation relating to the double bond formed at MS³ stage (m/z 770.5 Figure 4B).

Glycan (a) branching topology by (–) MSⁿ

Negative ion fragmentation is a proven technique for the determination of *N*-glycan branching patterns^{41–47}. Diagnostic D- and E-type ions provide information regarding the number and location of antennae linked to the trimannosyl core^{43, 46, 48}. MS² of native Glycan (a) yielded ions at m/z 903.3^{–2} (D-ion) and m/z 867.3^{–2} (0–3A of the central mannose), (Figures 5A, 5B), suggesting two antennae are attached to the 6-mannose arm, and two to the 3-mannose arm. The remaining fragment ions in the spectrum were located within the antenna, confirming the permethylated structure that was characterized by positive ion MSⁿ. MS³ of the C-ion (1006.5^{–4} 831.3^{–2}) yielded an E-ion (m/z 792.3, Figure 5C) indicating that two antennae are attached to C-2 and C-4 of the 3-mannose arm; whereas, MS⁴ of the C-ion via a specific isolation of the 6-mannose arm D-ion (1006.5^{–4} 903.3^{–2} 831.3^{–2}) did not yield an E-ion (5D, 5E). Absence of the E-ion from the 6-mannose arm of Glycan (a) indicated those two antennae are attached either at the C-2 and C-6 positions (depicted in Figure 5), or at C-4 and C-6.

Characterization of Glycan (b)

Positive ion MSⁿ of permethylated Glycan (b), (Figure 2), indicated the same antennal sequence as Glycan (a), and its molecular weight (4769.7 Da), suggested a total of five antennae. Negative ion MSⁿ revealed a D-ion indicating that three antennae are on the 6-mannose arm, and the E-ion indicating two antennae are on the 3-mannose arm (Supplementary Figure 4).

Characterization of Glycan (c)

Negative ion MS² of Glycan (c), (Figure 2), produced several singly charged ions (m/z 698.4, 758.4, 818.4, 901.4 and 919.5, Figure 6A) suggesting a terminal mono-methoxyl fucose (mono-methyl- *O*-Fuc, 160 Da) within its antennae. Positive ion MSⁿ disassembly of the perdeuteromethylated = antennary B-ion (m/z 1129.8) revealed mono-*O*-methyl-fucose residues at position C-2 of the 4-*O*-methyl-GlcA, on the basis of ions at m/z 947.6, 935.6, 772.3, and 711.5 (Figure 6B). Lacking specific diagnostic ions, the location of *O*-methyl within this terminal fucose residue was not defined. Negative ion MSⁿ disassembly generated D- and E-ions to assign Glycan (c) the same branching topology as Glycan (a).

MS-based structural characterization

Structural characterization of unusual *N*-glycans has long been recognized as a challenge due to their inherent heterogeneity and potential for complexity^{49–51}. The situation is more complicated for *N*-glycans originating from less well studied non-mammalian species, where established biosynthetic rules are less applicable. For this reason, characterization of the various structural elements of EP glycoprotein *N*-glycans required an overlapping set of complementary mass spectrometry techniques. The monosaccharide identity and composition was established by GC-MS and high resolution mass measurements, distribution of the antennae was by (–) MSⁿ of the native glycans and the antennary epitope sequences was carried out by (+) MSⁿ of permethylated and deuterium permethylated derivatives. Other than glycosidic bond anomericities, this combined approach provided nearly complete characterization of the overall glycan topologies. Glycosidic bond anomericities, however, would be resolvable by applying MSⁿ spectrum matching tools if the necessary synthetic oligosaccharide substructures were obtained or made available from appropriate biological sources⁵². For example, differentiation of Gal-(1-4)-Fuc and Gal-(1-4)-Fuc was carried out by MSⁿ spectral matching in the course of describing the unusual

N-glycan Gal-Fuc moieties attached to the core GlcNAcs in *C. elegans*; these unusual structures were later implicated in cell-surface galectin binding^{53,54}. Such a comprehensive MS-based strategy has broad capabilities for the *de novo* analysis of novel oligosaccharides, combining structural specificity with low material requirements.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to acknowledge the technical assistance provided by Sherry Castle for GC-MS methods development, and Drs. Yvette Garner and Yang Yin for their respective help with mussel collection and sample purification. Importantly, we also appreciate the understanding and insight from Dr. Steven Levery for bringing to our attention related NMR work by the Prestegard laboratory (33), and the fundamental structural work of Professor Hori and colleagues (34,35). Financial assistance was provided by NIH grants RO1 GM54045 (VNR), RO1 GM20194 (NDC) and Glycan Connections, LLC.

Abbreviations

CID	collision-induced dissociation
EI	electron impact
EP	extrapallial
Fuc	fucose
Gal	galactose
GC-MS	gas chromatography-mass spectrometry
GlcA	glucuronic acid
GlcNAc	<i>N</i> -acetyl-glucosamine HRG, histidine-rich glycoprotein
IT	Ion trap
MS	mass spectrometry
SPE	solid phase extraction

References

1. Marin F, Luquet G, Marie B, Medakovic D. Molluscan shell proteins: primary structure, origin, and evolution. *Current topics in developmental biology*. 2008; 80:209–76. [PubMed: 17950376]
2. Colfen H. Biomineralization: A crystal-clear view. *Nature materials*. 2010; 9(12):960–1.
3. Zuykov M, Pelletier E, Demers S. Colloidal complexed silver and silver nanoparticles in extrapallial fluid of *Mytilus edulis*. *Marine environmental research*. 2011; 71(1):17–21. [PubMed: 20950850]
4. Muller, WEG.; Xie, L-p; Zhu, F-j; Zhou, Y-j; Yang, C.; Zhang, R-q. *Molecular Biomineralization*. Springer; Berlin Heidelberg: 2011. *Molecular Approaches to Understand Biomineralization of Shell Nacreous Layer*; p. 331–352.
5. Binbin, Hu; ZXaZD. The Biomimetic Mineralization Closer to a Real Biomineralization. In: Cavrak, PM., editor. *Advances in Biomimetics*. 2011.
6. Ma Z, Huang J, Sun J, Wang G, Li C, Xie L, Zhang R. A novel extrapallial fluid protein controls the morphology of nacre lamellae in the pearl oyster, *Pinctada fucata*. *The Journal of biological chemistry*. 2007; 282(32):23253–63. [PubMed: 17558025]
7. Misogianes MJ, Chasteen ND. A chemical and spectral characterization of the extrapallial fluid of *Mytilus edulis*. *Analytical biochemistry*. 1979; 100(2):324–34. [PubMed: 230749]

8. Hattan SJ, Laue TM, Chasteen ND. Purification and characterization of a novel calcium-binding protein from the extrapallial fluid of the mollusc, *Mytilus edulis*. The Journal of biological chemistry. 2001; 276(6):4461–8. [PubMed: 11084027]
9. Yin Y, Huang J, Paine ML, Reinhold VN, Chasteen ND. Structural characterization of the major extrapallial fluid protein of the mollusc *Mytilus edulis*: implications for function. Biochemistry. 2005; 44(31):10720–31. [PubMed: 16060681]
10. Lopes-Lima M, Ribeiro I, Pinto RA, Machado J. Isolation, purification and characterization of glycosaminoglycans in the fluids of the mollusc *Anodonta cygnea*. Comparative biochemistry and physiology Part A, Molecular & integrative physiology. 2005; 141(3):319–26.
11. Moura G, Vilarinho L, Santos AC, Machado J. Organic compounds in the extrapallial fluid and haemolymph of *Anodonta cygnea* (L.) with emphasis on the seasonal biomineralization process. Comparative biochemistry and physiology Part B, Biochemistry & molecular biology. 2000; 125(3):293–306.
12. Moura G, Vilarinho L, Machado J. The action of Cd, Cu, Cr, Zn, and Pb on fluid composition of *Anodonta cygnea* (L.): organic components. Comparative biochemistry and physiology Part B, Biochemistry & molecular biology. 2000; 127(1):105–12.
13. Nair PS, Robinson WE. Calcium speciation and exchange between blood and extrapallial fluid of the quahog *Mercenaria mercenaria* (L). Biol Bull. 1998; 195(1):43–51. [PubMed: 9739548]
14. Nair PS, Robinson WE. Purification and characterization of a histidine-rich glycoprotein that binds cadmium from the blood plasma of the bivalve *Mytilus edulis*. Arch Biochem Biophys. 1999; 366(1):8–14. [PubMed: 10334857]
15. Abebe AT, Devoid SJ, Sugumaran M, Etter R, Robinson WE. Identification and quantification of histidine-rich glycoprotein (HRG) in the blood plasma of six marine bivalves. Comp Biochem Physiol B Biochem Mol Biol. 2007; 147(1):74–81. [PubMed: 17276716]
16. Devoid SJ, Etter R, Sugumaran M, Wallace GT, Robinson WE. Histidine-rich glycoprotein from the hemolymph of the marine mussel *Mytilus edulis* L. binds Class A, Class B, and borderline metals. Environ Toxicol Chem. 2007; 26(5):872–7. [PubMed: 17521131]
17. Nair PS, Robinson WE. Cadmium speciation and transport in the blood of the bivalve *Mytilus edulis*. Mar Environ Res. 2000; 50(1–5):99–102. [PubMed: 11460758]
18. Satish Nair P, Robinson WE. Histidine-rich glycoprotein in the blood of the bivalve *Mytilus edulis*: role in cadmium speciation and cadmium transfer to the kidney. Aquat Toxicol. 2001; 52(2):133–42. [PubMed: 11164535]
19. Nair PS, Robinson WE. Cadmium binding to a histidine-rich glycoprotein from marine mussel blood plasma: potentiometric titration and equilibrium speciation modeling. Environ Toxicol Chem. 2001; 20(7):1596–604. [PubMed: 11434304]
20. Ashline D, Singh S, Hanneman A, Reinhold V. Congruent strategies for carbohydrate sequencing. 1. Mining structural details by MSn. Anal Chem. 2005; 77(19):6250–62. [PubMed: 16194086]
21. Gutternigg M, Ahrer K, Grabher-Meier H, Burgmayr S, Staudacher E. Neutral N-glycans of the gastropod *Arion lusitanicus*. Eur J Biochem. 2004; 271(7):1348–56. [PubMed: 15030485]
22. Puanglarp N, Oxley D, Currie GJ, Bacic A, Craik DJ, Yellowlees D. Structure of the N-linked oligosaccharides from tridacnin, a lectin found in the haemolymph of the giant clam *Hippopus hippopus*. Eur J Biochem. 1995; 232(3):873–80. [PubMed: 7588729]
23. Khoo KH, Maizels RM, Page AP, Taylor GW, Rendell NB, Dell A. Characterization of nematode glycoproteins: the major O-glycans of *Toxocara* excretory-secretory antigens are O-methylated trisaccharides. Glycobiology. 1991; 1(2):163–71. [PubMed: 1823159]
24. Huang Y, Konse T, Mechref Y, Novotny MV. Matrix-assisted laser desorption/ionization mass spectrometry compatible beta-elimination of O-linked oligosaccharides. Rapid Commun Mass Spectrom. 2002; 16(12):1199–204. 2002. [PubMed: 12112272]
25. Shimizu Y, Nakata M, Kuroda Y, Tsutsumi F, Kojima N, Mizuochi T. Rapid and simple preparation of N-linked oligosaccharides by cellulose-column chromatography. Carbohydr Res. 2001; 332(4):381–8. [PubMed: 11438095]
26. Ciucanu I, Costello CE. Elimination of oxidative degradation during the per-O-methylation of carbohydrates. J Am Chem Soc. 2003; 125(52):16213–9. [PubMed: 14692762]

27. Mechref Y, Kang P, Novotny MV. Solid-phase permethylation for glycomic analysis. *Methods Mol Biol.* 2009; 534:53–64. [PubMed: 19277536]
28. Jiao J, Zhang H, Reinhold VN. High Performance IT-MS Sequencing of Glycans (Spatial Resolution of Ovalbumin Isomers). *Int J Mass Spectrom.* 2011 Jun 1; 303(2–3):109–117. [PubMed: 21686090]
29. Sweeley CC, Bentley R, Makita M, Wells WW. Gas-Liquid Chromatography of Trimethylsilyl Derivatives of Sugars and Related Substances. *J Amer Chem Soc.* 1963:85.
30. Becker DJ, Lowe JB. Fucose: biosynthesis and biological function in mammals. *Glycobiology.* 2003; 13(7):41R–53R.
31. Scarsdale JN, Prestegard JH, Ando S, Hori T, Yu RK. 1H-2D-nuclear magnetic resonance applied to the primary structure determination of a novel octasaccharide glycolipid isolated from the spermatozoa of bivalves. *Carbohydr Res.* 1986; 155:45–56. [PubMed: 3791301]
32. Hori T, Sugita M, Ando S, Tsukada K, Shiota K, Tsuzuki M, Itasaka O. Isolation and characterization of a 4-O-methylglucuronic acid-containing glycosphingolipid from spermatozoa of a fresh water bivalve, *Hyriopsis schlegelii*. *J Biol Chem.* 1983; 258(4):2239–45. [PubMed: 6822554]
33. Hori T, Sugita M, Ando S, Kuwahara M, Kumauchi K, Sugie E, Itasaka O. Characterization of a novel glycosphingolipid, ceramide nonasaccharide, isolated from spermatozoa of the fresh water bivalve, *Hyriopsis schlegelii*. *The Journal of biological chemistry.* 1981; 256(21):10979–85. [PubMed: 7287744]
34. Yu RK, Koerner TA, Scarsdale JN, Prestegard JH. Elucidation of glycolipid structure by proton nuclear magnetic resonance spectroscopy. *Chemistry and physics of lipids.* 1986; 42(1–3):27–48. [PubMed: 3549018]
35. Gielens C, Idakieva K, Van den Bergh V, Siddiqui NI, Parvanova K, Compennolle F. Mass spectral evidence for N-glycans with branching on fucose in a molluscan hemocyanin. *Biochem Biophys Res Commun.* 2005; 331(2):562–70. [PubMed: 15850797]
36. Idakieva K, Stoeva S, Voelter W, Gielens C. Glycosylation of *Rapana thomasiana* hemocyanin. Comparison with other prosobranch (gastropod) hemocyanins. *Comparative biochemistry and physiology Part B, Biochemistry & molecular biology.* 2004; 138(3):221–8.
37. Stanley, P.; Schachter, H.; Taniguchi, N. N-Glycans. In: Varki, A.; Cummings, RD.; Esko, JD.; Freeze, HH.; Stanley, P.; Bertozzi, CR.; Hart, GW.; Etzler, ME., editors. *Essentials of Glycobiology.* 2. Cold Spring Harbor (NY): 2009.
38. Bleton J, Mejanelle P, Sansoulet J, Goursaud S, Tchaplal A. Characterization of neutral sugars and uronic acids after methanolysis and trimethylsilylation for recognition of plant gums. *Journal of Chromatography A.* 1996; 720:27–49.
39. Stumpo KA, Reinhold VN. The N-glycome of human plasma. *Journal of proteome research.* 2010; 9(9):4823–30. [PubMed: 20690605]
40. Ashline DJ, Lapadula AJ, Liu YH, Lin M, Grace M, Pramanik B, Reinhold VN. Carbohydrate structural isomers analyzed by sequential mass spectrometry. *Anal Chem.* 2007; 79(10):3830–42. [PubMed: 17397137]
41. Harvey DJ. Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. *J Am Soc Mass Spectrom.* 2005; 16(5):622–30. [PubMed: 15862764]
42. Harvey DJ. Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans. *J Am Soc Mass Spectrom.* 2005; 16(5):631–46. [PubMed: 15862765]
43. Harvey DJ. Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans. *J Am Soc Mass Spectrom.* 2005; 16(5):647–59. [PubMed: 15862766]
44. Harvey DJ. Collision-induced fragmentation of negative ions from N-linked glycans derivatized with 2-aminobenzoic acid. *Journal of mass spectrometry: JMS.* 2005; 40(5):642–53. [PubMed: 15751107]
45. Harvey DJ, Baruah K, Scanlan CN. Application of negative ion MS/MS to the identification of N-glycans released from carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1). *Journal of mass spectrometry: JMS.* 2009; 44(1):50–60. [PubMed: 18788072]

46. Harvey DJ, Jaeken J, Butler M, Armitage AJ, Rudd PM, Dwek RA. Fragmentation of negative ions from N-linked carbohydrates, part 4. Fragmentation of complex glycans lacking substitution on the 6-antenna. *Journal of mass spectrometry: JMS*. 2010; 45(5):528–35. [PubMed: 20446311]
47. Harvey DJ, Royle L, Radcliffe CM, Rudd PM, Dwek RA. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. *Analytical biochemistry*. 2008; 376(1):44–60. [PubMed: 18294950]
48. Ritchie G, Harvey DJ, Stroeher U, Feldmann F, Feldmann H, Wahl-Jensen V, Royle L, Dwek RA, Rudd PM. Identification of N-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionisation time-of-flight and negative ion electrospray tandem mass spectrometry. *Rapid communications in mass spectrometry: RCM*. 2010; 24(5):571–85. [PubMed: 20131323]
49. Hart GW, Copeland RJ. Glycomics hits the big time. *Cell*. 2010; 143(5):672–6. [PubMed: 21111227]
50. Bertozzi, CR.; Rabuka, D. Structural Basis of Glycan Diversity. In: Varki, A.; Cummings, RD.; Esko, JD.; Freeze, HH.; Stanley, P.; Bertozzi, CR.; Hart, GW.; Etzler, ME., editors. *Essentials of Glycobiology*. 2. Cold Spring Harbor (NY): 2009.
51. Marino K, Bones J, Kattla JJ, Rudd PM. A systematic approach to protein glycosylation analysis: a path through the maze. *Nature chemical biology*. 2010; 6(10):713–23.
52. Reinhold V, Zhang H, Hanneman A, Ashline D. Toward a Platform for Comprehensive Glycan Sequencing. *Molecular & Cellular Proteomics*. 2013; 12(4):866–873. [PubMed: 23438731]
53. Maduzia LL, Yu E, Zhang Y. *Caenorhabditis elegans* Galectins LEC-6 and LEC-10 Interact with Similar Glycoconjugates in the Intestine. *Journal of Biological Chemistry*. 286(6):4371–4381. [PubMed: 21115491]
54. Hanneman AJ, Rosa JC, Ashline D, Reinhold VN. Isomer and glycomer complexities of core GlcNAcs in *Caenorhabditis elegans*. *Glycobiology*. 2006; 16(9):874–90. [PubMed: 16769777]

(-) ESI-IT-MS

Glycan	M.W.
a.	4027.4
b.	4769.7
c.	4667.8

Figure 1.

(-) ESI-IT-MS of the native *N*-glycans released from EP glycoprotein with the three most abundant structures labeled (**a**, **b**, and **c**). The respective molecular weights are shown at top right.

Figure 2. Structures assigned to the three most abundant *N*-glycans (*a*, *b*, and *c*) of *M. edulis* EP glycoprotein.

Figure 3.

(+) MSⁿ disassembly of the antennal B-ion of permethylated **Glycan (a)** revealing the tetrasaccharide moiety: 4-*O*-methyl-GlcA(1-4)[GlcNAc(1-3)]Fuc(1-4)GlcNAc. (A) MS² of the molecular ion m/z 1240.1⁽⁺⁴⁾ distinguished the antennal B-ion (m/z 919.4) consisting of four monosaccharides (additional MS² peaks are shown in supplementary Figure 3). (B) MS³ of m/z 919.4, indicating a branched fucose residue. (C) MS⁴ of the trisaccharide ion (m/z 660.3) after neutral loss of its non-reducing end GlcNAc, and (D) MS⁵ of the disaccharide (m/z 433.2) after neutral loss of its reducing end GlcNAc, indicating terminal 4-*O*-methyl-GlcA and GlcNAc linked to C-4 and C-3 of Fuc, respectively. The ion m/z 431

(observed as m/z 443 in the CD₃ methylated derivative) may indicate dehydrogenation at C4/C5 of 4-*O*-methyl-GlcA to yield conjugation. All the ions are sodium adducted.

Figure 4.

(+) MSⁿ disassembly of the perdeuteromethylated **Glycan (a)** antennal B-ion corroborating the site of endogenous methylation at C-4. (A) MS³ of the antennal B-ion (m/z 952.6) showed a preferential loss of the natural methyl group (m/z 920.5). (B) MS⁴ of m/z 920.5 then produced a prominent ion at m/z 770.5 from Retro Diels-Alder fragmentation. All the ions are sodium adducted.

Figure 5.

(-) MSⁿ disassembly of native **Glycan** (a) revealing its branching topology. (A) Diagram highlighting the diagnostic (-) MSⁿ fragment ions. (B) MS² (m/z 1006.4)⁽⁻⁴⁾ produced the D-Ion (m/z 903.3)⁽⁻²⁾ and ^{0,3}A-ion of the central mannose (m/z 867.3)⁽⁻²⁾, indicating two antennae on the 6-mannose arm. (C) Diagram of the E-ion (m/z 792.3)⁽⁻²⁾ obtained only from the 3-mannose arm (D) The E-ion was not observed in MS⁴ along the C-ion pathway (m/z 831.3)⁽⁻²⁾ derived specifically from the 6-mannose arm, indicating these two antennae are linked to positions C-2 and C-6 or positions C-4 and C-6 of the 6-mannose arm. (E) The E-Ion was observed in the MS³ spectrum of the C-ion, indicating these two antennae are linked to C-2 and C-4 of the 3-mannose arm.

Figure 6.

Spectra showing (-) and (+) MS^n disassembly of **Glycan (c)**, revealing the pentasaccharide antennal sequence: mono-*O*-methyl-Fuc1 2(4-*O*-methyl-)GlcA1 4 [GlcNAc 1 3]Fuc 1 4 GlcNAc. (A) (-) MS^2 spectrum of native **Glycan (c)**. Two ions not labeled in the diagram (m/z 991.5²⁻ and 982.5²⁻) represent C-type cleavages between the 3-mannose or 6-mannose arms to the central mannose including two antennae and one branched mannose. (B) (+) MS^3 of the deuterium permethylated (Na^+ adducted) antennal B-ion.