

Review

De Novo Structural Determination of the Oligosaccharide Structure of Hemocyanins from Molluscs

Pavlina Dolashka ^{1,*}, Asya Daskalova ¹, Aleksandar Dolashki ¹ and Wolfgang Voelter ²

¹ Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria; asya.daskalova@gmail.com or adaskalova@orgchm.bas.bg (A.D.); adolashki@orgchm.bas.bg (A.D.)

² Interfaculty Institute of Biochemistry, University of Tuebingen, 72074 Tuebingen, Germany; wolfgang.voelter@uni-tuebingen.de

* Correspondence: dolashka@orgchm.bas.bg or pda54@abv.bg; Tel.: +359-887193423

Received: 26 August 2020; Accepted: 23 September 2020; Published: 22 October 2020

Abstract: A number of studies have shown that glycosylation of proteins plays diverse functions in the lives of organisms, has crucial biological and physiological roles in pathogen–host interactions, and is involved in a large number of biological events in the immune system, and in virus and bacteria recognition. The large amount of scientific interest in glycoproteins of molluscan hemocyanins is due not only to their complex quaternary structures, but also to the great diversity of their oligosaccharide structures with a high carbohydrate content (2–9%). This great variety is due to their specific monosaccharide composition and different side chain composition. The determination of glycans and glycopeptides was performed with the most commonly used methods for the analysis of biomolecules, including peptides and proteins, including Matrix Assisted Laser Desorption/Ionisation–Time of Flight (MALDI-TOF-TOF), Liquid Chromatography - Electrospray Ionization-Mass Spectrometry (LC/ESI-MS), Liquid Chromatography (LC-Q-trap-MS/MS) or Nano- Electrospray Ionization-Mass Spectrometry (nano-ESI-MS) and others. The molluscan hemocyanins have complex carbohydrate structures with predominant *N*-linked glycans. Of interest are identified structures with methylated hexoses and xyloses arranged at different positions in the carbohydrate moieties of molluscan hemocyanins. Novel acidic glycan structures with specific glycosylation positions, e.g., hemocyanins that enable a deeper insight into the glycosylation process, were observed in *Rapana venosa*, *Helix lucorum*, and *Haliotis tuberculata*. Recent studies demonstrate that glycosylation plays a crucial physiological role in the immunostimulatory and therapeutic effect of glycoproteins. The remarkable diversity of hemocyanin glycan content is an important feature of their immune function and provides a new concept in the antibody–antigen interaction through clustered carbohydrate epitopes.

Keywords: hemocyanins; molluscs; mass spectrometric analysis

1. Introduction

Glycoproteins are of great importance for the optimal and proper functioning of many bioactivities in the human body. Absence or alteration of glycoproteins can cause a wide range of diseases, e.g., the glycosylated form of haemoglobin is responsible for the onset of diabetes [1]. Glycans also play an important role in cardiovascular disease, microbial and viral pathogenesis, development of tumors, tissue repair, strengthening of the immune system, and others [2–4]. As demonstrated by various studies [5,6], hemocyanins have gained a large amount scientific interest due to their interesting glycoprotein structures.

Hemocyanins are type-3 copper-binding proteins found freely dissolved in the hemolymph in two main phyla (mollusca and arthropoda), which have the same physiological function, but show significant differences in their structural organization and carbohydrate content.

Molluscan hemocyanins are among the largest known glycoproteins with huge cylindrical multimeric forms and molecular masses ranging from 3.3 to 13.5 MDa. They form decamers or multi-decamers of 330- to 550-kDa subunits organised by more than seven different functional units (FUs) with molecular masses ranging from 45 to 65 kDa. Structural subunits assemble to di-decamers as presented for keyhole limpet hemocyanin (KLH1) in (Figure 1) [7].

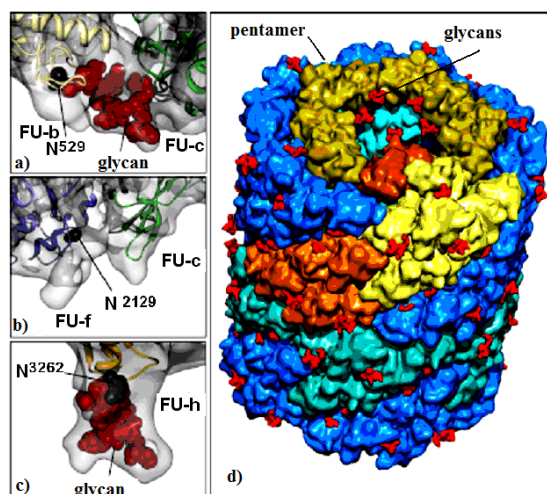


Figure 1. KLH1 subunit model. Location of Asp residue in the glycosylation site of: (a) FU-b and FU-c; (b) FU-f and FU-c; (c) FU-h; (d) di-decamer KLH1 model and the position of 120 glycans (marked in red) [7].

Their structure has been investigated using a combination of electron cryo-microscopy and X-ray crystallography (Figure 2). The represented X-ray crystal structure of the intact 3.8-MDa molecule of *Todarodes pacificus* squamous hemocyanin (TpH) shows the complex oligosaccharide structures of the glycoprotein [8]. Chemical structure analysis of *Todarodes pacificus* hemocyanin revealed that the two most abundant sequences, HexNAcMan₃GlcNAc₂ and HexHexNAcMan₃GlcNAc₂, accounted for more than 95% of all oligosaccharides tested in all individuals. The Cu₂O₂ cluster and the carbohydrates of functional unit “d” of TpH are represented in the 3D-model of Figure 2. The authors demonstrated the influence of glycans on the dissociation and re-association behavior of the decameric forms of TpH [9].

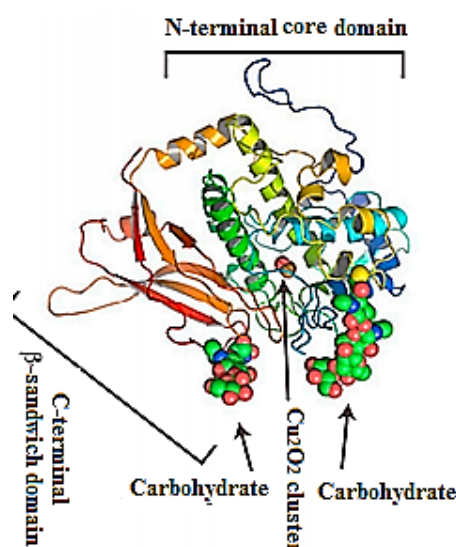


Figure 2. Representative ribbon diagram of FU-d of *Todarodes pacificus* squamous hemocyanin, colored as a ramp from blue (N-terminus) to red (C-terminus). The Cu₂O₂ cluster and carbohydrates [9].

Therefore, the great scientific interest in molluscan hemocyanins is due not only to their complex quaternary structures, but also to their great diversity in oligosaccharide structures, high carbohydrate content (2–9%), specific monosaccharide composition [10–13], and their well-established link with the immunostimulatory effect [14,15].

The studies on carbohydrate structures of hemocyanins from molluscs occupy an important part of this review, demonstrating significant differences in carbohydrate structures of molluscan hemocyanins.

2. Carbohydrate Structures of Hemocyanins

The first analyses of the carbohydrate structures of hemocyanins from the earth snail *Helix pomatia* (HpH) [10], freshwater snail *Lymnaea stagnalis* (LsH) [16], and the California mussel (keyhole limpet) *Megathura crenulata* (KLH) [13,17] present complex carbohydrate structures composed predominantly of the monosaccharides Xyl, Fuc, Man, Gal, GalNAc, and GlcNAc. Along with common monosaccharides, other monosaccharides such as Xyl and O-methyl-Gal have also been identified in HpH glycans [10,11,18]. Though the Xyl residue is less common in glycoproteins from animal organisms, their presence is essential for immunogenic effects [10].

Other specifically-modified structures have been identified that bind (β1-2)Xyl residues to β-Man and α-Fuc to 1,6-GlcNAc from the core of the glycan. Traces of large neutral N-glycans ending with 3-O-MeGal, Xyl, and/or Fuc monosaccharides have been found in *Planorbarius corneus*, *Achatina fulica*, and *Arion lusitanicus* [12,19]. The presented complete N-glycan spectrum of the slug *Arion lusitanicus* includes a combination of carbohydrate structures found in mammals, plants, insects, and the like [12,19]. Complex structures have also been published for some glycans of *Rapana thomasiana* hemocyanin (RtH), later renamed *Rapana venosa* (RvH) [11,20].

Furthermore, the studies on the carbohydrate structure of KLH, used in medicine, present two new N-linked glycans, Fuc(α1-3)GalNAc(β1-4)[Fuc(α1-3)]GlcNAc and Gal(β1-4)Gal(β1-4)Fuc(α1-6), as part of the huge protein molecule [13,17]. In one glycan, a Gal monosaccharide is linked to the final GalNAc, for which the clinical success of KLH in the treatment of bladder cancer is found. The carbohydrate chain Gal(β1-4)Gal(β1-4)Fuc(α1-6) is linked to the nucleus of another glycan, which also contributes to the use of KLH in immunotherapy.

The structural diversity of N-glycans in snails has been investigated very thoroughly by a number of authors who report predominantly high-mannose carbohydrate chains. Exceptions were found for some carbohydrate structures that bind sialic acid, α1-6- and α1-3-Fuc, β1-2-Xyl, MeMan,

MeGal, GlcNAc, GalNAc, and other monosaccharides, but significantly fewer glycans end up with a 3-O-MeMan residue [10,11,18].

A great variety of oligosaccharide structures has also shown for side chains linked to carbohydrate structures of hemocyanins from the abdominal mollusc *Lymnaea stagnalis*, the California mussel KLH, and others [12,14,16,19,21]. Specific enzymes, responsible for the formation of these complex structures, have been found in various organs and tissues of *Limax maximus*, *Lymnaea stagnalis*, and other representatives of the molluscs [12].

Important information for the positions of glycosylated Asn residues for three FU-s (FU-b, FU-f and FU-h) is provided by the di-decameric model of KLH1 (Figure 1a–d). The model shows 120 potential *N*-glycan motifs (NXT and/or NXS), located mainly on the surface of the molecule (Figure 1d). FU-c is of special interest due to its absence of a residual glycan (Figure 1b) [22]. This corresponds to KLH-c with only one *O*-glycosylation site, lacking an attached, glycan moiety [23]. The absence of glycan in FU-c was also observed in hemocyanins of *Haliothis tuberculata* (HtH), *Octopus dofleini* (OdH), and *Aplysia californica* (AcH) [24]. Published literature mostly presents *O*-glycosylated hemocyanins of the arthropodan species, while very few *O*-linked glycans have been found in hemocyanins of the Molluscs.

Another interesting feature for molluscan and arthropodan hemocyanins is the position of the putative glycosylation sites, located mainly on the surface of the molecule. Two *N*-linked sites of FU HtH1-h are located at rather unusual positions and are not found in other FUs, including HtH2-h. These glycans probably perform an important role related to the function of glycosylated FU-h, suggesting that they prevent binding of decamers to di-decamers and formation of multi-decamers in hemocyanins. The assumption is confirmed by the established absence of multi-decamers in the hemocyanins of AcH and HtH1 [24,25].

In this review, the oligosaccharide structures of three hemocyanins are presented: from the Black Sea snail *Rapana venosa*, the garden snail *Helix lucorum*, and the abalone *Haliothis tuberculata*, which normally exist under different living conditions. They have been compared with the well-studied oligosaccharide structure of keyhole limpet hemocyanin (KLH) from a mussel inhabiting the northern coast of America, and other hemocyanins with complex carbohydrate structures.

Mass spectral methods and techniques are appropriate for analysis of glycoproteins and identification of their structure. To determine molecular masses of amino acid sequences (AAS) of proteins and peptides, mass spectrometry is used as the nucleus of proteomics. It is one of the most widely used methods for the analysis of biomolecules like peptides, proteins, and glycoproteins, with a series of sophisticated instrumentations such as Matrix Assisted Laser Desorption/Ionisation–Time of Flight (MALDI-TOF-TOF), Liquid Chromatography-Electrospray Ionization-Mass Spectrometry (LC/ESI-MS), Liquid Chromatography (LC-Q-trap-MS/MS) or Nano- Electrospray Ionization-Mass Spectrometry (nano-ESI-MS).

3. Carbohydrate Structure of Hemocyanins from the Marine snail *Rapana venosa*

Recent research reported on the glycosylated nature of RvH, with a carbohydrate content of 8.9%, which differs for the two subunits RvH1 (12.4%) and RvH2 (4.4%) [26,27]. A new strategy for analysis and determination of the carbohydrate structure of subunits and FUs of RvH is presented here (Figure 3).

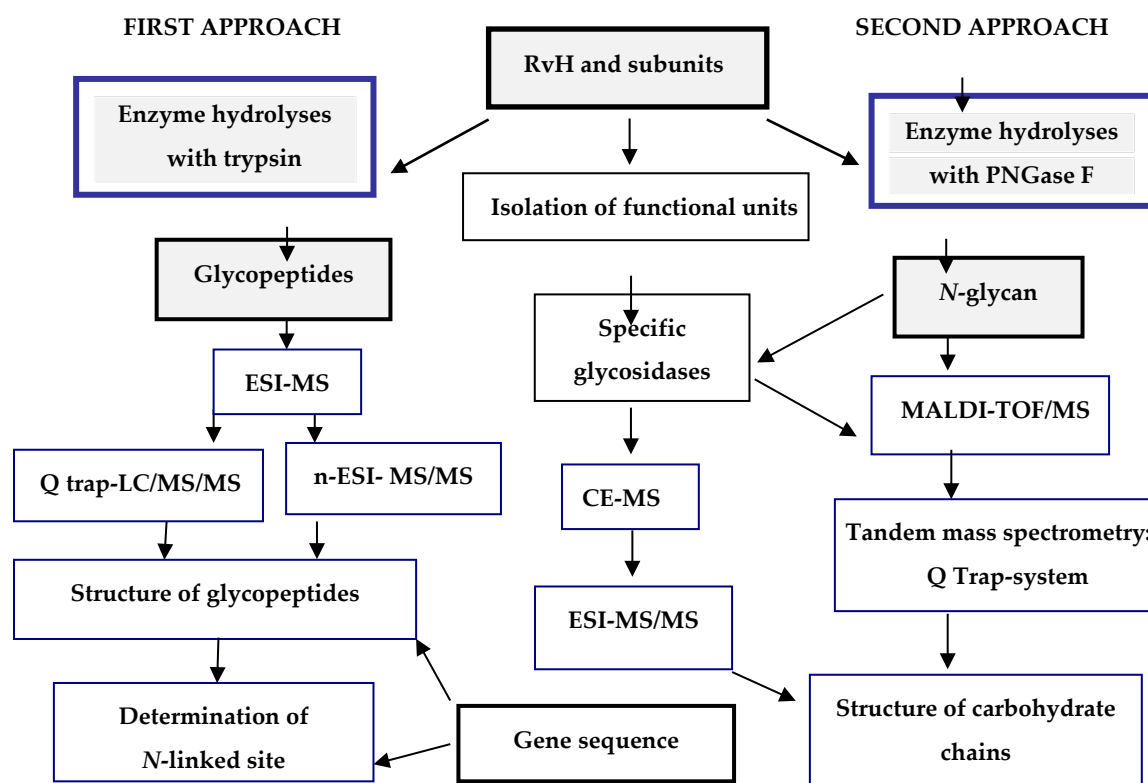


Figure 3. Scheme for analysis and determination of the carbohydrate structure of RvH1, RvH2 and FUs. On the basis of this strategy, two main approaches are involved.

In the first approach, functional units (FUs) of the two subunits RvH1 and RvH2 are treated with trypsin and the obtained glycopeptides are identified through LC/ESI-MS, LC-Q-trap-MS/MS or nano-ESI-MS [28–33]. The second approach includes isolation of glycans after hydrolysis of the subunits of RvH with specific glycosidases (PNGase F) and analysis of the glycan structures with MALDI-TOF-MS, CE-MS, and Q-trap instrumental systems [28,29,34].

3.1. First Approach for Determination of the Carbohydrate Structure of RvH

The developed strategy is suitable for the determination of the complex carbohydrate structures of molluscan hemocyanins. The first approach of the developed and implemented strategy provides information about the oligosaccharide structures only for the RvH.

3.1.1. Isolation and Characterization of Glycopeptides of FUs from RvH

Three FUs (RvH1-a, RvH1-f, and RvH1-b) are isolated after trypsinolysis of RvH1 and the resulting fragments are separated on a Superdex 300 column. The positive orcinol/H₂SO₄ test allows one to detect two (Glp1 and Glp2) glycopeptides in RvH1-a, but only one (Glp3) in RvH1-f. Confirmation of the complex carbohydrate structures of the three glycopeptides (Glp1, Glp2, and Glp3) is presented after treatment with various glycosidases and analysis with CE-MS and ESI-MS.

Detailed data for the structures of three glycopeptides from RvH1 (Glp1, Glp2, and Glp 3) are received after treatment with specific glycosidases and analysis of cleaved products with CE-MS and ESI-MS. That allows the identification of complex bi-antennary structures of the glycans including methylated galactoses bound to Glp1 and Glp2 differing from the structure of the glycopeptide Glp3 of RvH1-f [28,29].

With these published data on the carbohydrate structures of molluscan hemocyanins, methylated monosaccharides are mentioned, as found before, as constituents in seven related complex structures, constituents of the hemocyanins from *Lymnaea stagnalis* [16].

3.1.2. Structures of Glycopeptides from RvH, Determined through LC/ESI-MS, Nano-ESI-MS and Q-trap LC/MS/MS Systems

After trypsinolysis of the two subunits, RvH1 and RvH2, and separation of fragments on a Nucleosil 7 C18 column, mass spectrometric methods were applied for structure determination of glycopeptides [28]. The presented orcinol /H₂SO₄ test (Figure 4) gives positive reactions only for 58 eluted fractions from the column. The oligosaccharide structures of the glycopeptides were determined through tandem-MS after tracking of the fragmented ions from the MS/MS data of the double-charged ion [M+2H]²⁺ and scanning of the pseudo-MS/MS/MS.

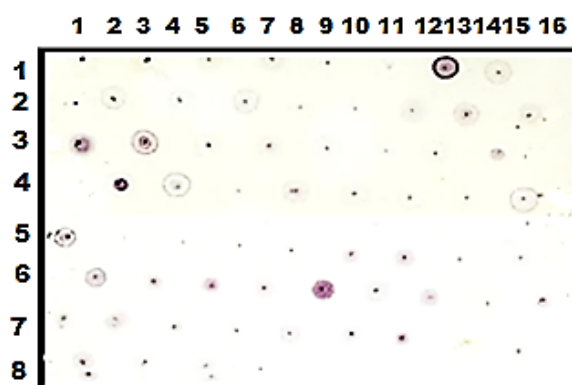


Figure 4. Orcinol /H₂SO₄ test of the obtained fractions of RvH1 after trypsinolysis.

The summary of the results obtained from μ LC/ESI-MS analysis after tryptic hydrolysis of RvH1 provide additional confirmation for the glycosylation of RvH1: three glycopeptides were identified with Mw 2745.5 Da, 2421.0 Da, and 2907.1 Da, and oligosaccharide structures of the high-mannose type [26,35]. Data for the carbohydrate structure of RvH was completed with the comprehensive nano-ESI-MS analysis of isolated RvH1 and RvH2 glycopeptides after trypsinolysis. This methodology gives simultaneous access to fundamental structural characteristics of the glycoproteins, such as their glycan structure, amino acid sequences (AAS), and binding sites of glycans to the peptide chains [31]. The structure of the glycopeptides are determined after the interpretation of the Y- and B-ions represented in the spectrum, giving information about the calculated glycopeptide, with Mw 2676.32 Da, and ion, with m/z 1338.16 [M+2H]²⁺ (Figure 5). Another important feature from the MS/MS spectrum is the bound glycan (GlcNAc residue) at the ion at m/z 1661.71 to the N-glycosylated peptide (R-) at m/z 1459.75 [M+H]⁺. The glycan-typical oxon ions [M+H]⁺ with m/z 204.08 (HexNAc) and m/z 366.15 (HexHexNAc) in the MS/MS spectrum show the exact structure of the glycan and the bound peptide. The glycopeptide characteristic is complemented by the represented different single fragment [M+H]⁺ glycan ions with m/z 528.20 (Hex₂HexNAc), m/z 690.27 (Hex₃HexNAc), m/z 852.31 (Hex₄HexNAc), and m/z 1014.46 (Hex₅HexNAc) (Figure 5).

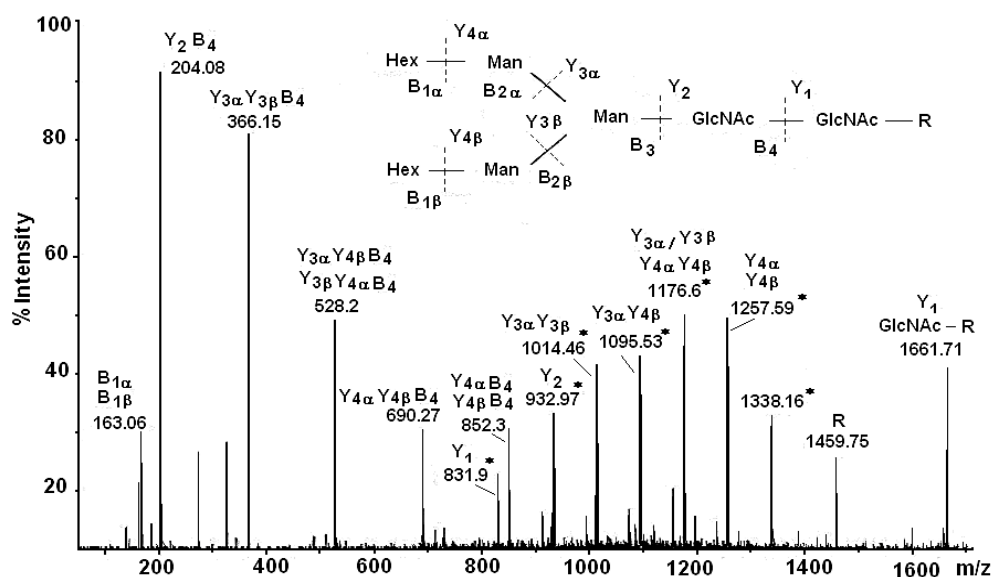


Figure 5. Nano-ESI-MS/MS analysis of the glycan G1, bound to glycopeptides, presented as an ion at m/z 1338.16 $[M+2H]^{2+}$ [31].

The results of the nano-ESI-MS analysis of the glycopeptide with Mw 2676.3 Da confirm the highly mannose-like carbohydrate structure of RvH [31]. The assumption of the carbohydrate structure of RvH1 was confirmed by the combined analysis of glycopeptides with LC/MS and the Q-trap system. The resulting LC/MS/MS chromatogram of the eluted fraction at 31.24 min at high scanning resolution expressed a dominant ion with m/z 837.97 $[M+3H]^{3+}$, corresponding to a glycopeptide with Mw 2511.91 Da. The fragment Y- and B-ions with m/z 204.1, 366.2, 528.3, 690.5, and 852.4 determine the carbohydrate structure as Hex₂Man₂GlcNAc₂ of a glycan with Mw 1054.0 Da. The other fragmented y- and b-ions reflect the amino acid sequence AAS of the peptide MGQYGD (I/L) STNNTR with two potential glycosylation sites (–D(L/I) S- and –NNT–) (Figure 6). According to LC/MS and Q-trap analysis, only one glycosylation site (–NNT) in a glycopeptide with Mw 2511.91 Da, to which a high mannose glycan Hex₂Man₂GlcNAc₂ with Mw 1054.0 Da is linked, can be detected [31].

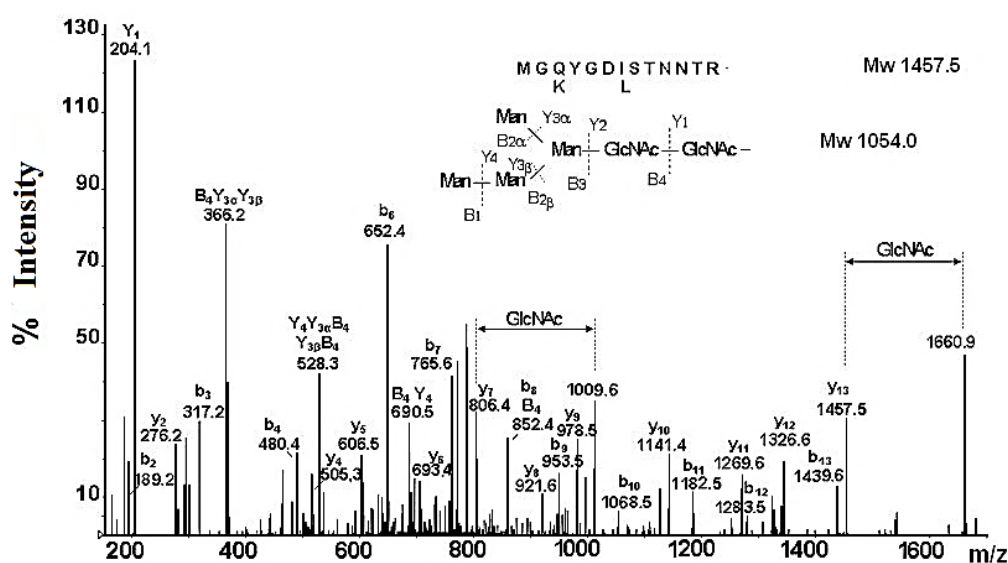


Figure 6. Enhanced product ion (EPI) - scanning of an ion with m/z 837.97 $[M+3H]^{3+}$ with a higher resolution.

3.2. Second Approach for Determination of Carbohydrate Structure of RvH

Interpretation of the structure of glycopeptides is often difficult due to the overloaded MS spectra via fragmented ions of both the peptide (y- and b-ions) and the glycan (Y- and B-ions). This disadvantage is solved by a second approach. After treatment with the specific glycosidase PNGase, F glycans are removed from the subunits RvH1 and RvH2 [32,33], eluted with 2 mL 25% ACN/ 0.05% TFA through a Carbograph column and finally analysed by CE and MALDI-TOF, presenting the high mannose character of glycan structures including Gal, Fuc and Xyl in the carbohydrate chains of RvH1 and RvH2 [32].

3.2.1. Analysis of N-glycans from RvH by Capillary Electrophoresis Combined with MS Analysis

The approach of CE with MS allows structure identification of glycans in very low concentration mixtures. Moreover, using this method, a new type of structure was presented for the ion at m/z 555.7 $[M-3H]^{3-}$ of the glycan from RvH1, which was different from structures published by other authors. Giles et al. 2005 published one MeHex residue connected to an inner Fuc residue (Figure 7A, B) [33].

The analyses of CE-MS/MS show the complex structures of the glycans of RvH1 presented by ions at m/z 555.7 $[M-3H]^{3-}$ and two isomeric forms of the glycan with identical MS/MS spectra. The identical MS/MS spectra demonstrate essential divergences in the structure of glycan in connection to the GlcNAc residue at α -1,3 or α -1,6-positions in the trisaccharide of the glycan. Moreover, this structure resembles the published structure from Gielens and co-authors with an internal GalNAc residue and MeHex [11] (Figure 7A).

The used methods cannot differentiate the two structures of MeHex and HexA because of the small difference in the Mw from 0.036 Da. This problem was solved by CE-MS/MS analyses, which determined the methylated structure of the oligosaccharide, whose behavior does not correspond to the glycan at m/z 555.7. Larger glycans will migrate faster than smaller ones, and charged oligosaccharides will migrate slower than their neutral counterparts. A glycan, detected at m/z 555.7 as a 4-fold negatively charged ion, migrates slower than glycan, which is smaller and detected at m/z 557.8 as 3 times negatively charged ion.

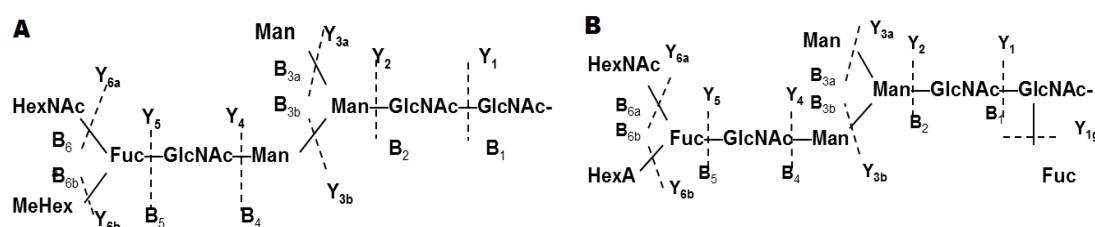


Figure 7. The complex structures presented for the ion at m/z 555.7 $[M-3H]^{3-}$ of RvH1 analyzed by CE-MS/MS and MS/MS spectra: (A) Structure presented by Giles et al. 2005 [11]; (B) Structure presented by Sandra et al. 2007 [33].

Based on this, we can conclude that the ion at m/z 555.7 corresponds to a charged N-glycan and a complex glycan with an internal fucose connecting a hexuronic acid (HexA) and an N-acetylhexosamine (HexNAc). Similar structures with HexA and Fuc residues are suggested from MS/MS analyses and the observed electrophoretic offset of the three ions $[M-4H]^{4-}$ in RvH. The specific structures with HexA in RvH1 are most possibly a result of the oxidizing environment [33].

3.2.2. Analysis of Glycans from RvH by 3-AP- and APTS-Labeled Glycans

A new combined approach of CE-MS and MS/MS analyses of labeled glycans of RvH1 with positively charged 3-aminopyrazole (3-AP) is suitable to determine glycan structures at very low concentrations in mixtures. A comparative analysis of the behaviour of the 3-AP-labeled glycans shows a shift of the elution time in the chromatogram [33].

Another important piece of information is the structural differences of glycans with identical isomer MS/MS spectra, which is probably due to the acid tetrasaccharide attached to the α -1,3 or α -

1,6-arm positions of the trimannose core. Proof for this is the reported difference from the additional Fuc residue in the MS/MS spectra of the ion with m/z 853.7 $[M+2H]^{2+}$ (Figure 8A) and with m/z 926.8 $[M+2H]^{2+}$ (Figure 8B).

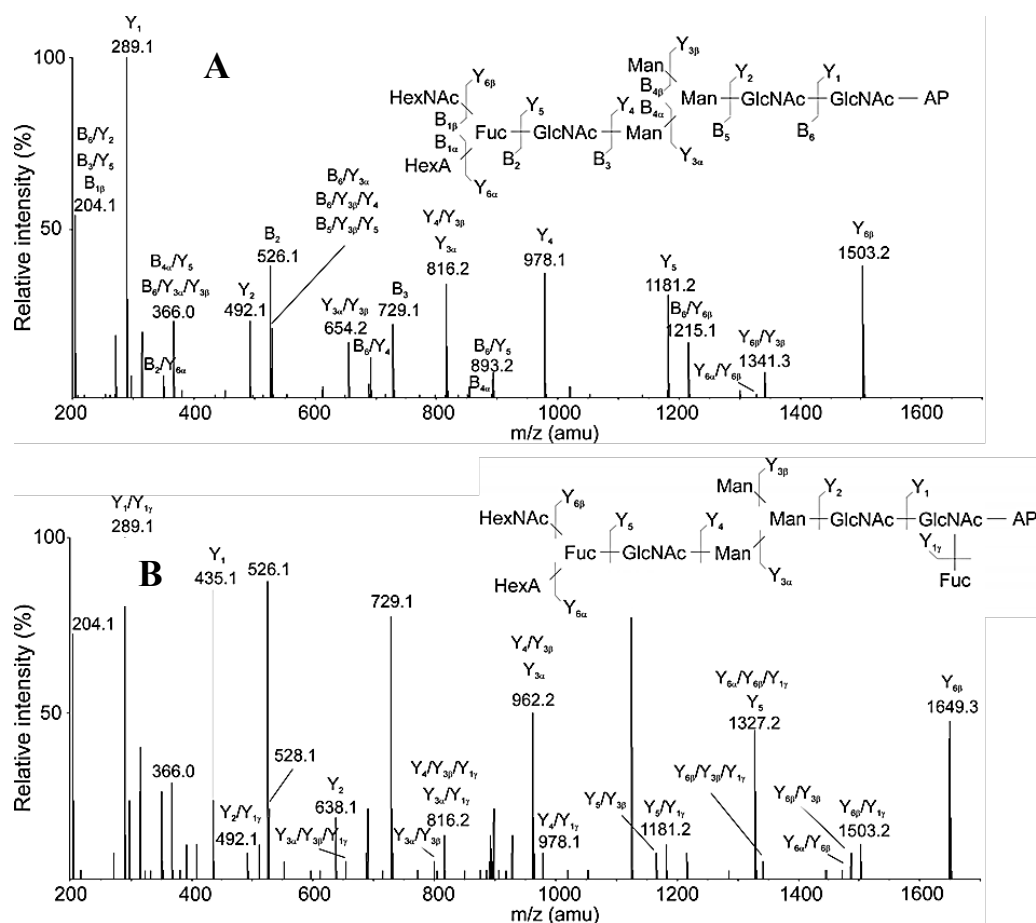


Figure 8. (A). MS/MS analysis of 3-AP-glycan from RvH with m/z 853.7 $[M+2H]^{2+}$ eluted at 14.87 min. (B) MS/MS analysis of 3-AP-glycan from RvH with m/z 926.8 $[M+2H]^{2+}$, eluting at 14.87 min.

Confirmation of the position of the triple moiety (HexNAcFucHexA) is ion B_2 with m/z 526.1 $[M+H]^+$, indicating the loss of the Fuc residue after separation from the HexNAc and HexA residue. The advantage of the applied approaches is the identification of the complex structures of two glycans FucHexAHexNAcGlcNAcMan₃GlcNAc₂ and Fuc₂HexAHexNAcGlcNAcMan₃GlcNAc₂, respectively.

3.2.3. Determination of Hexuronic Acid in the Structure of RvH1 after Amidation and Permethylaton of the Glycans

This new method is based on the modification of the carboxyl groups of the glycan with NH_4Cl and DMT-MM [4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride], whereby the carboxylic acid is converted to an amide, leading to a decrease in the Mw of acidic glycans by 0.9840 Da, identifying a bound HexA moiety of the glycan [33]. MALDI-MS analysis reported a decrease in Mw for eight N-linked glycans in RvH1 after amidation, demonstrating the involvement of eight HexA in glycans of RvH1. The conversion of HexA to MeHex shows an increase in Mw by 14 Da after permethylation of the four identified glycans, noted in Table 1 as 14, 17, 19 and 21 [32,33].

Table 1. Composition of N-glycans in RvH1. **(a)** N-glycans are represented as [M+Na]⁺ ions. Mw of permethylated glycans is indicated in parentheses. **(b)** *m/z* values of APTS-labeled glycans [33].

№	Structure	MALDI [M-H] ⁺	№	Structure	MALDI [M-H] ⁺
1	FucMan ₂ GlcNAc ₂	917.3	12	Man ₇ GlcNAc ₂	1581.7
2	Man ₃ GlcNAc ₂	933.3	13	Fuc ₂ HexNAc ₂ Man ₃ GlcNAc ₂	1631.8
3	FucMan ₃ GlcNAc ₂	1079.5	14	FucHexAHexNAc ₂ Man ₃ GlcNAc ₂	1661.7
4	Man ₄ GlcNAc ₂	1095.5	15	FucHexNAc ₃ Man ₃ GlcNAc ₂	1688.8
5	HexNAcMan ₃ GlcNAc ₂	1136.5	16	Man ₈ GlcNAc ₂	1743.7
6	FucMan ₄ GlcNAc ₂	1241.6	17	Fuc ₂ HexAHexNAc ₂ Man ₃ GlcNAc ₂	1807.8
7	Man ₅ GlcNAc ₂	1257.6	18	Fuc ₂ HexNAc ₃ Man ₃ GlcNAc ₂	1834.
8	FucHexNAcMan ₃ GlcNAc ₂	1282.6	19	FucHexAHexNAc ₃ Man ₃ GlcNAc ₂	1864.8
9	HexNAc ₂ Man ₃ GlcNAc ₂	1339.6	20	Man ₉ GlcNAc ₂	1905.8
10	Man ₆ GlcNAc ₂	1419.6	21	Fuc ₂ HexAHexNAc ₃ Man ₃ GlcNAc ₂	2010.8
11	FucHexNAc ₂ Man ₃ GlcNAc ₂	1485.7			-

Another important method for determination of complex structures of the glycans in RvH is Q-trap analyses of permethylated glycans. The presented MS/MS spectra of structures with tetrasaccharide are expressed by two ions with m/z 660.3 [M+Na]⁺(B₃/Y_{6β}) and 919.4 [M+Na]⁺(B₃) (Figure 9). The position of the Fuc residue in the chain was confirmed by additional MS³ analysis of the permethylated B₃ ion. The reported change in Mw by 160 Da [145+14] probably reflects methylation of only one hydroxyl group of the glycan because the other three hydroxyl groups participate in other bonds. Evidence of acidic carbohydrate structures is also presented in the RvH2 subunit. They are reported based on the change in molecular masses with 1 D of two glycans from the glycan mixture of RvH2 before and after amidation [32].

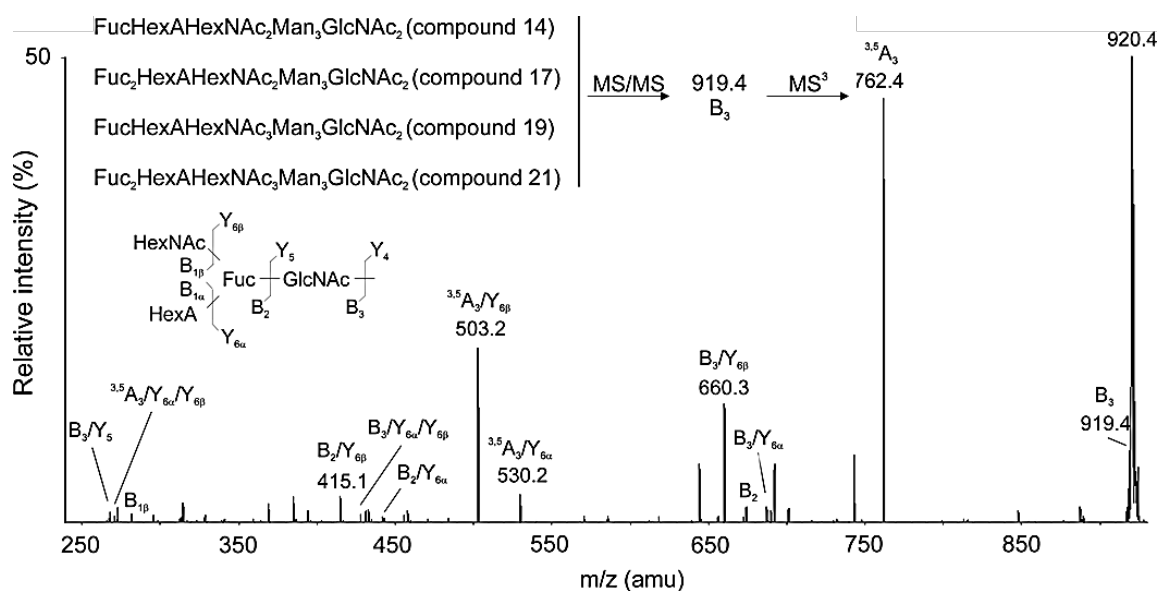


Figure 9. MS³ spectrum of the B₃ ion with m/z 919.4 [M+Na]⁺ (tetrasaccharide), after methylation of acid glycan from RvH1 with m/z 1038.5 [M+2Na]²⁺ (structural information including permethylated sites).

Further interesting carbohydrate structures are also presented in RvH1 and RvH2 by Q-trap analyses of the obtained glycans after treatment of both subunits with PNGase F glycosidase, specific for N-glycosylation. This approach proved to be very suitable for determination of the structures of 25 glycans from RvH1 and 28 glycans from RvH2. Glycan structures with one HexNAc residue and HexA bound in an oligosaccharide of RvH to an internal Fuc residue have been demonstrated [32,33].

3.2.4. Structures of the N-Glycans of RvH1 and RvH2 Determined by the Q-trap System

The comparison of the data obtained from the various employed methods and techniques determines the complex carbohydrate structure of *Rapana venosa* hemocyanin, similar to that observed for most molluscan hemocyanins. With the new methods developed, rare carbohydrate moieties of glycans, such as MeHex residues, have been detected. For the first time, a new class of glycans with an α 1,6-Fuc residue bound to HexA and HexNAc has been discovered for hemocyanins.

4. Carbohydrate Structure of Hemocyanins from *Helix lucorum*

New methods and approaches for analyzing the carbohydrate structure of hemocyanin from the snail *Helix lucorum* (HIH) show a complex carbohydrate structure (Figure 10). The glycans obtained after treatment of β c-HIH with PNGase F are determined by MALDI-MS analyses, mainly as triple charged $[M-3H]^{3+}$ ions [36–38]. The determined methylated monosaccharides in glycan structures may be MeMan or MeGal found in β c-HIH. It is important to note the absence of acid glycans with HexA at β c-HIH (Figure 11) [38].

The obtained results from conducted experiments of APTS-labeled *N*-glycans and CE-MS also confirm the complex structures of HIH [37]. Similar structures have been published for other molluscan hemocyanins, such as *Helix pomatia*, *Lymnaea stagnalis* [16], and *Arion lusitanicus* [19], suggesting that methylated *N*-glycans perform an important function in the body.

Important information is presented on the complex structure of some glycopeptides, such as the designated Xyl moiety attached to the Man residue from inside the glycan, as well as the *N*-linked glycosylation site to which the glycan is attached.

The position of the glycans in the hemocyanin is also expressed after mass spectral analysis of some β c-HIH glycopeptides. An important advantage of analyzing the carbohydrate structure of HIH is its determined nucleotide sequence, which helps determining the motifs “NXT” and “NXS”, and glycosylation of the sites accordingly.

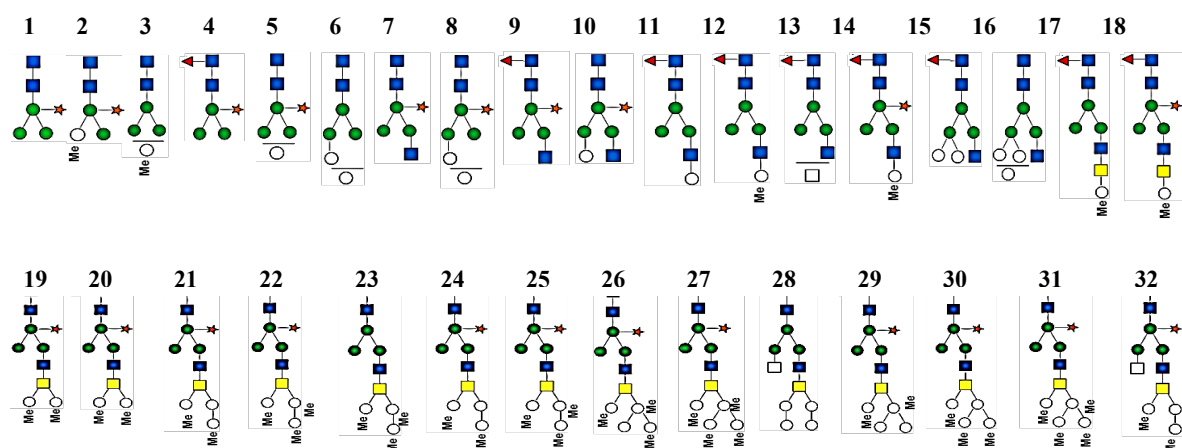


Figure 10. Neutral *N*-attached glycans from β c-HIH, obtained after treatment with PNGase. Monosaccharides, presented through nomenclature of the consortium of functional glycomics.

№	Glycan's structures from β -HIH	Q-Trap/MS (m/z)	MALDI-TOF/MS (m/z)
1		[M+Na] ⁺ 1065,3	[M+Na] ⁺ 1065,4
2		[M+Na] ⁺ 1079,1	[M+Na] ⁺ 1079,2
3		[M+Na] ⁺ 1109,2	[M+Na] ⁺ 1109,3
4		[M+Na] ⁺ 1211,3	[M+Na] ⁺ 1211,4
5		[M+H] ⁺ 1205,4	[M+Na] ⁺ 1227,4
6		[M+H] ⁺ 1235,3	[M+Na] ⁺ 1257,4
7		[M+Na] ⁺ 1268,5	[M+Na] ⁺ 1268,4
8		[M+Na] ⁺ 1367,5	[M+Na] ⁺ 1389,4
9		[M+Na] ⁺ 1414,3	[M+Na] ⁺ 1414,4
10		[M+H] ⁺ 1378,3	[M+Na] ⁺ 1430,4
11		[M+H] ⁺ 1422,3	[M+Na] ⁺ 1444,4
12		[M+Na] ⁺ 1458,3	[M+Na] ⁺ 1458,4
13		[M+Na] ⁺ 1485,4	[M+Na] ⁺ 1485,5
14		[M+H] ⁺ 1588,2	[M+Na] ⁺ 1590,3
15		[M+Na] ⁺ 1606,5	[M+Na] ⁺ 1606,6
16		[M+Na] ⁺ 1622,4	[M+Na] ⁺ 1622,5
17		[M+H] ⁺ 1639,5	[M+Na] ⁺ 1661,5
18		[M+2Na] ²⁺ 908,2	[M+Na] ⁺ 1793,5

Figure 11. Structures of glycans from β -H1H, determined through different methods [37].

Based on this information, 13 *N*-binding sites in β c-HIH, 14 in α D-HIH, and 8 in α N-HIH are identified on different positions in their FUs. Only one potential site is located in β c-HIH-e, two in β c-HIH-a, -f, and -g, and three *N*-binding sites in β c-HIH-d and -h (Figure 11) [39].

5. Carbohydrate Structure of Hemocyanins from *Haliotis tuberculata*

The complex carbohydrate structure of subunit HtH1 from molluscan hemocyanin *Haliotis tuberculata* (HtH) is also presented. Preliminary data from MALDI-MS, performed before and after treatment of the glycans from subunit HtH1 with the specific exoglycosidases β 1-3,4,6-galactosidase and α -1,6- and α -2,3,4-fucosidase, confirm the complex structure of the glycoprotein. After application of the second approach, 15 new structures of N-linked glycans to HtH1 were determined by Q-trap tandem mass spectrometry [40].

One complex structure of HtH1 is presented in Figure 12 with a glycan at m/z 1002.8 $[M+2Na]^{2+}$ and structure of MeHex₂HexMan₃GlcNAc₄Fuc₂, including two Fuc and two terminal MeHex residues bound to internal GlcNAc.

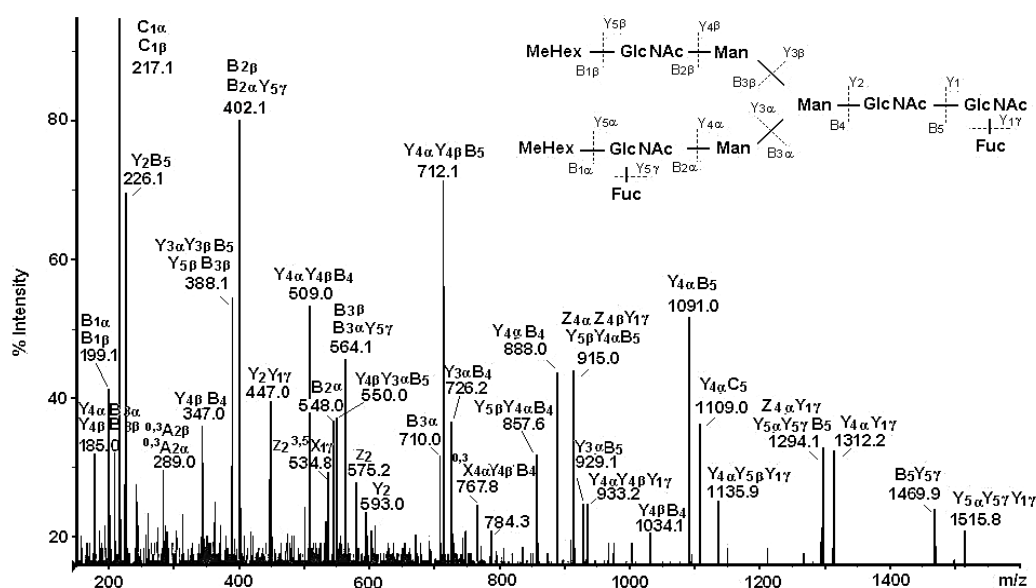


Figure 12. MS/MS spectrum of HtH1 glycan represented by ion with m/z 1002.8 $[M+2Na]^{2+}$ and complex carbohydrate structure [40].

The large set of ions in the MS/MS spectrum determines the structure of the glycan, the positions of two Fuc residues in the said chains, and one bound Fuc(α 1-6) residue to the pentasaccharide core in the glycan. On the basis of the obtained data, a conclusion about the wide variety of structures of the 15 N-linked glycans of HtH1 was made (Figure 13).

As a general feature of the structures of glycans 1, 5, 8, 10, 12, 13, 14, and 15, a bound Fuc (α 1-6) residue to the pentasaccharide nucleus is outlined. Another interesting feature is the structures of glycans 4, 12 and 15 of HtH1 with MeHex, which is also found in other hemocyanins of the molluscan species.

No	Glycan structures from HtH1	Q-Trap MS(m/z)	MALDI-TOF MS(m/z)	No	Glycan structures from HtH1	Q-Trap MS(m/z)	MALDI-TOF MS(m/z)
1		$[M+H]^+$ 1056.8	$[M+Na]^+$ 1079.4	12		$[M+2Na]^{2+}$ 923.2	$[M+Na]^+$ 1823.4
2		$[M+H]^+$ 1235.0	$[M+Na]^+$ 1257.4	13		$[M+2Na]^{2+}$ 1002.8	$[M+Na]^+$ 1983.2
3		$[M+H]^+$ 1248.6	$[M+Na]^+$ 1271.4	14		$[M+2Na]^{2+}$ 1082.2	$[M+Na]^+$ 2101.4
4		$[M+Na]^+$ 1285.0	$[M+Na]^+$ 1285.4	15		$[M+2Na]^{2+}$ 1089.2	$[M+Na]^+$ 2115.5
5		$[M+H]^+$ 1422.0	$[M+Na]^+$ 1444.4	16		$[M+2Na]^{2+}$ 1076.3	$[M+Na]^+$ 2129.5
6		$[M+H]^+$ 1436.0	$[M+Na]^+$ 1458.2				
7		$[M+H]^+$ 1568.0	$[M+Na]^+$ 1590.1				
8		$[M+H]^+$ 1582.0	$[M+Na]^+$ 1604.5				
9		$[M+H]^+$ 1639.0	$[M+Na]^+$ 1661.2				
10		$[M+2Na]^{2+}$ 878.9	$[M+Na]^+$ 1736.5				
11		$[M+2Na]^{2+}$ 902.8	$[M+Na]^+$ 1782.4				

Figure 13. Structure of the glycans isolated from HtH1

The advantage of the new methods and approaches developed should be noted, which allows the determination of complex structures of glycans 6, 8, 9, 10, 12, 13, 15, and 16 in which a novel MeMan [Fuc(α 1-3)-]GlcNAc structural motif is linked to an internal GlcNAc residue.

The discovered modifications in glycan 7 are interesting and complex, where a Xyl residue is attached to carbon-2 (C2) of Man from the core Man₃GlcNAc₂ or a Fuc residue is attached to C6 of GlcNAc. Chemical structure analysis of the oligosaccharides of *Todarodes pacificus* hemocyanin revealed that the two most abundant sequences HexNAc+Man₃GlcNAc₂ and HexHex+NAcMan₃GlcNAc₂ accounted for >95% of oligosaccharides in all individuals tested [9].

6. Glycosylated Sites of Hemocyanins from Molluscs

A general characteristic of the hemocyanins is the potential glycosylation sites in the molecule. The published nucleotide sequences can help us to suggest the features of the carbohydrate structures of the H1H, KLH, and HtH1 subunits [25,39].

6.1. Glycosylation Sites in *Helix lucorum* Hemocyanins

The established positions on the active sites in H1H are different for the α D-H1H subunit, where α D-H1H-c has one site, two are found in α D-H1H-d and -e, and three in FUs α D-H1H-a, -g, and -h. Partial information about potential binding sites is presented for the α N-H1H subunit, because of the indefinite full amino acid sequences, with one site located in FUs α N-H1H-a, -b, -g, and -h and two in α N-H1H-d and α N-H1H-e.

In confirmation of the data on the location of the putative sites on the surface of the molecule, the positions of the glycans in the constructed 3D-model of FU β c-H1H-g are presented [41] (Figure 14). The presented 3D-model of FU β c-H1H-g is composed of two domains, the active site and one putative N-linked glycosylation site at positions 125–127, which are located in/on one of the “central” domains. Another site is located at position 372–374 in the “ β -sandwich” domain with both putative sites exposed to the surface of the molecule. They are mainly concentrated between β 11- and β 12-sheet structures, as reported for the FUs β c-H1H-a, -d, -e and -g [39]. Different positions have important roles in the oligomerization of the molecule and the formation of di-decameric forms. N-glycosylation sites located at the C-end of FU “h” from α D-H1H and β c-H1H also perform a similar function. As a general characteristic of the three subunits of H1H, a potential glycosylation site in the loop between the α 16-helical and β 11-sheet structures of the β -sandwich domain can be noted. Another feature of almost all molluscan hemocyanins is the absence of a potential N-linked glycosylation site in FU “c”. This information is confirmed for both β c-H1H and α N-H1H subunits with N- glycosylation sites also not found for FU-s β c-H1H-b, α D-H1H-b, α D-H1H-f, and α N-H1H-f.

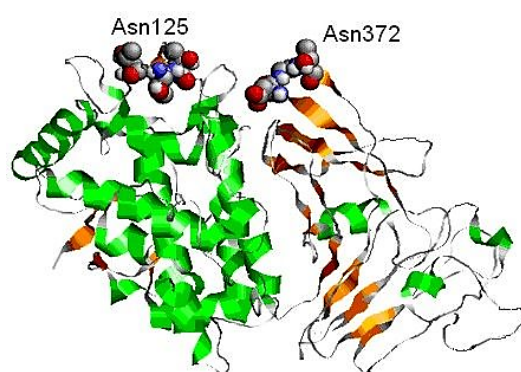


Figure 14. 3D-model of FU β c-H1H-g, built through the Swiss PDB viewer program and compared with the *Octopus dofleini* (OdH-g) FU model “g” [41]. Two possible glycosylation sites, to Asn125 and Asn245 residue, are indicated [41].

A similar motif, -Asn-Pro-Thr-, is found in FUs β c-H1H-g and α N-H1H-g, but glycan binding is blocked by a Pro residue.

This conclusion is valid for all molluscan hemocyanins, but not for the FU α D-HIH with one putative glycosylation site at position 331–333 (-AsnAspThr-). This site corresponds to the site of position 143 (-AsnThrSer-), determined by the nucleotide sequence of RvH-c, which was further demonstrated by the orcinol/H₂SO₄ assay [31]. The location of two glycans on the surface of the 3D-model of β c-HIH-g bonded to the glycosylation sites Asn125 and Asn245 corresponds to the published crystallographic structures of the Odh-g and Rth2-e FUs [42,43].

The nucleotide sequences of HtH1 and HtH2 subunits are also known [24,44], and were applied to construct 3D-models of eight FUs in which glycans are located on the surface of the molecule (Figure 15). Two models deserve special attention: one is FU HtH1-c, in which a potential glycosylation site is absent, and HtH1-h, where one of the two presented glycans is exposed on the surface of FU HtH1-h [40].

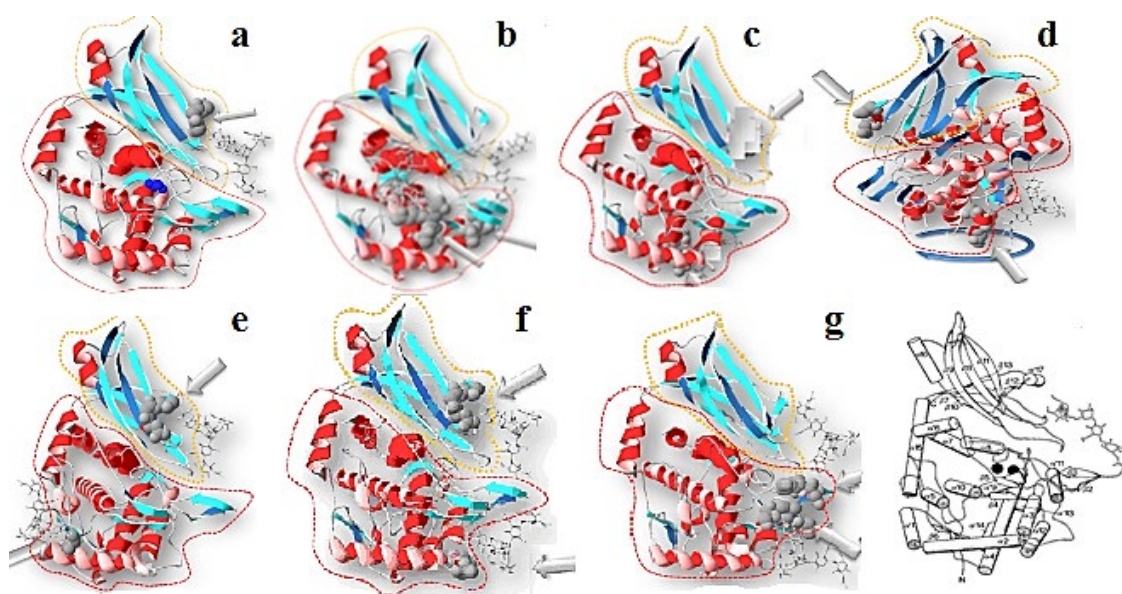


Figure 15. 3D-models of FU from «a» to «g» of HtH1 and potential N-linked sites, depicted on the basis of the structure of ODH-g. β -sandwich domain (yellow); α -nucleus (red); N-glycosylated site (grey). The glycans are marked with arrows [40].

Based on the presented data, an important role for this glycan has been suggested, which probably impedes the re-association of dissociated HtH1 subunits that do not fibrillate into di-decamers or larger aggregates as observed in some molluscan hemocyanin carbohydrate structures. The new information on the structure of hemocyanin from the garden snail *Helix lucorum* and *Haliothis tuberculata* is that they have an area of glycoproteins with predominantly methylated carbohydrate structures, which do not include acidic glycans.

The full characteristics of the complex oligosaccharide structure of hemocyanins from Molluscs are presented after identifying the connecting sites for the glycans to the polypeptide chains of RvH. The absence of information for the amino acid sequence (AAS) of RvH makes the determination of the glycosylation sites difficult. However, this is presented after comparative analysis of drawn results from different applied methods and approaches:

- Edman degradation of glycopeptides;
- Analysis of the peptides after de-glycosylation with specific glycosidases;
- Separation of glycopeptides on the columns and determination of AAS after MS/MS analyses.

N-linked glycosylation sites have been determined in isolated glycopeptides from FUs and subunits of RvH1 using a new method, including labelling of the N-linking sites, followed by treatment of the glycopeptides with PNGase F enzyme in buffer containing 50% H₂¹⁸O [26,29,33].

6.2. Glycosylation Sites in RvH

6.2.1. Glycosylation Sites in RvH Determined by nano-ESI-MS and Q-trap-LC/MS/MS Systems

The information provided on the glycosylated structure of RvH has also been confirmed after the determination of glycosylation sites with the development of a new method of analysis of the separated glycans from the polypeptide chain to which the first *N*-linked HexNAc residue remains. An important advantage of the method is fragmentation of the peptide chain by nano-ESI-MS and Q-trap-LC/MS/MS systems, which takes place in the absence of glycans in the cuvette [33].

There are two possible glycosylation sites in the eluted 31.24 min glycopeptide with AAS MGQYGNXSTNNTR. After triple quadrupole scanning, only one site was shown to be *N*-glycosylated (-NXS-). A drawback of the method is the insufficient information to demonstrate the glycosylation sites of all tested glycopeptides. This problem is solved after applying a new approach that involves MS/MS analyses of nano-LC-ESI-MS/MS triple charged ^{18}O -labeled peptides and proof of the glycosylation site.

6.2.2. Glycosylation Sites in RvH Determined after ^{18}O -labelling of Peptides

An additional characteristic of the glycosylation sites in RvH is presented by applying a new approach after separation of the *N*-glycans from the protein by glycosidase PNGase F and marking of the glycosylation sites of RvH1 in buffer with 50% H_2O . The glycosylation sites show a difference of 2 Da between *m/z* of the ^{18}O -labeled and unlabeled ions of the MS/MS spectra [33].

This new method is very sensitive and suitable for the analysis and the recognition of glycosylation sites in other glycoproteins, but the presented limited amount of data from the MS/MS spectrum does not allow for *de novo* analysis to determine the AAS of glycopeptides. This is achieved by a nano-LC-ESI-Q-trap system where, at slow fragmentation over time (TOF) of ^{18}O -labeled peptides, very low ions are detected in the MS spectra (Figure 16).

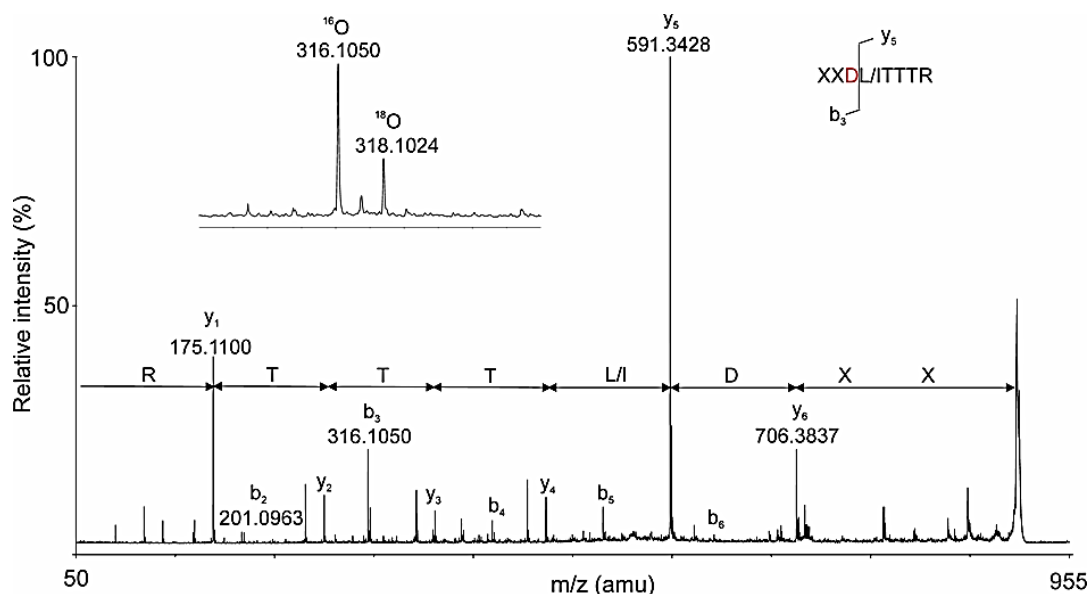
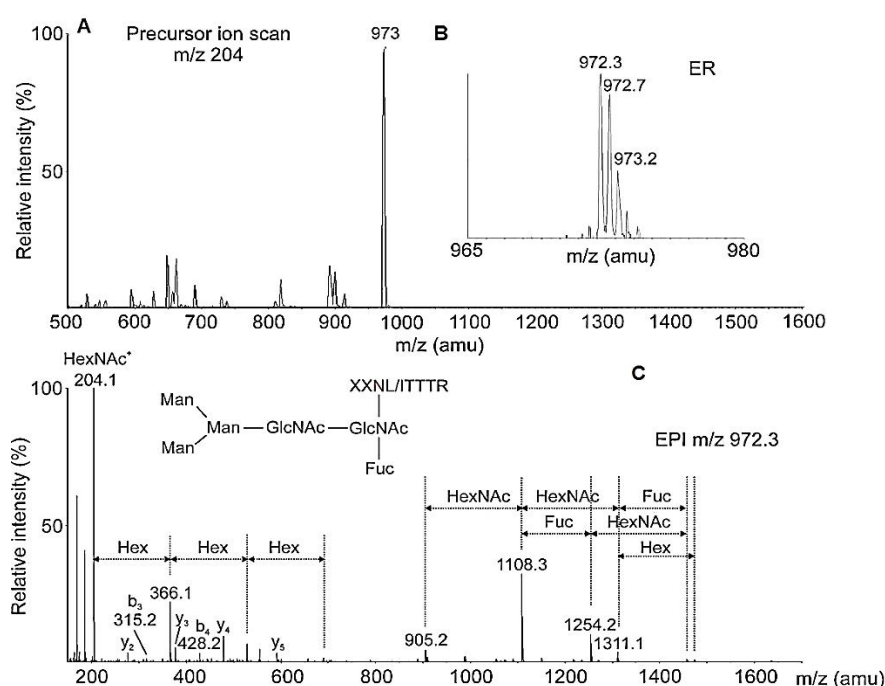


Figure 16. TOF spectrum of ^{18}O -labeled peptide, presented from an ion with *m/z* 966 $[\text{M}-2\text{H}]^{2+}$; collision energy 28 eV and TDF time 8 ms. B) The labeled ion b_3 with *m/z* 316.1050 is depicted on the figure.

The *y*-ions and the labeled b_3 with *m/z* 316.1050 $[\text{M}+\text{H}]^+$ in the TDF-MS spectrum of the ^{18}O -marked peptide with *m/z* 966.0 $[\text{M}-2\text{H}]^{2+}$ AAS with one *N*-linked glycosylation site -DLT- was determined [33]. Results from the advantage of the chosen method for detecting ^{18}O -labeled glycopeptides in biological solution are presented in Table 2. Preliminary ion scanning at 204 *m/z* of the 23-min elution fraction showed the most intensive ion with *m/z* 973 $[\text{M}+2\text{H}]^{2+}$, which represented a glycopeptide with Mw 1943.6 Da $([972.3 \times 2] - 1 = 1943.6)$ (Figure 17A, B) [33].

Table 2. AAS and carbohydrate structures of glycopeptides from RvH analyzed by ESI-MS, Q-Trap-LC/MS/MS and Q-Trap system. X is unknown IF [28,29,33].

Glycopeptides		Glycans	Mass/Charge (<i>m/z</i>)
<i>CE</i>			
1	XVYSVNGTLLGAHVLGSR	Man ₃ GlcNAc ₂	941 [M+3H] ³⁺
		HexA HexNAc ₂ Fuc ₂ Man ₃ GlcNAc ₂	1233 [M+3H] ³⁺
2	X—XFSWVDGHNTSR	GlcNAcMan ₃ GlcNAc ₂	1168 [M+3H] ³⁺
		FucMan ₃ GlcNAc ₂	1149 [M+3H] ³⁺
		Man ₃ GlcNAc ₂	1104 [M+3H] ³⁺
3	FQNDTSLDGYQAVAEFHGLPAK	FucMan ₃ GlcNAc ₂	1153 [M+3H] ³⁺
4	FQNDTSLDGFQAVAEFHGLPPK	GlcNAcMan ₃ GlcNAc ₂	1010 [M+4H] ⁴⁺
5	LHSYSGSYLNASLLHX—X	Man ₃ GlcNAc ₂	968 [M+2H] ²⁺
6	XNGTELSR / XNASELSR	Hex ₅ Man ₃ GlcNAc ₂	1373 [M+2H] ²⁺
<i>ESI-MS</i>			
7	[QK]AENLTTTR	FucMan ₃ GlcNAc ₂	1036 [M+2H] ²⁺
8	AENLTTTR	Fuc Man ₃ GlcNAc ₂	972 [M+2H] ²⁺
9	HHGHV[...K...N...]R	Fuc Man ₃ GlcNAc ₂	1396 [M+2H] ²⁺
10	FSWVDGHNTSR	Man ₃ GlcNAc ₂	1099 [M+2H] ²⁺
11	YE[IL]HAVNGST[IL]AA[IL]	Hex ₃ Man ₃ GlcNAc ₂	1339 [M+2H] ²⁺
12	YE[IL]HAVNGST[IL]AA[IL]	Hex ₃ Man ₃ GlcNAc ₂	1419 [M+2H] ²⁺
<i>Q trap-LC/MS/MS</i>			
13	MGQYGDLSNTR	Hex Man ₃ HexNAc ₂	837.9[M+2H] ³⁺
14	SVNGTLLGSQILGKPY	Fuc Man ₃ GlcNAc ₂	896 [M+3H] ³⁺
15	FSWVDGHNTSR	Man ₃ GlcNAc ₂	1099 [M+2H] ²⁺
16	AENITTTTR	Fuc Man ₃ GlcNAc ₂	972 [M+2H] ²⁺
17	FANATSIDGPNA	SO ₄ MeHexAMeHexNAc ₂ Man ₃ GlcNAc ₂	2786 [M+H] ⁺
18	EMLTLNGTNLA	MeHex ₂ AHexNAc ₂ Man ₃ GlcNAc ₂	2846 [M+H] ⁺
19	IHSYSGSYINASLLHGPSII	MeManMan ₂ GlcNAc ₂	2848 [M+H] ⁺

**Figure 17.** (A) Ion scanning of 23-min eluted fraction. (B) ER scanning of the intensive ion; (C) CEPI scanning of fragmenting ions. Increase of the mass of ion b₃ with 1 Da in EPI spectra is compared to the MS/MS of the labeled peptides [33].

Analysis of the ion with m/z 905.2 $[M+H]^+$ (a peptide without glycan, Figure 17C) from the MS/MS spectrum proves that one glycan FucMan₃GlcNAc₂ is attached to the (-NI/LT-) site of the glycopeptide (Table 2). Mass spectra report a difference of 1 Da between the values obtained for a peptide after glycan release with PNGase F and those of ¹⁸O-labeled ions. The difference reported is evidence of a glycosylation site represented by a group of $[M+H]^+$ ions resulting from the sequential fragmentation of the bound peptide with GlcNAc [33].

6.2.3. Glycosylation Sites in RvH, Determined after Fragmentation of the Genome

The positions of glycosylation sites are also demonstrated by the nucleotide sequences of hemocyanin of *Haliothis tuberculata*, *Aplysia californica*, *Nucula nucleus* and *Nautilus pompilius* [7,22,24,25,44,45], and two fragments of the RvH genome that partially encode FU-b and FU-c [31]. The negative Orcinol/H₂SO₄ test provides information on the absence of an N-linked site in RvH-b, confirmed by the obtained 316 amino acid residues (AARs) of RvH-b, as the positive test for sugars corresponds to one N-linked site (-Asp-Thr-Ser-) at the 143rd position identified in the nucleotide sequence of RvH-c.

Data obtained from MS/MS analysis of the glycopeptide with m/z 1099.19 $[M-2H]^{2+}$ (Table 2, № 10) for site AAS (FSGEVDGHNTSR) were corrected by nucleotide sequence. It is known that the Gly residue is not always reflected in the MS/MS spectrum. This was confirmed by the nucleotide analysis of the glycopeptide, where the Trp residue is replaced by -GlyGlu-. The information obtained is consistent with the published putative N-linked glycosylation sites for the FUs HtH2-c, OdH-c, and NpH-c (Figure 18) and serves as a basis for determination of other RvH characteristics.

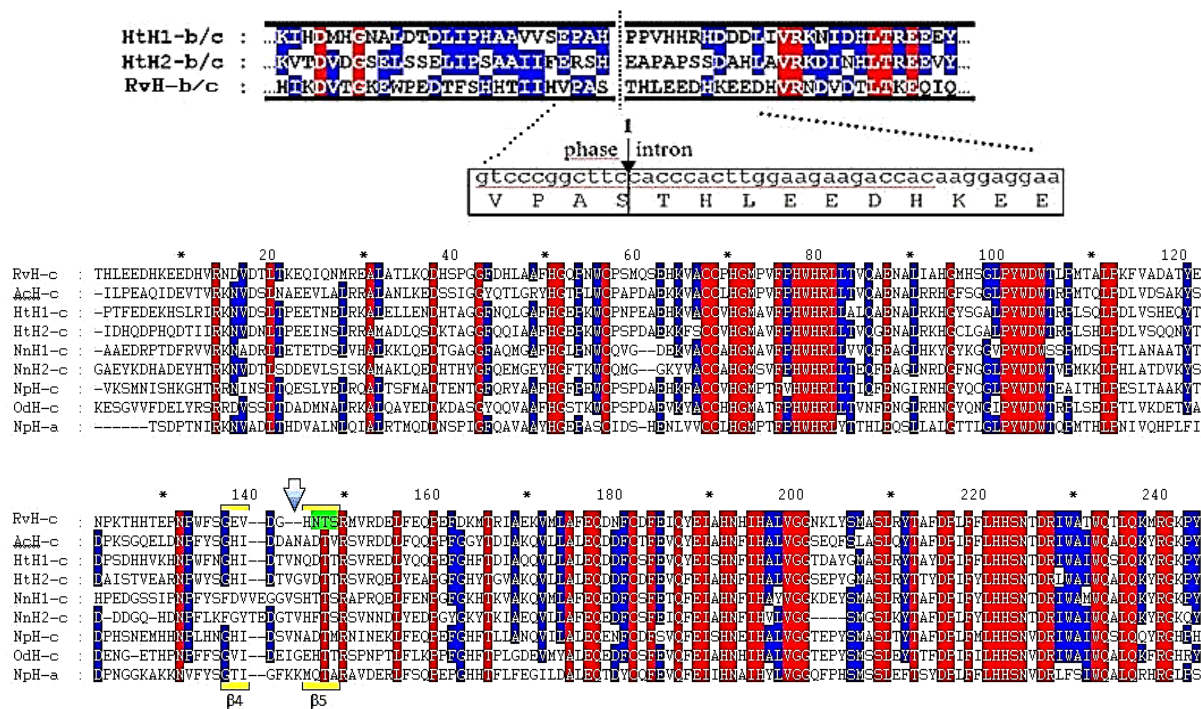


Figure 18. Comparative analysis of the AAS of RvH-c with FUs from other hemocyanins. Conservative sites of the AAS (in blue and red) and N-linked glycosylation site of RvH-c (in green). [31]

The results of the constructed 3D-model of RvH-c represent the located glycans on the surface of the molecule (Figure 19) [31], which is confirmed by the crystallographic structure of hemocyanin FUs of *Todarodes pacificus* (Figure 1) [9]. The location of N-linked glycosylation sites in molluscan hemocyanins is most commonly observed between β 11- and β 12-sheet structures in the β -sandwich domain of the FUs [25]. For the OdH-d and NpH-d FUs, these positions are similar, but for OdH-g and NpH-g they differ, which is related to the influence of the formed N-glycan site of the opposite FU OdH-g. One site in NpH-g located at the same position, but with no bound glycan, has been demonstrated [22].

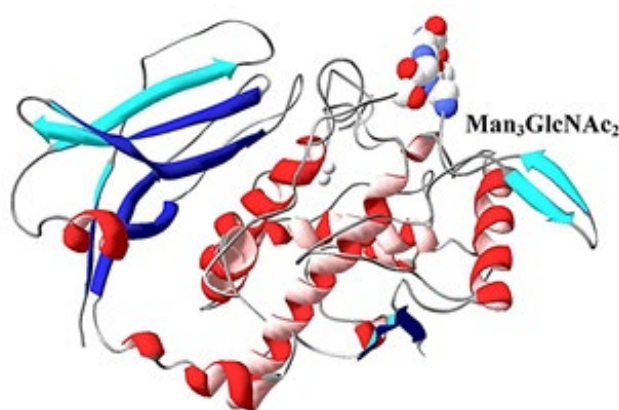


Figure 19. 3D-model of FU RvH-c, depicted on the basis of the structure of the hemocyanin from *Octopus dofleini* [41]. Glycan, attached to the polypeptide chain of FU-c. [31]

The crystal structure of *Todarodes pacificus* hemocyanin revealed *N*-linked glycosylation at five of seven possible sites, including Asn387 for FU-a, Asn806 for FU-b, Asn1498 and Asn1636 for FU-d, and Asn2472 for FU-e. Four of these sites are located at identical positions, however, FU-d possesses an additional glycan (Figure 20) [9].

The oligosaccharides were all located on the surface of the wall and formed carbohydrate clusters. The carbohydrates conjugated at Asn387 and Asn806 of protomer A (hereafter, glycan-387 A and glycan-806 A), and formed a cluster with glycan 1498 which was located at the interface among three protomers, i.e., protomers A, I, and J. Furthermore, this site is located at the interface between two plate-like protomer-dimers.

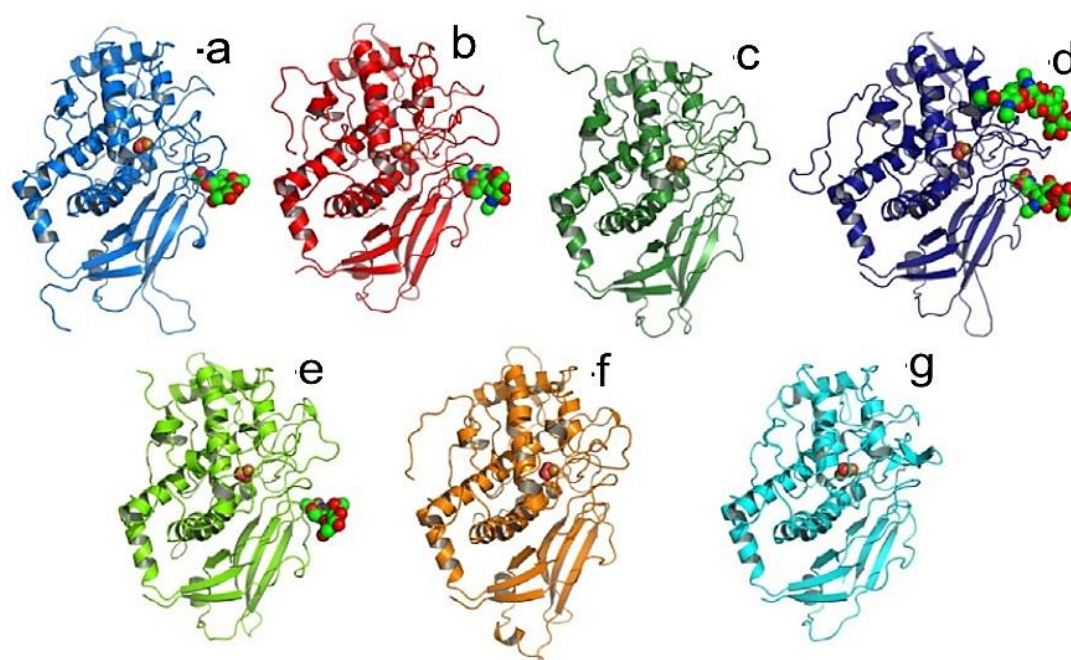


Figure 20. *N*-linked glycosylation sites of functional units of *Todarodes pacificus* hemocyanin [9].

Molluscan hemocyanins are oligomeric glycoproteins with complex dodecameric quaternary structures and heterogeneous glycosylation patterns, primarily consisting of mannose-rich *N*-glycans, which contribute to their structural stability and immunomodulatory properties in mammals [46]. The enzyme-catalyzed *N*-deglycosylation of *Concholepas concholepas* [47], *Fissurella latimarginata* [2], and KLH disrupts their quaternary structure and impairs their immunogenic effects. Biochemical analyses confirm the importance of glycans by showing that deglycosylation does not alter the secondary structure of hemocyanin, but alters their re-folding mechanism and dodecameric structure.

7. Conclusions and Future Direction

The data provide new information on the carbohydrate structures of oxygen-binding glycoproteins, ‘hemocyanins’, with two copper atoms in the active site determined by a combination of new methods and techniques. Complex carbohydrate structures have a fucosylated pentasaccharide core and a high degree of methylation.

It is important to note that, along with the common structures, all molluscan hemocyanins stand out with specific glycan structures. Of great interest are the first discovered acid glycans in RvH1 and RvH2, where at two positions from the internal Fuc residue, HexNAc and HexA is bound. The novel MeHex [Fuc (α 1-3)] GlcNAc structural motif has also been demonstrated in HtH1 glycans, which is composed of one or two MeHex and (α 1-3) Fuc bound to internal GlcNAc and an additional Xyl residue.

One of the main characteristics of hemocyanins from the snails *Helix lucorum*, *Helix pomatia*, *Lymnaea stagnalis*, and *Arion lusitanicus* is the proven predominantly methylated glycan structures of 3-O-methyl-D-mannose and 3-O-methyl-D-galactose. Similar structures have been published for other molluscan hemocyanins, suggesting that methylated N-glycans perform an important function in the body.

Comparative analyses of the primary structures of HtH1, β c-HtH, and RvH present a high degree of homology to the putative N-linked glycosylation sites located on the surface of the FUs. The position of the glycans on the surface of the molecule was also proven by the constructed 3D models of β c-HtH-g, RvH-c, as well as the eight FUs of HtH1, which correspond to the represented positions of the crystallographic structures of OdH-g and RtH2-e [42,43].

The obtained data from these studies provide new information on the carbohydrate structures of molluscan hemocyanins combining typical structural characteristics of various higher organisms (mammals, plants, insects, nematodes and trematodes). This explains the scientists’ interest in these glycoproteins, which are a suitable target for determining the effect of glycans on the structure, properties and function of glycoproteins.

Author Contributions: P.D. conceptualized this study, A.D. (Aleksandar Dolashki) and A.D. (Asya Daskalova) performed the literature search and drafted the manuscript, and W.V. edited and provided critical revisions to the finalized review. All authors have read and agreed to the published version of the review.

Funding: This work was supported by the Bulgarian Ministry of Education and Science (Grant D01-217/30.11.2018) under the National Research Programme “Innovative Low-Toxic Bioactive Systems for Precision Medicine (BioActiveMed)” approved by DCM#658/14.09.2018 and with the support of a project under contract No.DN 01/14 of 19.12.2016, funded by the Scientific Research Fund of the Ministry of Education and Science in the Republic of Bulgaria.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reily, C.; Stewart, T.J.; Renfrow, M.B.; Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **2019**, *15*, 346–366.
2. Arancibia, S.; Espinoza, C.; Salazar, F.; Del Campo, M.; Tampe, R.; Zhong, T.Y.; De Ioannes, P.; Moltedo, B.; Ferreira, J.; Lavelle, E.C.; et al. A novel immunomodulatory hemocyanin from the limpet *Fissurella latimarginata* promotes potent anti-tumor activity in melanoma. *PLoS ONE* **2014**, *9*, e87240.
3. Ayyagari, S.V.; Sreerama, K. Spectroscopic Studies on the Conformational Stability of Hemocyanin of *Pila virens* (Lamarck, 1822) in the Presence of Temperature and Detergents. *Proc. Zool. Soc.* **2019**, *73*, 302–311.
4. Ulagesan, S.; Kim, H.J. Antibacterial and Antifungal Activities of Proteins Extracted from Seven Different Snails. *Appl. Sci.* **2018**, *8*, 1362.
5. Mendoza-Porras, O.; Kamath, S.; Harris, J.O.; Colgrave, M.L.; Huerlimann, R.; Lopata, A.L.; Wade, N.M. Resolving hemocyanin isoform complexity in haemolymph of black tiger shrimp *Penaeus monodon*—implications in aquaculture, medicine and food safety. *J. Proteom.* **2020**, *218*, 103689.
6. Zhang, Z.; Wang, F.; Chen, C.; Zheng, Z.; Awewa, J.J.; Zhang, Y. Glycosylation of hemocyanin in *Litopenaeus vannamei* as an antibacterial response feature. *Immunol. Lett.* **2017**, *192*, 42–47.

7. Gatsogiannis, C.; Markl, J. Keyhole limpet hemocyanin: 9 Å CryoEM structure and molecular model of the KLH1 didecamer reveal the interfaces and intricate topology of the 160 functional units. *J. Mol. Biol.* **2009**, *385*, 963–983.
8. Tanaka, Y.; Kato, S.; Stabrin, M.; Raunser, S.; Matsui, T.; Gatsogiannis, C. Cryo-EM reveals the asymmetric assembly of squid hemocyanin. *IUCrJ* **2019**, *6*, 426–437.
9. Gai, Z.; Matsuno, A.; Kato, K.; Kato, S.; Khan, M.R.I.; Shimizu, T.; Yoshioka, T.; Kato, Y.; Kishimura, H.; Kanno, G.; et al. Crystal Structure of the 3.8-MDa Respiratory Supermolecule Hemocyanin at 3.0 Å Resolution. *Structure* **2015**, *23*, 2204–2212.
10. Lommerse, J.P.M.; Thomas-Oates, J.E.; Gielens, C.; Préaux, G.; Kamerling, J.P.; Vliegthart, J.F.G. Primary Structure of 21 Novel Monoantennary and Diantennary N-Linked Carbohydrate Chains from αD-Hemocyanin of *Helix pomatia*. *Eur. J. Biochem.* **1997**, *249*, 195–222.
11. Gielens, C.; Idakieva, K.; Van Den Bergh, V.; Siddiqui, N.I.; Parvanova, K.; Compennolle, F. Mass spectral evidence for N-glycans with branching on fucose in a molluscan hemocyanin. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 562–570.
12. Gutternigg, M.; Bürgmayr, S.; Pörtl, G.; Rudolf, J.; Staudacher, E. Neutral N-glycan patterns of the gastropods *Limax maximus*, *Cepaea hortensis*, *Planorbis corneus*, *Arianta arbustorum* and *Achatina fulica*. *Glycoconj. J.* **2007**, *24*, 475–489.
13. Wuhler, M.; Robijn, M.L.; Koeleman, C.A.; Balog, C.I.; Geyer, R.; Deelder, A.M.; Hokke, C.H. A novel Gal(β1-4)Gal(β1-4)Fuc(α1-6)-core modification attached to the proximal N-acetylglucosamine of keyhole limpet haemocyanin (KLH) N-glycans. *Biochem. J.* **2004**, *378*, 625–632.
14. Kurokawa, T.; Wuhler, M.; Lochnit, G.; Geyer, H.; Markl, J.; Geyer, R. Hemocyanin from the keyhole limpet *Megathura crenulata* (KLH) carries a novel type of N-glycans with Gal(β1-6)Man-motifs. *Eur. J. Biochem.* **2002**, *269*, 5459–5473.
15. Eckmair, B.; Jin, C.; Abed-Navandi, D.; Paschinger, K. Multistep Fractionation and Mass Spectrometry Reveal Zwitterionic and Anionic Modifications of the N- and O-glycans of a Marine Snail. *Mol. Cell. Proteom.* **2016**, *15*, 573–597.
16. Hall, R.L.; Wood, E.J.; Kamberling, J.P.; Gerwig, G.J.; Vliegthart, F.G. 3-O-methyl sugars as constituents of glycoproteins. Identification of 3-O-methylgalactose and 3-O-methylmannose in pulmonate gastropod haemocyanins. *Biochem. J.* **1977**, *165*, 173–176.
17. Wuhler, M.; Dennis, R.D.; Doenhoff, M.J.; Geyer, R. A fucose-containing epitope is shared by keyhole limpet haemocyanin and *Schistosoma mansoni* glycosphingolipids. *Mol. Biochem. Parasitol.* **2000**, *110*, 237–246.
18. Gielens, C.; De Geest, N.; Compennolle, F.; Préaux, G. Glycosylation sites of hemocyanins of *Helix pomatia* and *Sepia officinalis*. *Micron* **2004**, *35*, 99–100.
19. Gutternigg, M.; Ahrer, K.; Grabher-Meier, H.; Bürgmayr, S.; Staudacher, E. Neutral N-glycans of the gastropod *Arion lusitanicus*. *Eur. J. Biochem.* **2004**, *271*, 1348–1356.
20. Stoeva, S.; Idakieva, K.; Betzel, C.; Genov, N.; Voelter, W. Amino Acid Sequence and Glycosylation of Functional Unit Rth2-e from *Rapana thomasi* (Gastropod) Hemocyanin. *Arch. Biochem. Biophys.* **2002**, *399*, 149–158.
21. Stepan, H.; Pabst, M.; Altmann, F.; Geyer, H.; Geyer, R. O-Glycosylation of snails. *Glycoconjugate J.* **2012**, *29*, 189–198.
22. Gatsogiannis, C.; Moeller, A.; Depoix, F.; Meissner, U.; Markl, J. *Nautilus pompilius* hemocyanin: 9 Å cryo-EM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units. *J. Mol. Biol.* **2007**, *374*, 465–486.
23. Stoeva, S.; Schütz, J.; Gebauer, W.; Hundsdoerfer, T.; Manz, C.; Markl, J.; Voelter, W. Primary structure and unusual carbohydrate moiety of functional unit 2-c of keyhole limpet hemocyanin (KLH). *Biochim. Biophys. Acta* **1999**, *1435*, 94–109.
24. Lieb, B.; Boissguérin, V.; Gebauer, W.; Markl, J. cDNA sequence, protein structure, and evolution of the single hemocyanin from *Aplysia californica*, an opisthobranch gastropod. *J. Mol. Evol.* **2004**, *59*, 536–545.
25. Lieb, B.; Altenhein, B.; Markl, J. The sequence of a gastropod hemocyanin (HtH1 from *Haliotis tuberculata*). *J. Biol. Chem.* **2000**, *275*, 5675–5681.
26. Dolashka-Angelova, P.; Beck, A.; Dolashki, A.; Stevanovic, S.; Beltramini, M.; Salvato, B.; Hristova, R.; Velkova, L.; Voelter, W. Carbohydrate moieties of molluscan *Rapana venosa* hemocyanin. *Micron* **2004**, *35*, 101–104.

27. Idakieva, K.; Stoeva, S.; Voelter, W.; Gielens, C. Glycosylation of *Rapana thomasiana* hemocyanin. Comparison with other prosobranch (gastropod) hemocyanins. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2004**, *138*, 221–228.
28. Beck, A.; Hillen, N.; Dolashki, A.; Stevanovic, S.; Salvato, B.; Voelter, W.; Dolashka-Angelova, P. Oligosaccharide structure of a functional unit RvH1-b of *Rapana venosa* hemocyanin using HPLC/electrospray ionization mass spectrometry. *Biochimie* **2007**, *89*, 938–949.
29. Dolashka-Angelova, P.; Beck, A.; Dolashki, A.; Beltramini, M.; Stevanovic, S.; Salvato, B.; Voelter, W. Characterization of the carbohydrate moieties of the functional unit RvH1-a of *Rapana venosa* haemocyanin using HPLC/electrospray ionization MS and glycosidase digestion. *Biochem. J.* **2003**, *374*, 185–192.
30. Dolashka-Angelova, P.; Stevanovic, S.; Dolashki, A.; Devreese, B.; Tzvetkova, B.; Voelter, W.; Van Beeumen, J.; Salvato, B. A challenging insight on the structural unit 1 of molluscan *Rapana venosa* hemocyanin. *Arch. Biochem. Biophys.* **2007**, *459*, 50–58.
31. Dolashka-Angelova, P.; Lieb, B.; Velkova, L.; Heilen, N.; Sandra, K.; Nikolaeva-Glomb, L.; Dolashki, A.; Galabov, A.S.; Van Beeumen, J.; Stevanovic, S.; et al. Identification of glycosylated sites in *Rapana* hemocyanin by mass spectrometry and gene sequence, and their antiviral effect. *Bioconjugate Chem.* **2009**, *20*, 1315–1322.
32. Dolashka, P.; Velkova, L.; Shishkov, S.; Kostova, K.; Dolashki, A.; Dimitrov, I.; Atanasov, B.; Devreese, B.; Voelter, W.; Van Beeumen, J. Glycan structures and antiviral effect of the structural subunit RvH2 of *Rapana* hemocyanin. *Carbohydr. Res.* **2010**, *345*, 2361–2367.
33. Sandra, K.; Dolashka-Angelova, P.; Devreese, B.; Van Beeumen, J. New insights in *Rapana venosa* hemocyanin N-glycosylation resulting from on-line mass spectrometric analyses. *Glycobiology* **2007**, *17*, 141–156.
34. Dolashka, P. Tandem Mass Spectrometry and Glycoproteins. In *Tandem Mass Spectrometry—Applications and Principles*; Jeevan, K. Prasain, Ed.; InTech, Croatia 2012; Chapter 6, pp. 105–127.
35. Dolashka-Angelova, P.; Schwarz, H.; Dolashki, A.; Stevanovic, S.; Fecker, M.; Saeed, M.; Voelter, W. Oligomeric stability of *Rapana venosa* hemocyanin (RvH) and its structural subunits. *Biochim. Biophys. Acta Proteins Proteom.* **2003**, *1646*, 77–85.
36. Velkova, L.; Dolashki, A.; Dolashka, P. Analysis of a glycopeptide from structural subunit (β c-HIH) of *Helix lucorum* hemocyanin by mass spectrometry. In Proceedings of the Thirty-Third European Peptide Symposium, Sofia, Bulgaria, 31 August–5 September 2014; pp. 288–289.
37. Velkova, L.; Dolashki, A.; Dolashka, P. Carbohydrate structure of molluscan hemocyanins from snails *Helix lucorum* and *Rapana venosa*, determined by mass spectrometry. *J. Biosci. Biotechnol.* **2015**, SE/ONLINE, 75–85.
38. Velkova, L.; Dolashka, P.; Van Beeumen, J.; Devreese, B. N-glycan structures of β -HIH subunit of *Helix lucorum* hemocyanin. *Carbohydr. Res.* **2017**, *449*, 1–10.
39. Velkova, L.; Dolashka, P.; Lieb, B.; Dolashki, A.; Voelter, W.; Van Beeumen, J.; Devreese, B. Glycan structures of the structural subunit (HtH1) of *Haliotis tuberculata* hemocyanin. *Glycoconj. J.* **2011**, *28*, 385–395.
40. De Smet, L.; Dimitrov, I.; Debyser, G.; Dolashka-Angelova, P.; Dolashki, A.; Van Beeumen, J.; Devreese, B. The cDNA sequence of three hemocyanin subunits from the garden snail *Helix lucorum*. *Gene* **2011**, *487*, 118–128.
41. Kostadinova, E.; Dolashka, P.; Velkova, L.; Dolashki, A.; Stevanovic, S.; Voelter, W. Positions of the glycans in molluscan hemocyanin, determined by fluorescence spectroscopy. *J. Fluoresc.* **2013**, *23*, 4, 753–760.
42. Cuff, M.E.; Miller, K.I.; van Holde, K.E.; Hendrickson, W.A. Crystal structure of a functional unit from Octopus hemocyanin. *J. Mol. Biol.* **1998**, *278*, 855–870.
43. Perbandt, M.; Guthöhrlein, E.W.; Rypniewski, W.; Idakieva, K.; Stoeva, S.; Voelter, W.; Genov, N.; Betzel, C. The structure of a functional unit from the wall of a gastropod hemocyanin offers a possible mechanism for cooperativity. *Biochemistry* **2003**, *42*, 6341–6346.
44. Altenhein, B.; Markl, J.; Lieb, B. Gene structure and hemocyanin isoform HtH2 from the mollusc *Haliotis tuberculata* indicate early and late intron hot spots. *Gene* **2002**, *301*, 53–60.
45. Bergmann, S.; Lieb, B.; Ruth, P.; Markl, J. The hemocyanin from a living fossil, the cephalopod *Nautilus pompilius*: Protein structure, gene organization, and evolution. *J. Mol. Evol.* **2006**, *62*, 362–374.
46. Salazar, L.M.; Jiménez, M.J.; Villar, J.; Rivera, M.; Báez, M.; Manubens, A.; Inés Becker, M. N-glycosylation of mollusk hemocyanins contributes to their structural stability and immunomodulatory properties in mammals. *J. Biol. Chem.* **2019**, *294*, 19546–19564.

47. De Ioannes, P.; Moltedo, B.; Oliva, H.; Pacheco, R.; Faunes, F.; De Ioannes, E.A.; Becker, M.I. Hemocyanin of the molluscan *Concholepas concholepas* exhibits an unusual heterodecameric array of subunits. *J. Biol. Chem.* **2004**, *279*, 26134–26142.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).