

# N-Glycan structures of $\beta$ -HIH subunit of *Helix lucorum* hemocyanin

Lyudmila Velkova,<sup>a,\*</sup> Pavlina Dolashka,<sup>a</sup> Jozef Van Beeumen,<sup>b</sup> and Bart  
Devreese<sup>b</sup>

<sup>a</sup> *Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, 9 G. Bonchev St., Sofia 1113, Bulgaria*

<sup>b</sup> *Laboratory of Protein Biochemistry and Biomolecular Engineering, Ghent University, KL Ledeganckstraat 35, 9000 Ghent, Belgium*

\* Corresponding author.

Email address: [lyudmila\\_velkova@abv.bg](mailto:lyudmila_velkova@abv.bg)

## Abstract

The carbohydrate structures of molluscan hemocyanins have recently received particular interest due to their specific monosaccharide composition, as well as their immunostimulatory properties and application in clinical studies. For the first time, we investigated N-glycans of the structural subunit  $\beta$ -HIH of hemocyanin isolated from *Helix lucorum*. In total, 32 different glycans were enzymatically liberated and characterized by tandem mass spectrometry using a Q-Trap mass spectrometer. Our study revealed a highly heterogeneous mixture of glycans with composition  $\text{Hex}_{3-7}\text{HexNAc}_{2-5}\text{MeHex}_{0-4}\text{Pent}_{0-1}\text{Fuc}_{0-1}$ . The oligosaccharide chains are mostly modified at the inner core by  $\beta$ 1-2-linked xylose to  $\beta$ -mannose, by  $\alpha$ 1-6-fucosylation of the innermost GlcNAc residue (the Asn-bound GlcNAc), and by methylation. The glycans of  $\beta$ -HIH mainly contain a terminal MeHex residue; in some cases even two, three or four of these residues occur. Several carbohydrate chains in  $\beta$ -HIH are core-fucosylated without Xyl and also possess a high degree of methylation. This study shows the presence of mono- and bi-antennary N-glycans as well as hybrid type structures with or without core-fucosylation.

*Keywords:* molluscan hemocyanins; *Helix lucorum* hemocyanin (HIH);  $\beta$ -HIH structural subunit; mass spectrometry; Q-Trap; N-glycan

## 1. Introduction

Hemocyanins are oligomeric blue copper-containing respiratory glycoproteins with extremely high molecular weight and complex quaternary structure. They play a role as dioxygen carriers in the hemolymph of different species of molluscs and arthropods.<sup>1,2</sup> An important feature of hemocyanin structures is their carbohydrate content, an essential feature for their biomedical application. The relevance of glycosylation on hemocyanins as a factor for their immunostimulatory properties has been revealed from studies on the hemocyanin isoforms KLH1 and KLH2 of *Megathura crenulata*, which are widely used in experimental immunology and clinical practice.<sup>3-7</sup> The clinical success of intravesical administration of KLH to patients with bladder carcinoma is assumed to be based on the presence of the disaccharide Gal( $\beta$ 1-3)GalNAc determinants which are cross reactive with an equivalent epitope on the bladder tumor cell surface. The proposed mechanism of action involves immune activation due to the presence of epitopes such as the Thomson–Friederich antigen and N-linked oligosaccharides carbohydrates motifs, as well as the enhancement of T helper type-1 immunity.<sup>8-10</sup>

The versatile properties of KLH in biomedical and clinical applications have led to a growing interest in finding other hemocyanins with similar or even more potent immunomodulatory properties. Hemocyanins from different species of molluscs, such as *Rapana venosa* (RvH), *Helix lucorum* (HIH), previously called *Helix vulgaris*, and *Concholepas concholepas* have also been shown to be immunogenic and to display significant antitumor activities.<sup>10-14</sup> It was also demonstrated that HIH and RvH have a direct antiproliferative effect on CAL-29 and T-24 bladder cancer cell lines, and that the antitumor properties of HIH are even superior to KLH.<sup>15,16</sup> Moreover, also the antiviral activity of molluscan hemocyanins (RvH,  $\beta$ -HIH, KLH, and *Haliotis tuberculata* hemocyanin (HtH)), which is also based on the presence of carbohydrate structures, has recently drawn considerable attention.<sup>17-25</sup>

The carbohydrate content of molluscan hemocyanins (2–9%, w/w) is higher than that of most arthropod hemocyanins.<sup>26,27</sup> Moreover, hemocyanins of *Helix pomatia* (HpH) and *Lymnaea stagnalis* contain monosaccharides not usually found in animal glycoproteins (*H. pomatia*: D-xyl and 3-O-methyl-D-galactose (3MeGal)<sup>28,29</sup>; *L. stagnalis*: D-Xyl, 3-O-methyl-D-mannose (3MeMan), and 3MeGal)<sup>30</sup>. 4-O-Methylgalactose was detected for the first time in HpH.<sup>29</sup> Fucosylated LacdiNAc, core xylose, galactosylated

(i.e., Gal( $\beta$ 1-4)Gal( $\beta$ 1-4)Fuc( $\alpha$ 1-6)) as well as Gal( $\beta$ 1-3)-GalNAc and Gal( $\beta$ 1-6)Man motifs have been determined in KLH glycans.<sup>8,9,31,32</sup> In other species, the glycans present on the haemocyanins vary from the 'less exciting' hybrid and Man<sub>5</sub>GlcNAc<sub>2</sub> structures in *Panulirus interruptus*<sup>33</sup> to oligosaccharides with methylated mannose in *Hippopus hippopus*<sup>34</sup>, methylated Lewis-like motifs, including MeHex and Fuc residues  $\alpha$ 1-3-linked to an internal GlcNAc residue (MeHex[Fuc( $\alpha$ 1-3)]GlcNAc motif) in HtH,<sup>35</sup> sulfated mannose, methylated GlcNAc and methylated galactose in RvH.<sup>17,36,37</sup> Recent structural studies of both isoforms of RvH demonstrated the presence of unusual N-glycan structures with an internal fucose residue ( $\beta$ 1-2)-connecting GalNAc and a hexuronic acid.<sup>18,20,38</sup> In the snail *Biomphalaria glabrata*, the intermediate host of the trematode *Schistosoma mansoni*, over 100 N-glycans have been isolated from haemolymph proteins, including biantennary glycans with core xylose, core  $\alpha$ 1-6-linked fucose, methylmannose and fucosylated LacdiNAc motifs. These features provide cross-reactivity with anti-schistosoma antibodies.<sup>39</sup> Methylated galactose and core xylose have been found on HpH and *L. stagnalis* haemocyanin with peripheral blood group H disaccharide (Fuc( $\alpha$ 1-2)Gal $\beta$ ) in the former and core  $\alpha$ 1-6-linked fucose in the latter.<sup>28-30</sup>

Native haemocyanin of a land snail *Helix lucorum* is organized by three different structural subunits:  $\beta$ -HIH,  $\alpha$ <sub>D</sub>-HIH, and  $\alpha$ <sub>N</sub>-HIH. Each of them, ranging from 350 to 450 kDa, includes eight globular-folded domains known as functional units (FUs) with molecular masses of about 50–60 kDa.<sup>40</sup> Previously, based on the gene sequence of  $\beta$ -HIH, we have shown that this subunit contains 13 potential N-glycosylation sites.<sup>41</sup> While  $\beta$ -HIH manifested immunological, potential antitumor and antiviral activities, information on the structure of its glycans is still incomplete.<sup>42</sup> The aim of the present study was therefore to identify and to characterize the carbohydrate structures of  $\beta$ -HIH using tandem mass spectrometry. This information is essential for understanding biomedical properties of  $\beta$ -HIH, and is a prerequisite for producing recombinant glycoprotein therapeutics.

## 2. Results and Discussion

### 2.1. Identification of glycans

The subunit  $\beta$ -HIH was subjected to PNGase F digestion and the glycans were separated from the protein. MALDI-TOF-MS was used for determination of the molecular

mass and the relative abundance of the native glycans and provided first data on their monosaccharide composition (Fig. 1a,b and Table 1). Figure 1a shows eight predominant peaks for the  $[M+Na]^+$  ions at  $m/z$  2145.6, 2131.6, 1211.4, 1999.6, 1065.4, 2161.6, 1227.4, and 1414.4, corresponding to complex and high-mannose type N-linked glycans (Fig. 1b).

## 2.2. Sequencing of the glycans by Q-Trap analysis

Most of the glycans presented in Fig.1a were identified using a ESI-Q-Trap system (Table 1). Although it is likely that we did not detect all glycan structures of  $\beta$ -HIH, there is sufficient data (Table 1) to compare the inferred structures with those found in other gastropodan hemocyanins.

Sequencing and determination of configurations of the 32 N-glycans released from  $\beta$ -HIH after PNGase F treatment were performed by Q-Trap tandem mass spectrometry. The glycan sequence and branching was predominantly obtained by analysis of the dominant B, C, and Y, Z ions, and those from the internal oligosaccharide chain YB, YC in the MS/MS spectra. Furthermore, cross-ring fragments A and X allowed elucidation of the linkage positions in the glycans as described<sup>38</sup>. Analysis of the carbohydrate structure of  $\beta$ -HIH was complicated by the vast heterogeneity of its glycans resulting in the expression of a great variety of different isomeric and/or isobaric structures.

The results (Fig.1a,b; Table 1) revealed that  $\beta$ -HIH carries predominantly complex and oligomannose types of oligosaccharide structures (glycans Nos. 4, 9, 21, 22, 24-26, and 1-3, 5, 6, 8, respectively) as was observed in other molluscan hemocyanins. Most of them are methylated. The structures appeared to be based on various larger N-glycans in which 3-O-methylhexoses were found, in addition to carbohydrates containing fucose and/or xylose residues. Methylated sugars were also identified in other molluscan hemocyanins, such as 3-O-methyl-D-mannose and 3-O-methyl-D-galactose in HpH and *L. stagnalis* haemocyanin.<sup>28-30</sup> It has been suggested that a high degree of methylation in the gastropod *Arion lusitanicus* is an important regulating event in this organism.<sup>43</sup> The hybrid type structures of glycans Nos. 10, 15 and 16 are present at relatively lower abundances than those of complex and high-mannose type N-linked glycan structures

(Fig.1a,b). Below, elucidation of these unusual carbohydrate structures is illustrated by interpretation of the mass spectra of four glycans.

Several glycans of the high mannose type were identified in  $\beta$ -HIH. The MS/MS spectrum of the N-glycan with the  $[M+Na]^+$  ion at  $m/z$  1079.1 (Fig. 2 a,b; glycan No. 2 in Table 1) reveals a core-xylosylated structure, comprising one xylose residue and a terminal MeHex residue, both linked to  $\beta$ -mannose of the inner core. The presence of these two monosaccharides was easily revealed from two fragment peaks indicative for the sequential loss of Xyl and GlcNAc,  $Y_{3\gamma}$  at  $m/z$  947.2 and  $Y_{3\gamma}B_3$  at  $m/z$  726.0, respectively. The terminal position of MeHex was revealed from the  $C_{1\beta}$  fragment at  $m/z$  217.0. Probably it is MeMan as this residue is common in similar glycan structures of  $\alpha$ -HpH.<sup>28,29</sup> Other fragments, such as  $Z_1$  (or  $Y_2B_3$ ) at  $m/z$  226.0 ( $GlcNAc+Na^+$ ),  $Y_2$  at  $m/z$  447.0, and  $Y_{\gamma}B_3$  at  $m/z$  726.0 ( $MeHex_1Man_2GlcNAc_1+Na^+$ ), defined the structure of the glycan. The linkage positions of outer hexose and methylhexose residues were assigned by analysis of fragment ions due to cross-ring cleavages, which are also present in the spectrum (Fig. 2 a, b). Cross-ring fragment ions  $^{3,5}A_2$  ( $m/z$  273.2) and  $^{0,4}A_2$  ( $m/z$  259.0) confirmed the linkage position of MeHex. The presence of a D-ion was diagnostic for the side-chain composition; it was  $Y_{3\beta}B_2$  ( $m/z$  492.9) corresponding to  $Y_{3\beta}Y_{3\gamma}B_2$  ( $m/z$  361.0), after loss of one Xyl residue (132 Da). The cross-ring fragment ion  $^{0,2}X_{3\gamma}C_2$  at  $m/z$  583.0 confirmed the position of Xyl and the presence of MeHex. The fragment ions  $B_2Y_{3\gamma}$  at  $m/z$  523.1 and  $^{0,2}X_{3\gamma}C_2$  at  $m/z$  583.0 confirmed the methylhexose residue. The presence of outer hexose, MeHex and Xyl residues was further corroborated by the fragment ions  $B_2$  at  $m/z$  655.1,  $B_3$  at  $m/z$  858.1 and  $^{0,2}A_4$  at  $m/z$  977.9. The MS/MS spectrum clearly demonstrated the presence of a core-linked xylose as well as a terminal methylhexose attached to the  $\beta$ 1-4-linked Man of the core.

Oligosaccharide structures with partly or fully methylated terminal hexoses were detected in  $\beta$ -HIH. The MS/MS spectrum of the second glycan with the  $[M+Na]^{2+}$  ion at  $m/z$  1092.3 (Fig. 3 a,b, glycan No. 26), clearly demonstrates the presence of three MeHex residues, two of them being terminally and one internally linked, one xylose  $\beta$ 1-2-linked to the core mannose, and one terminal Hex linked to the internal MeHex. The fragment ion at  $m/z$  217.1 obviously again corresponds to a terminal MeHex, and  $Y_{7\alpha'}[2]C_{2\alpha'}$  at  $m/z$  393.1 fits the structure of a terminal MeHex linked to an internal MeHex residue. The adjacent ion  $Y_{7\alpha'}[1]C_{2\alpha'}$  at  $m/z$  379.1 justifies this interpretation. We also

observe an ion at  $m/z$  402.1 which corresponds to MeHex-HexNAc (Fig.3 a,b). This finding suggests that a second terminal MeHex is linked at an internal HexNAc residue. Subsequent ions  $C_{2\alpha}'$  at  $m/z$  555.1 (corresponding to MeHex[MeHex]Hex),  $Y_{6\alpha}'B_{4\alpha}$  at  $m/z$  605.2,  $Y_{7\alpha}'_{[1]}B_{3\alpha}'$  (or  $Y_{6\alpha}''B_{3\alpha}$ ) at  $m/z$  740.3,  $Y_{7\alpha}'_{[2]}B_{3\alpha}$  at  $m/z$  754.2, and  $B_{3\alpha}$  at  $m/z$  916.2 confirm the proposed structure. The  $B_{4\alpha}$  ion was due to the presence of an internal HexNAc residue. On the basis of similarity with the structure of glycans from HpH, we assume this residue to be GalNAc.<sup>29,44</sup> The D-ion  $Y_{3\alpha}C_6$  at  $m/z$  497.0 indicates the presence of a core  $\beta$ 1-2-linked xylose. The observed mass difference between the ion  $Y_{4\alpha}C_6$  at  $m/z$  659.2 and fragment ion  $Y_{4\alpha}B_6Y_{3\gamma}$  at  $m/z$  509.0 is in agreement with the loss of one Xyl residue (132 Da) and a molecule of water. Further evidence for the presence of a xylose residue linked the core mannose is ions  $Y_{4\alpha}B_7$  (or  $Y_{5\alpha}B_6$ ) at  $m/z$  844.3,  $Z_{4\alpha}$  ( $Y_{5\alpha}B_7$ ) at  $m/z$  1047.4, and  $Y_{4\alpha}$  (and  $Y_{5\alpha}C_7$ ) at  $m/z$  1065.2. The presence of the cross-ring fragment ions  $^{0,2}A_{3\alpha}$  at  $m/z$  833.0,  $^{0,4}A_{3\alpha}$  at  $m/z$  259.0,  $^{3,5}A_{2\alpha}'$  at  $m/z$  273.0,  $^{2,4}X_{5\alpha}Y_{6\alpha}''$  at  $m/z$  1421.5, and the ions at  $m/z$  1250.2 and 1268.4 (Fig.3 a,b) confirmed the proposed structure of glycan No. 26.

Glycan No. 24 that displayed molecular mass 2131.6 in the MALDI-MS spectrum was sequenced as a doubly-charged  $[M+2Na]^{2+}$  species at  $m/z$  1077.2. As shown in the structure (Fig. 4), one pentose, one deoxyhexose and two terminal MeHex residues could be demonstrated, thus indicating a molecule with composition MeHex<sub>2</sub>Hex<sub>1</sub>HexNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>Fuc<sub>1</sub>Xyl<sub>1</sub>. The fragment ion  $C_{1\alpha}'$  (or  $C_{2\alpha}''$ ) at  $m/z$  217.1 and the cross-ring fragment ion  $^{3,5}A_{2\alpha}'$  (or  $^{3,5}A_{3\alpha}$ ) at  $m/z$  273.0 in the MS/MS spectrum obviously correspond to a terminal MeHex and are followed by the ions  $B_{2\alpha}'$  at  $m/z$  361.1 and  $C_{2\alpha}'$  at  $m/z$  379.1, which refer to MeHex-Hex. The sodium adduct ion  $Y_{6\alpha}'B_{3\alpha}$  at  $m/z$  402.1 suggests that the second terminal MeHex is linked to the internal HexNAc residue. The inferred structure with two terminal methylhexose residues is further confirmed by ion  $B_{3\alpha}$  at  $m/z$  740.1 corresponding to MeHex-Hex-[(MeHex)]HexNAc, followed by the ions  $B_{4\alpha}$  and  $B_{5\alpha}$ . The ion  $B_6$  at  $m/z$  1561.2 derives from the structure with two additional hexose residues and one core  $\beta$ 1-2-linked xylose. An additional evidence for the  $\beta$ 1 $\rightarrow$ 2-linked pentose residue in the partial structure Xyl( $\beta$ 1 $\rightarrow$ 2)Man( $\beta$ 1 $\rightarrow$ 4)GlcNAc( $\beta$ 1-4)GlcNAc is the ion  $Y_{3\alpha}B_6$  at  $m/z$  479.2, and fragment ions  $Y_{4\alpha}B_6$  at  $m/z$  641.2 and  $Y_{4\alpha}C_6$  at  $m/z$  658.9 (in accordance with the composition Xyl<sub>1</sub>Man<sub>3</sub>),  $Y_{4\alpha}B_7$  (or  $Y_{5\alpha}B_6$ ) at  $m/z$

844.3 (Xyl<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>1</sub>), Y<sub>5α</sub>B<sub>7</sub> and Z<sub>4α</sub>Y<sub>1γ</sub> at *m/z* 1047.5 (Xyl<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>), Y<sub>5α</sub>Y<sub>1γ</sub> at *m/z* 1268.3 (Xyl<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>3</sub>), and Y<sub>6α</sub>'Y<sub>1γ</sub> at *m/z* 1647.2 (Xyl<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>HexNAc<sub>2</sub>MeHex<sub>1</sub>). The fragment ions Z<sub>4α</sub> at *m/z* 1193.4 and Y<sub>4α</sub> at *m/z* 1211.3 (Xyl<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>Fuc<sub>1</sub>), Z<sub>5α</sub> at *m/z* 1396.3 and Y<sub>5α</sub> at *m/z* 1414.4, as well as the ion Y<sub>6α</sub>'Y<sub>6α</sub>' at *m/z* 1617.3 indicate the presence of Xyl and Fuc residues. The cross-ring fragment ion <sup>0,2</sup>X<sub>2</sub> at *m/z* 767.2 confirms the presence of a β1-2-linked xylose residue attached to β-mannose and of a core-linked Fuc. α1-6-linkage of Fuc to the terminal GlcNAc residue of the core followed from the cross-ring fragment ion <sup>2,4</sup>X<sub>0</sub> at *m/z* 330.1, resulting from ion <sup>1,5</sup>X<sub>1</sub> at *m/z* 418.1. These results are in agreement with glycan structures of HpH and other gastropods' hemocyanins.<sup>29,45-47</sup>

The last glycan under discussion is of hybrid-type and contains a core-linked Fuc(α1-6)GlcNAc disaccharide (Fig. 5 a, b, glycan No. 15). Attachment of a side-chain Fuc to the core follows from the Z<sub>1</sub> ion at *m/z* 372.0. and the Y<sub>1</sub> ion at *m/z* 390.2 (GlcNAc<sub>1</sub>Fuc<sub>1</sub>), as well as the Z<sub>2</sub> ion at *m/z* 575.2 and the Y<sub>2</sub> ion at *m/z* 593.2 (GlcNAc<sub>2</sub>Fuc<sub>1</sub>). The presence of the fragment ion Y<sub>4α</sub>'Y<sub>4α</sub>'Y<sub>4β</sub> at *m/z* 1079.2 (Man<sub>3</sub>GlcNAc<sub>2</sub>Fuc<sub>1</sub>) followed in increasing mass order by ions Y<sub>3α</sub> at *m/z* 1120.4 (Man<sub>2</sub>GlcNAc<sub>3</sub>Fuc<sub>1</sub>) and Y<sub>4α</sub>'Y<sub>4α</sub>' at *m/z* 1282.1 confirm the core-linked Fuc. The observed mass difference (146 Da) between the [M+Na]<sup>+</sup> ion at *m/z* 1606.5 and the fragment ion Y<sub>1γ</sub> at *m/z* 1460.5 also is in agreement with the loss of the Fuc residue.

The fragment ions <sup>0,4</sup>A<sub>5</sub> at *m/z* 229.1, <sup>0,2</sup>A<sub>5</sub>Y<sub>1</sub> at *m/z* 289.1, and <sup>2,4</sup>X<sub>0</sub> at *m/z* 330.1 confirm the presence of a terminal fucose residue α1-6-linked to the Asn-bound terminal GlcNAc residue of the core. An additional evidence is the presence of the cross-ring fragment ions <sup>1,5</sup>X<sub>1</sub> at *m/z* 418.1 and <sup>0,2</sup>X<sub>2</sub> at *m/z* 635.2. The rest of the sequence follows from the fragment ions B<sub>1α</sub>' (or B<sub>1α</sub>'') at *m/z* 185.0 and C<sub>1α</sub>' (or C<sub>1α</sub>'') at *m/z* 203.0, which obviously correspond to a hexose, and B<sub>1β</sub> (or Y<sub>2</sub>B<sub>4</sub>) at *m/z* 226.0, which stands for one GlcNAc. They are followed in increasing mass order by the ion B<sub>2β</sub> at *m/z* 388.1 (Hex-GlcNAc) and C<sub>2α</sub> at *m/z* 527.2, corresponding to three Hex residues (three Man residues by analogy). Furthermore, the D-ion Y<sub>3β</sub>B<sub>3</sub> at *m/z* 671.2 (Hex<sub>2</sub>Man<sub>2</sub>) and the ion Y<sub>4β</sub>B<sub>3</sub> at *m/z* 833.2 (Hex<sub>2</sub>Man<sub>3</sub>), as well as the B<sub>3</sub> ion (or Y<sub>4α</sub>B<sub>4</sub>) at *m/z* 1036.3, corresponding to Hex<sub>2</sub>Man<sub>3</sub>GlcNAc, and the B<sub>4</sub> ion at *m/z* 1239.4 (Hex<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>) confirmed the suggested structure. The presence of the cross-ring fragment ions <sup>0,4</sup>A<sub>2α</sub> at *m/z* 245.0 and <sup>3,5</sup>A<sub>2α</sub> at *m/z* 259.0, as well as <sup>3,5</sup>X<sub>3α</sub>C<sub>4</sub> (or <sup>3,5</sup>X<sub>3α</sub>Y<sub>4β</sub>Y<sub>1γ</sub>) at *m/z* 1184.4 indicates that



the terminal hexose (mannose) of the nonreducing end is  $\alpha$ 1-6-linked to the mannose at the  $\alpha$ 1-6-antennae of the N-glycan (Fig.5 b). A similar hybrid-type structure with a fucose  $\alpha$ 1-6-linked to the terminal GlcNAc residue of the core was established in mammals and is also a potential biosynthetic intermediate (as FucT needs GlcNAcT to act first), which remains unprocessed by  $\alpha$ -mannosidase II. Nevertheless, hybrid glycans are generally underrepresented on haemocyanins, which otherwise tend to possess high-mannose and core fucosylated complex structures.<sup>48</sup> The N-glycans of hemocyanins are characterized with mainly high-mannose and complex types structures, and thus hybrid core fucosylated structures are out of the ordinary.<sup>8,9,29,35</sup>

### 2.3. Localization of the carbohydrate linkage sites in 3D-model of $\beta$ -HIH-g

Glycosylation not only contributes to the physical properties of proteins, such as conformational stability, protease resistance, charge and hydrophilicity, but glycans may also function as recognition determinants in host-pathogen relationships, protein targeting and cell-cell interactions.<sup>46,47</sup> Therefore, the position of the glycosylation site in the amino acid sequence of the protein is an important feature.

Analysis of the gene sequence of the  $\beta$ -subunit reveals that there are 13 potential N-glycosylation sites, based on the N-glycan motifs NXT and/or NXS. Fourteen sites are located in the  $\alpha$ <sub>D</sub>-subunit, and seven in the  $\alpha$ <sub>N</sub>-subunit, but only some of these are effectively glycosylated.<sup>41</sup> The distribution of these glycosylation sites in several functional units is different: in  $\beta$ -HIH we observed a single potential site in FU-e, two in FU-a, FU-f and FU-g, and three in FU-d and FU-h, whereas FU-b and FU-c have none.<sup>41,49</sup> Our finding that the orcinol/H<sub>2</sub>SO<sub>4</sub> glycosylation test was negative for the latter two units and positive for the others (data not shown) thus corroborates the gene sequence data.

A model of  $\beta$ -HIH-g, based on the known structure of the functional unit 'g' from *Octopus dofleini* hemocyanin (OCH-g), is presented in Fig. 6. It shows the typical presence of two domains, one being the so-called copper-containing 'central' domain containing the active center, the other known as the ' $\beta$ -sandwich' domain located at the periphery of the molecule. The model clearly shows that the carbohydrate chains are exposed on the surface of the functional unit. One of the glycosylation sites is located in the domain at Asn125, while the other resides in the  $\beta$ -sandwich domain at position Asn

372. The model indicates that Asn125 is the most accessible one. This might be essential for 'g' to establish and maintain any contact with other functional units, and might contribute to the overall stability of the oligomeric hemocyanin by allowing the oligosaccharide trees to organize and stabilize the structural subunit of molluscan hemocyanins. This remains to be determined through further studies. Moreover, the hemocyanin antiviral activity has been attributed to van der Waals interactions of surface-exposed hemocyanin carbohydrate chains with surface-exposed amino acid or carbohydrate residues of viruses.<sup>18,19,24,25</sup>

#### **2.4. Comparison with other hemocyanins and functional consequences.**

The first analyses of a snail N-glycan were carried out on *Helix pomatia*  $\alpha$ -hemocyanin by <sup>1</sup>H NMR spectroscopy.<sup>28</sup> It was the first detection of xylose as a component of an animal N-glycan.<sup>28</sup> The major low molecular weight N-glycan of this protein was found to be a Man<sub>3</sub>GlcNAc<sub>2</sub>-core both with a Fuc residue  $\alpha$ 1-6-linked to the inner GlcNAc and a ( $\beta$ 1-2-linked) xylose residue attached to the  $\beta$ -Man residue.<sup>28</sup> Furthermore, an extension of the trimannosyl-N,N'-diacetylchitobiose core element for N-glycosylation by one or two antennas has been described.<sup>29</sup> The predominant antenna corresponds to a pentasaccharide consisting of a central GalNAc moiety substituted with (i) 3-O-methyl-D-galactose and (ii) a disaccharide of 3-O-methyl-D-galactose linked to GlcNAc at the reducing end.<sup>29,45</sup> Researches on the *Lymnea stagnalis*, *Cepaea hortensis*, *Planorbarius corneus*, *Arianta arbustorum* and *Achatina fulica* confirmed  $\beta$ 1-2-linked Xyl being a typical component of snail N-glycans.<sup>29,30,45</sup>

On the basis of this data and results from mass spectrometric analysis, we have determined the carbohydrate compositions of N-glycans shown in Table 1. Our analysis of the glycans of  $\beta$ -HIH revealed that the sugar chains also are modified at the inner core by xylose  $\beta$ 1-2-linked to  $\beta$ -mannose. We also demonstrated the attachment of an additional Fuc residue to position 6 of the Asn-bound GlcNAc and the presence of MeHex residues (Table 1).

Our study shows the presence of monoantennary and diantennary N-glycans with partly or fully methylated terminal hexoses, which for reasons of similarity with HpH glycans<sup>28,29</sup> are likely 3MeMan and/or 3MeGal residues. We also found hybrid structures, similar to those occurring in mammals<sup>48</sup> (glycan No. 15) and KLH<sup>9</sup> (glycan

No. 16). Besides a common core with an  $\alpha$ 1-6-linked fucose on the reducing GlcNAc and a  $\beta$ 1-2-linked xylose linked to  $\beta$ -mannose, we inferred, by similarity, that one of the other two  $\alpha$ -mannose residues might be substituted with GalNAc( $\beta$ 1-4)GlcNAc( $\beta$ 1-2) elements containing two to four  $\beta$ 1-3- or  $\beta$ 1-6-linked hexoses (mainly galactoses) with or without 3- or 4-O-methyl groups.<sup>28,29,45</sup> The 3-O-MeGal( $\beta$ 1-3)GalNAc( $\beta$ 1-4)GlcNAc( $\beta$ 1-2) sequence from HpH<sup>29</sup>, and *L. stugnulis* hemocyanin<sup>30</sup> are also observed in  $\beta$ -HIH (e.g. glycans Nos. 17, 21), as well as the 3-O-MeGal( $\beta$ 1-6)GalNAc( $\beta$ 1-4)GlcNAc( $\beta$ 1-2) structural fragment (similarly to glycans Nos. 21-27), but some display different motifs, containing MeHex( $\beta$ 1-4)GlcNAc( $\beta$ 1-2)Man( $\alpha$ 1-3) (glycans Nos. 12 and 14), Hex( $\beta$ 1-3)GalNAc( $\beta$ 1-4)GlcNAc( $\beta$ 1-2) (glycans Nos. 21, 24, 30 and 28), MeHex( $\beta$ 1-3)-[Hex( $\beta$ 1-6)-]-MeHex( $\beta$ 1-3)GalNAc (glycans Nos. 26 and 29), and Hex( $\beta$ 1-3)[Hex( $\beta$ 1-6)]GalNAc( $\beta$ 1-4)GlcNAc( $\beta$ 1-2) (glycan No. 28).

Some of the N-glycan structures of  $\beta$ -HIH are similar to glycans from  $\alpha$ -HpH (for example glycans Nos. 1, 4, 9, 18, 22, 25, 27, and 31)<sup>28,29</sup>, but there are also significant differences. We have established several N-glycans of  $\beta$ -HIH containing core-fucosylation without Xyl (glycans Nos. 11, 12, 13, 15, 17 and 23), which were not earlier observed in glycan structures in HpH.<sup>28,29</sup> In our preliminary experiments we checked for the possibility of the presence of hexuronic acid residues instead of methylhexose residues in  $\beta$ -HIH glycans, as we have found in RvH.<sup>18,20,38</sup> Because the mass difference between MeHex and HexA is only 0.036 Da, we differentiated methylated structures from acidic glycans containing HexA by using CE-MS, as in electrophoresis they migrate more slowly than their neutral counterparts. The N-glycans of  $\beta$ -HIH do not migrate as negatively charged glycans during electrophoresis, in contrast to the glycans of RvH. The presence of hexuronic acid instead of methylhexose is therefore unlikely. Particularly, an N-glycan with an internal fucose residue and containing hexuronic acid, as was detected in RvH,<sup>18,20,38</sup> seems to be absent from  $\beta$ -HIH. Besides, in contrast to KLH<sup>51</sup> and HtH1<sup>35</sup> and similarly to HpH,<sup>28,29,45</sup> neither LewisX-containing structures nor terminal Fuc( $\alpha$ 1-3)GalNAc structural motif have been detected in  $\beta$ -HIH.

Core fucosylated N-glycans are widely distributed in a variety of glycoproteins but little information is available about  $\alpha$ 1-6-fucosylation in land snails.<sup>28,29,43,45,48,50</sup> Core-fucosylated modifications are related to several physiological and pathological issues, e.g. they play important roles in cell signaling, and the enzymatic activity and protein

expression of FUT8 are increased in tumor tissues of human colorectal carcinoma.<sup>52</sup> Fucosylation also seems to play a role during apoptosis. Increased levels of fucose residues and changes in fucosylation patterns can act as specific markers for developmental antigens, particularly in inflammatory processes and cancers<sup>9,8,46,53</sup>, and have an effect on the efficacy of medicinal antibodies.<sup>11,13,51</sup>

We also found a plenty of core-xylosylated structures comprising a xylose residue  $\beta$ 1-2-linked to  $\beta$ -mannose of the inner core (glycans Nos. 1, 2, 4, 5, 7- 10, 14, 18-22, 24-27, 29-32, Table 1). N-Glycans comprising  $\beta$ 1-2-linked xylose residues were also found in HpH and *L. stagnalis* haemocyanin as major carbohydrate constituents.<sup>28-30</sup> It has been demonstrated that a xylose residue  $\beta$ 1-2-linked to a  $\beta$ -mannose unit of the core of an N-linked carbohydrate chain is a highly immunogenic epitope for mammalian species.<sup>43,50</sup> In contrast to the glycan structures of  $\alpha$ <sub>D</sub>-HpH, for which it has been shown that when xylose is absent, fucose is absent as well<sup>29</sup>, we found six completely different core-fucosylated structures without xylose residue in  $\beta$ -HIH (glycans Nos.11-13, 15, 17, 23).

Methylation of sugar residues is an unusual modification. As far as we know, mammals do not carry this glycan modification but many other organisms such as nematodes and molluscs do. The methylated structures are heterogenous in terms of involved monosaccharides and the position of methylation. Most of the analyzed glycans of  $\beta$ -HIH contain mainly a terminal and/or inner MeHex residue; in some cases even several such residues are present, e.g. up to 4 of them in the glycans Nos. 31 and 27. Methylation appears to play a role in some recognition phenomena, but details remain unknown.<sup>43,46,50,54</sup> Highly methylated complex structures have also been recognized in many glycan structures of gastropod hemocyanins, suggesting that this glycan modification is an important feature in these organisms.<sup>8,9,28-30,35-37,39,43,50</sup>

### 3. Conclusions

An extremely diverse set of 32 N-glycans (high-mannose, complex and hybrid types) has been isolated from the structural  $\beta$ -subunit of *H. lucorum* hemocyanin and identified. It represents a highly heterogeneous mixture of glycans with the compositions Hex<sub>3-7</sub>HexNAc<sub>2-5</sub>MeHex<sub>0-4</sub>Pent<sub>0-1</sub>Fuc<sub>0-1</sub>. The glycans have predominantly monoantennary and biantennary structures of the complex type containing MeHex units.

Hybrid type structures (some of them with  $\alpha$ 1-6-fucosylation of the Asn-bound GlcNAc) have been surprisingly found in the molluscan hemocyanins too.

As in most molluscan hemocyanins, the glycans of  $\beta$ -HIH mainly contain a terminal MeHex, in some cases even several MeHex residue. Further modifications of the core structure include a substitution of the central Man at position 2 by Xyl and/or the attachment of an additional Fuc residue to the innermost GlcNAc. We have detected several glycans, which are both core-fucosylated and have a high degree of methylation. Similar structures are not observed in *H. pomatia* hemocyanin. The isolation and characterization of the N-linked glycans performed in this study revealed in part novel structural motifs which might contribute to the pronounced immunogenicity of this gastropod glycoprotein. These and published data show that gastropods have a wide capacity to modify the basic biantennary N-glycan structure with many species-specific peculiar structures.

The structural model of the functional unit  $\beta$ -HIH-g, based on the model of the functional unit OdH-g, demonstrates that the glycans and the putative glycosylated sites are exposed on the surface in both domains of FU-g. Such site is missing in FU-b and FU-c, but there are three putative sites in the functional units FU-h and FU-d, and two in FU-a, FU-f and FU-g. As the glycans exposure in FU-h may prevent the formation of larger aggregates, we suggest that the native complex of  $\beta$ -HIH is not able to attach further decamers to its didecameric core structure.

Different species of molluscan hemocyanins have been considered for use in diverse biomedical and clinical applications. In particular, the glycan moieties play diverse roles in biological systems that make them relevant for use as biotherapeutics. Medical investigations of gastropod glycoproteins can be subdivided into two groups: one in which interaction of an intermediate host gastropod with a parasite is analysed, and the second one which comprises studies for their use in cancer and antiviral therapy. It is known for about twenty years that snails and parasites share some epitopes, which cause cross-reactivity *in vitro* and antibody production *in vivo*.<sup>9,11,13,14,31,54,55</sup> This reaction can be used on the one hand for diagnosis, and on the other hand for the design of drugs against the parasites.<sup>32,39,51,54,55</sup> Some gastropod glycoproteins bind specifically to certain types of cancer cells, enabling diagnosis and prognosis.<sup>46</sup> Natural or modified gastropod glycoproteins are also used in the fight

against cancer cells by stimulating the human immune response.<sup>3-7,10-16,31,32</sup> Our results may be considered as the basic information for further investigations on *H. lucorum* hemocyanin and its potential applications in therapy.

## 4. Experimental

### 4.1. Isolation of glycans from structural subunits of HIH

Intact structural subunit  $\beta$ -HIH was obtained from the whole HIH as described.<sup>40</sup> For deglycosylation, approximately 4 mg of  $\beta$ -HIH were dissolved in 50  $\mu$ l of denaturing solution (1% SDS, 0.5 M mercaptoethanol, 0.1 M EDTA), followed by incubation at 25 °C for 30 min. A volume of 300  $\mu$ l of Na-phosphate buffer (200 mM, pH 8.6) was added and the solution was heated in a boiling water bath for 5 min. After cooling to 25 °C, 50  $\mu$ l of Triton X100 and 5  $\mu$ l of PNGase F (2 units) (Roche Diagnostics GmbH, Mannheim, Germany) were added. This mixture was incubated during 20 h at 37 °C. The released N-glycans were isolated by solid phase extraction on a Carbograph column (Alltech, Lokeren, Belgium), and the glycans were eluted with 2 ml of 25% acetonitrile/0.05% TFA. The collected fraction was dried and dissolved in 20  $\mu$ l H<sub>2</sub>O (stock solution) for further analyses.

### 4.2. MALDI-TOF/TOF-MS analysis of glycans

The isolated oligosaccharides were analysed by MALDI-TOF-MS. The matrix was a dihydroxybenzoic acid solution in 50% ACN (10 mg.ml<sup>-1</sup>). The analyses were carried on a 4700 Proteomics Analyser with TOF/TOF optics (Applied Biosystems, Framingham, MA). The mass spectrometer had a 200 Hz frequency-tripled Nd-YAG laser operating at a wavelength of 355 nm. A total of 1500 shots were acquired in the MS mode. Spectra from  $m/z$  900 to 3000 were recorded. Glycans were detected as [M+Na]<sup>+</sup> ions.

### 4.3. MS and MS/MS Q-Trap analysis of glycans

Off-line ESI-MS and MS/MS measurements of the glycans were performed on an Absciex 4000 Q-Trap mass spectrometer (Applied Biosystems), equipped with a nanospray ion source (Proxeon, Odense, Denmark) and using Proxeon medium nanospray needles. Typically, 10  $\mu$ l of sample in 50% MeOH was introduced. The needle voltage was set at 1000 V. In the product ion-scanning mode, the scan speed

was set to 1000 Da/s, with Q-trapping being activated. The trap fill-time was 200 ms in the MS/MS scan mode. The resolution of Q1 was set to 'low'. The excitation time was set at 100 ms.

#### *4.4. 3D modelling and glycan structure localization*

The model of the tertiary structure of functional unit  $\beta$ -HIH-g of *H. lucorum* hemocyanin<sup>41</sup> was generated based on the X-ray structure of functional unit 'g' of *Octopus dofleini* hemocyanin<sup>56</sup> using SwissPdb Viewer.<sup>57</sup> Potential N-glycosylation sites were identified by their consensus sequences NXT or NXS.

## References

1. Van Holde, K.E.; Miller, K.I.; Decker, H. *J. Biol. Chem.* **2001**, *276*, 15563-15566.
2. Salvato, B.; Beltramini, M. *Life Chem. Rep.* **1990**, *8*, 1-47.
3. Linn, J.; Black, P.; Derksen, K.; Rubben, H.; Thuroff, J.W. *Eur. Urol.* **2000**, *37*, 34-40.
4. McFadden, D.W.; Riggs, D.R.; Jackson, B.J.; Vona-Davis, L. *Am. J. Surg.* **2003**, *186*, 552-555.
5. Krug, L.M.; Ragupathi, G.; Hood, C.; Kris, M.G.; Miller, V.A.; Allen, J.R.; Keding, S.J.; Danishefsky, S.J.; Gomez, J.; Tyson, L.; Pizzo, B.; Baez, V.; Livingston, P.O. *Clin. Cancer Res.* **2004**, *10*, 6094-6100.
6. Lamm, D.L. *J Urol. Urogynäkol.* **2003**, *10*, 18-21.
7. Lammers, R.J.; Witjes, W.P.; Janzing-Pastors, M.H.; Caris, C.T.; Witjes, J.A. *J. Clin. Oncol.* **2012**, *30*, 2273-9.
8. Wuhrer, M.; Robijn, M.L.M.; Koeleman, C.A.M.; Balog, C.I.A.; Geyer, R.; Deelder, A.M.; Hokke, C.H. *Biochem. J.* **2004**, *378*, 625-632.
9. Kurokawa, T.; Wuhrer, M.; Lochnit, G.; Geyer, H.; Markl, J.; Geyer, R. *Eur. J. Biochem.* **2002**, *269*, 5459-5473.
10. Moltedo, B.; Faunes, F.; Haussmann, D.; De Ioannes, P.; De Ioannes, A.E.; Puente, J. et al. *J. Urol.* **2006**, *176*, 2690-2695.
11. Dolashka-Angelova, P.; Stefanova, T.; Livanidou, E.; Velkova, L.; Klimentzou, P.; Stevanovic, S.; Salvato, B.; Neychev, H.; Voelter, W. *Immunol. Invest.*, **2008**, *37*, 822–840.
12. Dolashka, P.; Velkova, L.; Iliev, I.; Beck, A.; Dolashki, A.; Yossifova, L. et al. *Immunol. Invest.* **2011**, *40*, 130-149.
13. Iliev, I.; Toshkova, R.; Dolashka-Angelova, P. et al. *Compt. Rend. Acad. Bulg. Sci.* **2008**, *61*, 203-210.
14. Del Campo, M.; Arancibia, S.; Nova, E.; Salazar, F.; González, A.; Moltedo, B. et al. *Rev. Med. Chil.* **2011**, *139*, 236-246.
15. Boyanova, O.; Dolashka, P.; Toncheva, D.; Rammensee, H.G.; Stevanović, S. *Biomed. Rep.* **2013**, *1*, 235-238.
16. Antonova, O.; Dolashka, P.; Toncheva, D.; Rammensee, H.G.; Floetenmeyer, M.; Stevanovic, S. *Z. Naturforsch. C* **2014**, *69*, 325-34.



17. Dolashka-Angelova, P.; Lieb, B.; Velkova, L.; Heilen, N.; Sandra, K.; Nikolaeva-Glomb, L. et al. *Bioconjug. Chem.* **2009**, *20*, 1315-1322.
18. Dolashka, P.; Velkova, L.; Shishkov, S.; Kostova, K.; Dolashki, A.; Dimitrov, I.; Atanasov, B.; Devreese, B.; Voelter, W.; Van Beeumen, J. *Carbohydr. Res.* **2010**, *345*, 2361-2367.
19. Velkova, L.; Nikolaeva-Glomb, L.; Mukova, L.; Dolashki, A.; Dolashka, P.; Galabov, A.S. *Antivir. Res.* **2011**, *90*, A47.
20. Velkova, L.; Todorov, D.; Dimitrov, I.; Shishkov, S.; Van Beeumen, J.; Dolashka-Angelova, P. *Biotech. Biotech. Equip.* **2009**, *23*, 606-610.
21. Nesterova, N.; Zagorodnya, S.; Moshtanska, V.; Dolashka, P.; Baranova, G.; Golovan, A.; Kurova, A. *Antivir. Res.* **2011**, *90*, A38.
22. Zagorodnya, S.; Dolashka, P.; Baranova, G.; Golovan, A.; Nesterova, N. *Antivir. Res.* **2011**, *90*, A66.
23. Dang, V.T.; Benkendorff, K.; Speck, P. *J. Gen. Virol.* **2011**, *92*, 627-637.
24. Dang, V.T.; Green, K.B.T.; Speck, P. *J. Virol.* **2015**, *89*, 8114-8118.
25. Zanjani, N.T.; Sairi, F.; Marshall, G.; Saksena, M.M.; Valtchev, P.; Gomes, V.G.; Cunningham, A.L.; Dehghani, F. *Eur. J. Pharm. Sci.* **2014**, *53*, 77-85.
26. Van Kuik, J.A.; Kamerling, J.P.; Vliegthart, J.F.G. In *Invertebrate dioxygen carriers*. Préaux G, Lontie R, editors. Leuven, Belgium, Leuven University Press, **1990**, 157-163.
27. Hall, R.L.; Wood, E.J. *Biochem. Soc. Trans.* **1976**, *4*, 307-309.
28. Van Kuik, J.A.; van Halbeek, H.; Kamerling, J.P.; Vliegthart, J.F.G. *J. Biol. Chem.* **1985**, *260*, 13984-13988.
29. Lommerse, J.P.M.; Thomas-Oates, J.E.; Gielens, C.; Préaux, G.; Kamerling, J.P.; Vliegthart, J.F.G. *Eur. J. Biochem.* **1997**, *249*, 195-222.
30. Van Kuik, J.A.; Sijbesma, R.P.; Kamerling, J.P.; Vliegthart, J.F.; Wood, E.J. *Eur. J. Biochem.* **1987**, *169*, 399-411.
31. Wirguin, I.; Suturkova-Milosević L.; Briani, C.; Latov, N. *Cancer Immunol. Immunother.* **1995**, *40*, 307-310.
32. Geyer, H.; Wuhler, M.; Resemann, A.; Geyer, R. *J. Biol. Chem.* **2005**, *280*, 40731-40748.
33. Van Kuik, J.A.; Van Halbeek, H.; Kamerling, J.P.; Vliegthart, J.F. *Eur. J. Biochem.* **1986**, *159*, 297-301.

34. Puanglarp, N.; Oxley, D.; Currie, G.J.; Bacic, A.; Craik, D.J.; Yellowlees, D. *Eur. J. Biochem.* **1995**, *232*, 873-80.
35. Velkova, L.; Dolashka, P.; Lieb, B.; Dolashki, A.; Voelter, W.; Van Beeumen, J.; Devreese, B. *Glycoconj J.* **2011**, *28*, 385-395.
36. Dolashka-Angelova, P.; Beck, A.; Dolashki, A.; Beltramini, M.; Stevanovic, S.; Salvato, B.; Voelter, W. *Biochem. J.* **2003**, *374*, 185-192.
37. Dolashka-Angelova, P.; Beck, A.; Dolashki, A.; Stevanovic, S.; Beltramini, M.; Salvato, B.; Hristova, R.; Velkova, L.; Voelter, W. *Micron* **2004**, *35*, 101-104.
38. Sandra, K.; Dolashka-Angelova P.; Devreese B.; Van Beeumen, J. *Glycobiology*, **2007**, *17*, 141-156.
39. Lehr, T.; Geyer, H.; Maass, K.; Doenhoff, M.J.; Geyer, R. *Glycobiology*, **2007**, *17*, 82–103.
40. Velkova, L.; Dimitrov, I.; Schwarz, H.; Stevanovic, S.; Voelter, W.; Salvato, B.; Dolashka-Angelova, P.: *Comp. Biochem. Physiol. Part B: Biochem. Mol. Biol.* **2010**, *157*, 16–25.
41. De Smet, L.; Dimitrov, I.; Debyser, G.; Dolashka- Angelova, P.; Dolashki, A.; Van Beeumen, J. et al. *Gene*, **2011**, *487*, 118-128.
42. Velkova, L.; Dolashki, A.; Dolashka, P. *J. BioSci. Biotechnol.*, **2015**, *se/online*: 75-85.
43. Gutternigg, M.; Ahrer, K.; Grabher-Meier, H.; Burgmayr, S.; Staudacher, E. *Eur. J. Biochem.*, **2004**, *271*, 1348-1356.
44. Domon, B.; Costello, C.E. *Glycoconj. J.* **1988**, *5*, 397-409.
45. Siddiqui, N.I.; Yigzaw, Y.; Préaux, G.; Gielens C. *Biochimie.* **2009**, *91*, 508-516.
46. Staudacher, E.; Stepan, H.; Gutternigg, M. *Curr. Top. Biochem. Res.* **2009**, *11*, 29-39.
47. Staudacher, E.; Altmann, F.; Wilson, I.B.; Marz, L. *Biochim. Biophys. Acta*, **1999**, *1473*, 216-236.
48. Kronman, C.; Chitlaru, T.; Elhanany, E.; Velan, B.; Shafferman, A. *J. Biol. Chem.* **2000**, *275*, 29488-29502
49. Kostadinova, E.; Dolashka, P.; Velkova, L.; Dolashki, A.; Stevanovic, S.; Voelter, W. *J. Fluoresc.* **2013**, *23*, 753-760.
50. Gutternigg, M.; Bürgmayr, S.; Pörtl, G.; Rudolf, J.; Staudacher, E.: *Glycoconj. J.* **2007**, *24*, 475-489.
51. Kantelhardt, S.R.; Wuhler, M.; Dennis R.D.; Doenhoff, M.J.; Bickle, Q.; Geyer, R. *Biochem. J.* **2002**, *366*, 217-223.

52. Muinelo-Romay, L.; Vázquez-Martín, C.; Villar-Portela, S.; Cuevas, E.; Gil-Martín, E.; Fernández Briera, A. *Int. J. Cancer*. **2008**, *123*, 641-646.
53. Varki, A.; Kannagi, R.; Toole, B.P. In *Essentials of Glycobiology*. 2nd edition, Varki A.; Cummings R.D.; Esko J.D.; Freeze H.H.; Stanley P.; Bertozzi C.R.; Hart G.W.; Etzler M.E. editors. Cold Spring Harbor Laboratory Press (NY), **2009**, 580-670.
54. Schiller, B.; Hykollari, A.; Yan, S.; Paschinger, K.; Wilson, I.B. *Biol. Chem*. **2012**, *393*, 661-673.
55. Thors, C.; Linder, E. *J. Histochem. Cytochem.*, **2003**, *51*, 1367–1373.
56. Cuff, M.E.; Miller, K.I.; van Holde, K.E.; Hendrickson, W.A. *J. Mol. Biol*. **1998**, *278*, 855-870.
57. Guex, N.; Peitsch, M.C. *Electrophoresis*, **1997**, *18*, 2714–2723.



**Table 1.** Oligosaccharide structures of N-glycans isolated from structural subunit  $\beta$ -HIH. The residue HexNAc in the glycans Nos. 17 to 32 is assigned as GalNAc on the basis of the similarity with hemocyanin of *H. pomatia*.<sup>29</sup>

### Legends to Figures

**Figure 1. a)** MALDI-TOF-MS spectrum of the neutral N-linked glycans from  $\beta$ -HIH released by treatment with PNGase F. The spectrum was recorded using 2,5-DHB as the matrix. Each of the glycans is detected as the  $[M+Na]^+$  ion. **b)** The monosaccharides are presented according to the nomenclature of the Consortium for Functional Glycomics; Me indicates 3-O-methyl/4-O-methyl group. Key symbols used for the constituent monosaccharides: ●=Man; Me○=methylated hexose (MeHex); ■=GalNAc; ■=GlcNAc; □=HexNAc; ▲=Fuc; ★ =Xyl.

**Figure 2. a)** MS/MS spectrum and structure with fragmentation pattern of the  $[M+Na]^+$  ion at  $m/z$  1079.1 of the glycan No. 2 of  $\beta$ -HIH (MeHex<sub>1</sub>Man<sub>2</sub>GlcNAc<sub>2</sub>Xyl<sub>1</sub>). Collision energy (CE) was set at 40 eV. Fragments are assigned using the Domon/Costello nomenclature.<sup>44</sup> **b)** Scheme showing the main cleavage sites and cross-ring fragment ions of the oligosaccharide chain. The methylated hexose is presented as methylated mannose.

**Figure 3.** MS/MS spectrum and structure with fragmentation pattern of the double charged  $[M+2Na]^{2+}$  ion at  $m/z$  1092.3 of the glycan No. 26 with composition MeHex<sub>3</sub>Hex<sub>1</sub>HexNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>Xyl<sub>1</sub> (CE 60 eV). **b)** Scheme showing the main cleavage sites and cross-ring fragment ions of a part of the oligosaccharide chain at the non-reducing end.

**Figure 4. a)** MS/MS spectrum and structure with fragmentation pattern of the double charged  $[M+2Na]^{2+}$  ion of the glycan No. 24 at  $m/z$  1077.2 with composition MeHex<sub>2</sub>Hex<sub>1</sub>HexNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>Fuc<sub>1</sub>Xyl<sub>1</sub> (CE 60 eV). **b)** Scheme showing the main cleavage sites and cross-ring fragment ions. Methylated hexoses and hexose at the non-reducing end are presented as methylated galactose and galactose, respectively.

**Figure 5.** a) MS/MS spectrum and structure of the  $[M+Na]^+$  ion at  $m/z$  1606.5 of the glycan No. 15 with composition  $Hex_2HexNAc_1Man_3GlcNAc_2Fuc_1$  (CE 50 eV). b) Molecular structure with the main fragment ions and cross-ring fragment ions. Two terminal hexose residues at the non-reducing end are presented as mannose.

**Figure 6.** 3D-model of functional unit  $\beta$ -HIH-g, created by using the Swiss PDB viewer and the model of functional unit 'g' from *Octopus dofleini* hemocyanin. Glycans and the putative glycosylated sites N125 and N372 are presented as balls.<sup>49</sup>

# Table

№	Glycan structures in $\beta$ - HIH	Q-Trap/MS (m/z)	MALDI-TOF/MS (m/z)
1		[M+Na] <sup>+</sup> 1065,3	[M+Na] <sup>+</sup> 1065,4
2		[M+Na] <sup>+</sup> 1079,1	[M+Na] <sup>+</sup> 1079,2
3	MeMan +	[M+Na] <sup>+</sup> 1109,2	[M+Na] <sup>+</sup> 1109,3
4		[M+Na] <sup>+</sup> 1211,3	[M+Na] <sup>+</sup> 1211,4
5	Man +	[M+H] <sup>+</sup> 1205,4	[M+Na] <sup>+</sup> 1227,4
6	Man +	[M+H] <sup>+</sup> 1235,3	[M+Na] <sup>+</sup> 1257,4
7		[M+Na] <sup>+</sup> 1268,5	[M+Na] <sup>+</sup> 1268,4
8	Man +	[M+Na] <sup>+</sup> 1367,5	[M+Na] <sup>+</sup> 1389,4
9		[M+Na] <sup>+</sup> 1414,3	[M+Na] <sup>+</sup> 1414,4
10		[M+H] <sup>+</sup> 1408,3	[M+Na] <sup>+</sup> 1430,4
11	Hex -	[M+H] <sup>+</sup> 1422,3	[M+Na] <sup>+</sup> 1444,4
12	MeHex -	[M+Na] <sup>+</sup> 1458,3	[M+Na] <sup>+</sup> 1458,4
13	HexNAc +	[M+Na] <sup>+</sup> 1485,4	[M+Na] <sup>+</sup> 1485,5
14	MeHex -	[M+H] <sup>+</sup> 1568,2	[M+Na] <sup>+</sup> 1590,3
15		[M+Na] <sup>+</sup> 1606,5	[M+Na] <sup>+</sup> 1606,6
16	Hex +	[M+Na] <sup>+</sup> 1622,4	[M+Na] <sup>+</sup> 1622,5
17	MeHex -	[M+H] <sup>+</sup> 1639,5	[M+Na] <sup>+</sup> 1661,5
18	MeHex -	[M+2Na] <sup>2+</sup> 908,2	[M+Na] <sup>+</sup> 1793,5

№	Glycan structures in $\beta$ - HIH	Q-Trap/MS (m/z)	MALDI-TOF/MS (m/z)
19	MeHex -	[M+2Na] <sup>2+</sup> 923,2	
20	MeHex -	[M+2Na] <sup>2+</sup> 996,4	
21	MeHex -	[M+2Na] <sup>2+</sup> 1004,3	
22	MeHex -	[M+2Na] <sup>2+</sup> 1011,4	
23	MeHex -	[M+2Na] <sup>2+</sup> 1018,3	
24	MeHex -	[M+2Na] <sup>2+</sup> 1077,2	
25	MeHex -	[M+2Na] <sup>2+</sup> 1084,3	
26	MeHex -	[M+2Na] <sup>2+</sup> 1092,3	
27	MeHex -	[M+2Na] <sup>2+</sup> 1099,3	
28	Hex -	[M+2Na] <sup>2+</sup> 1106,5	
29	Hex -	[M+2Na] <sup>2+</sup> 1158,4	
30	MeHex -	[M+2Na] <sup>2+</sup> 1165,4	
31	MeHex -	[M+2Na] <sup>2+</sup> 1172,4	
32	MeHex -	[M+2Na] <sup>2+</sup> 1185,8	

Figure 1.  
a)

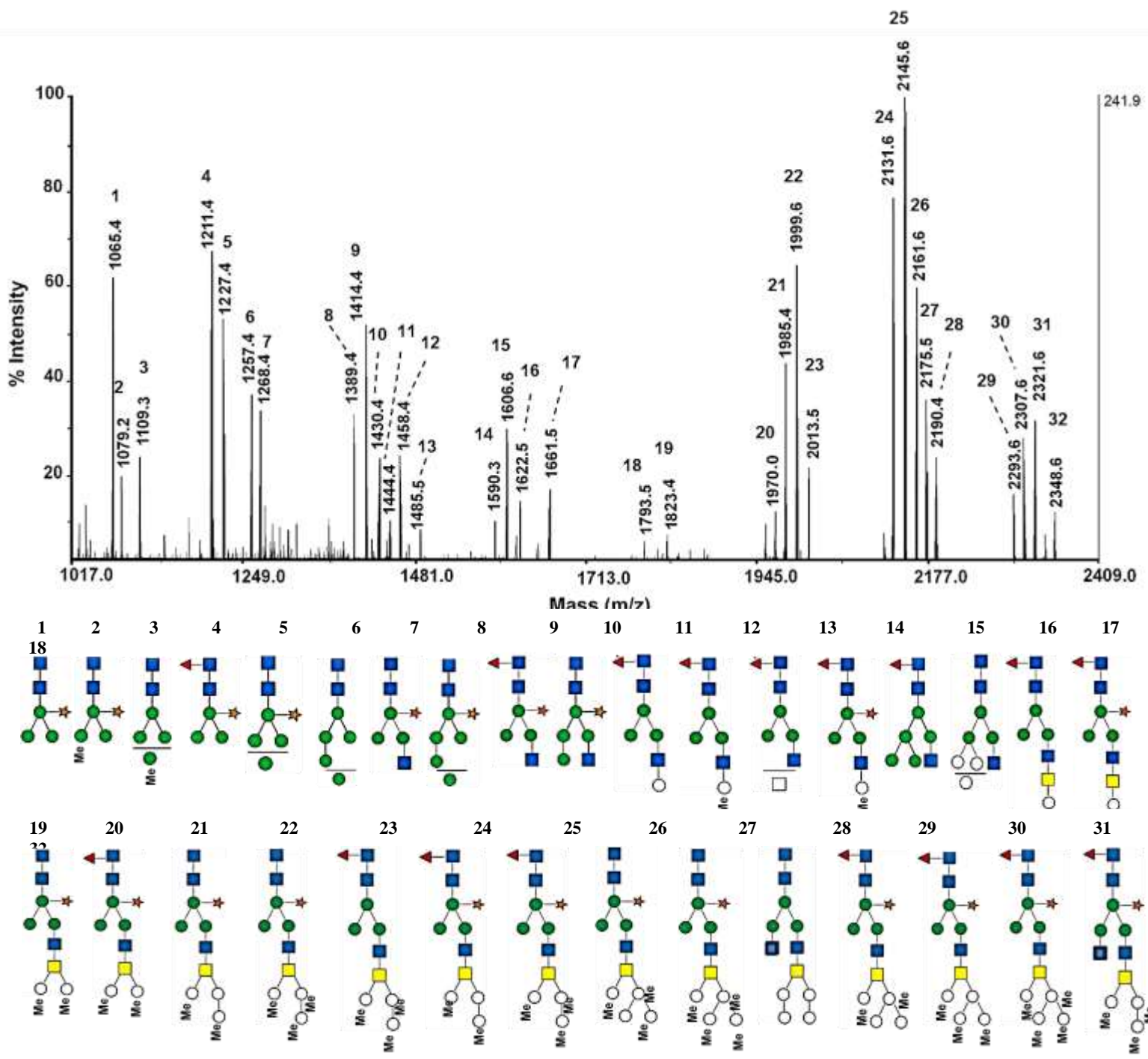




Figure 2.

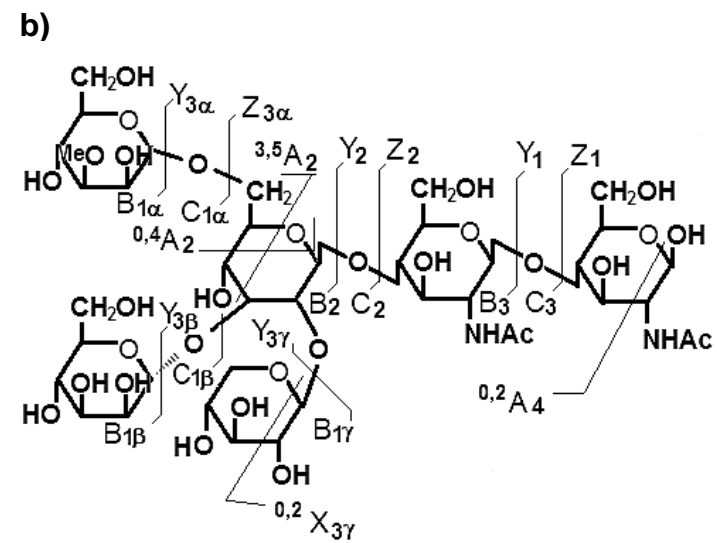
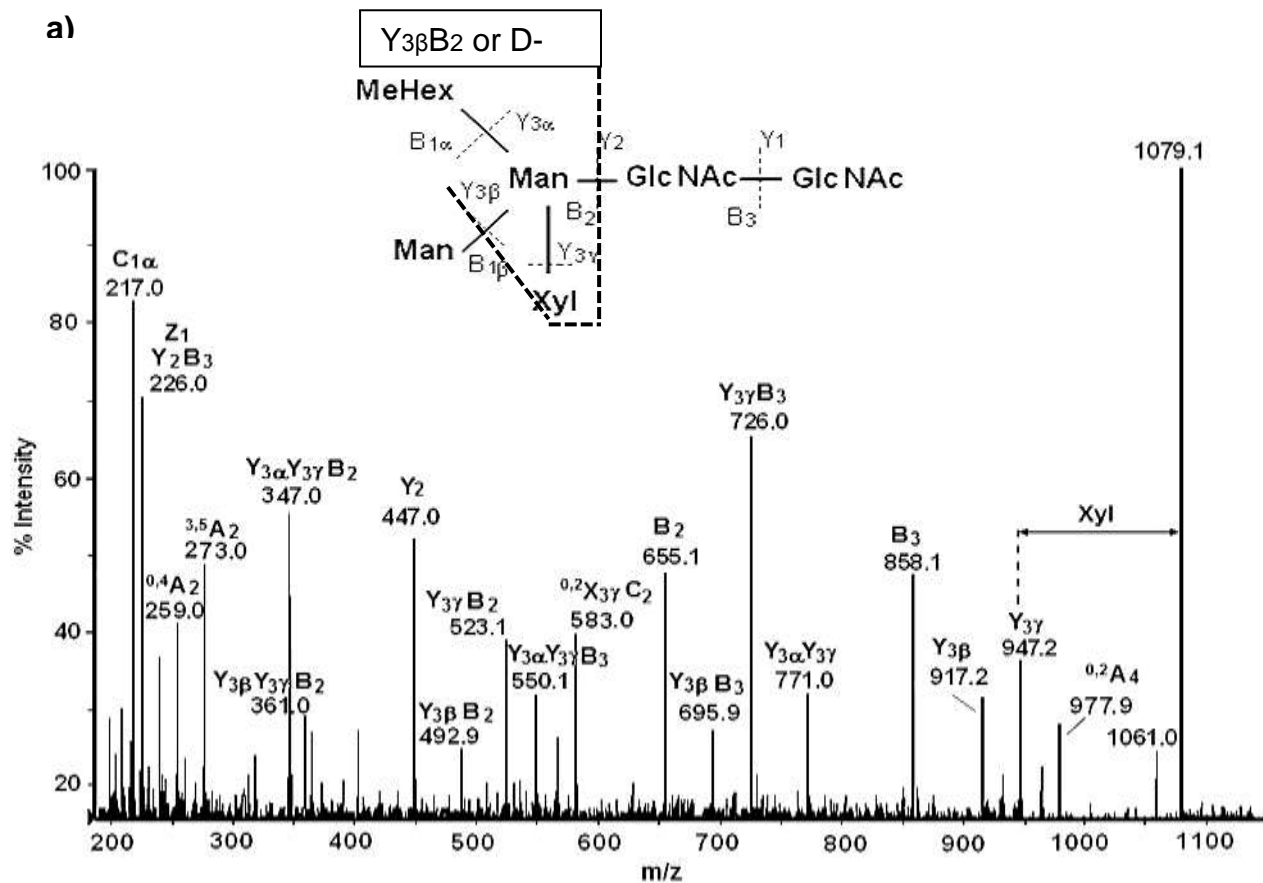
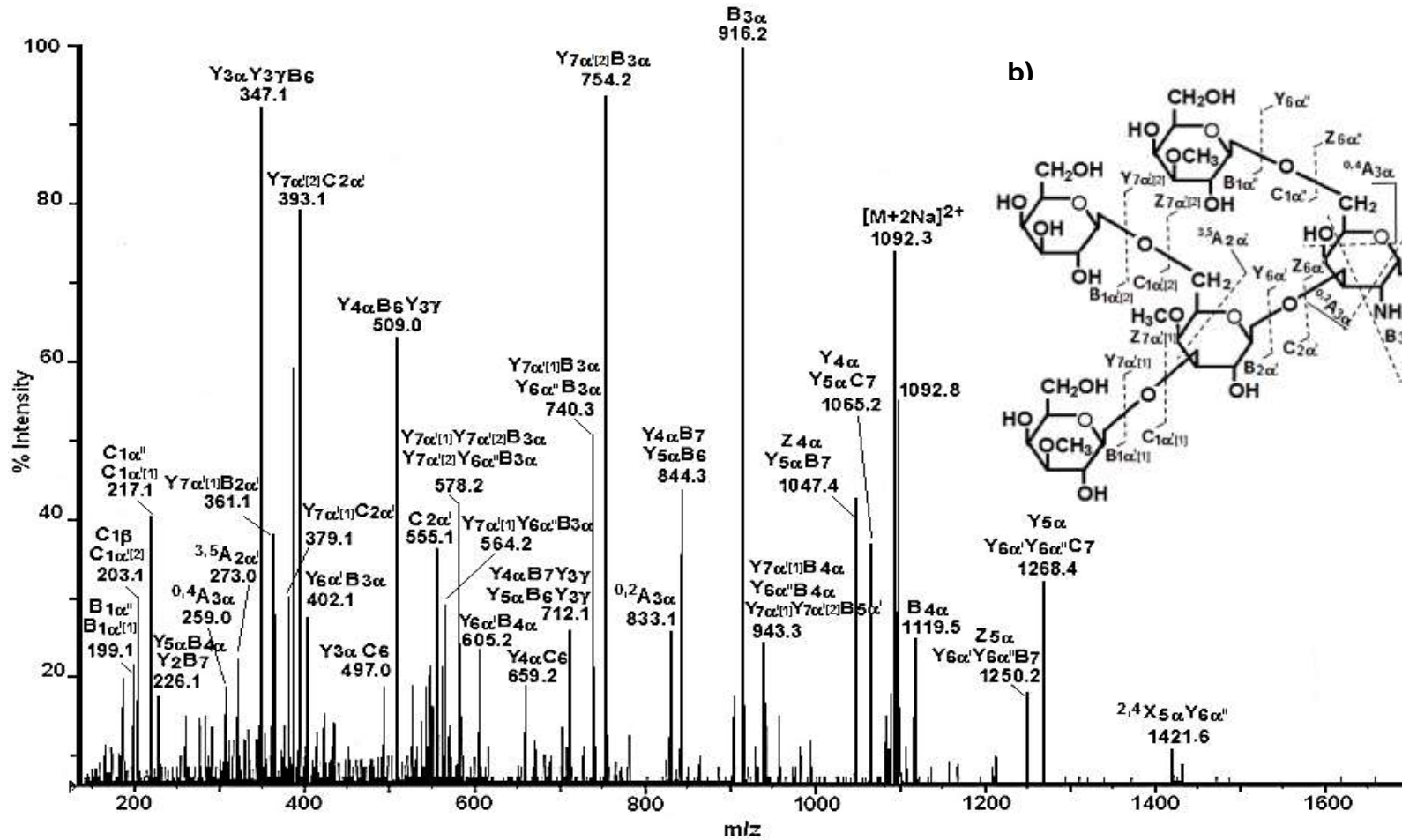
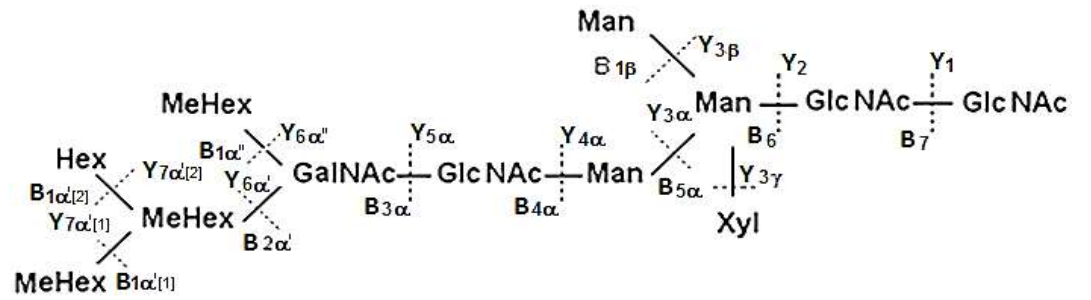


Figure 3.

a)



b)

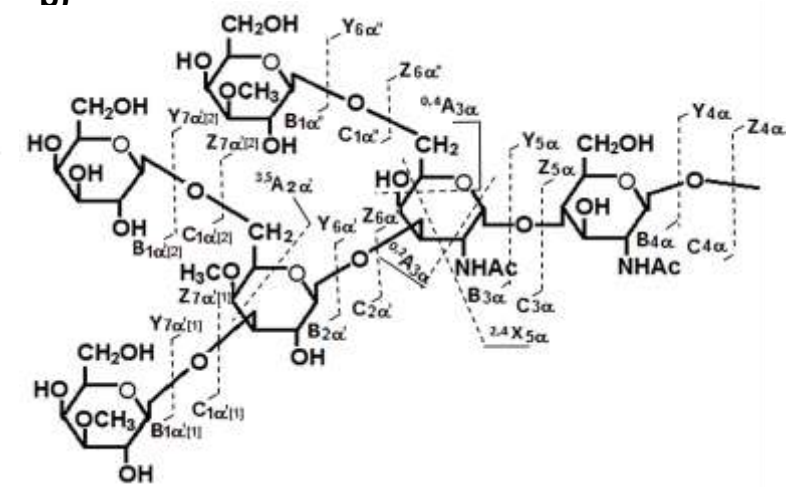
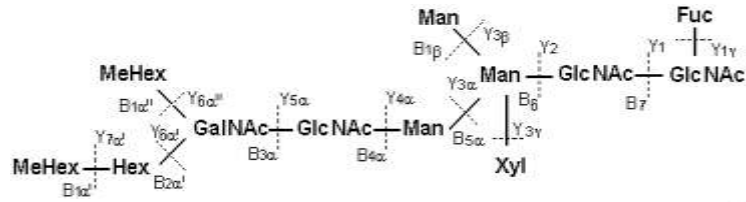


Figure 4.

a)



b)

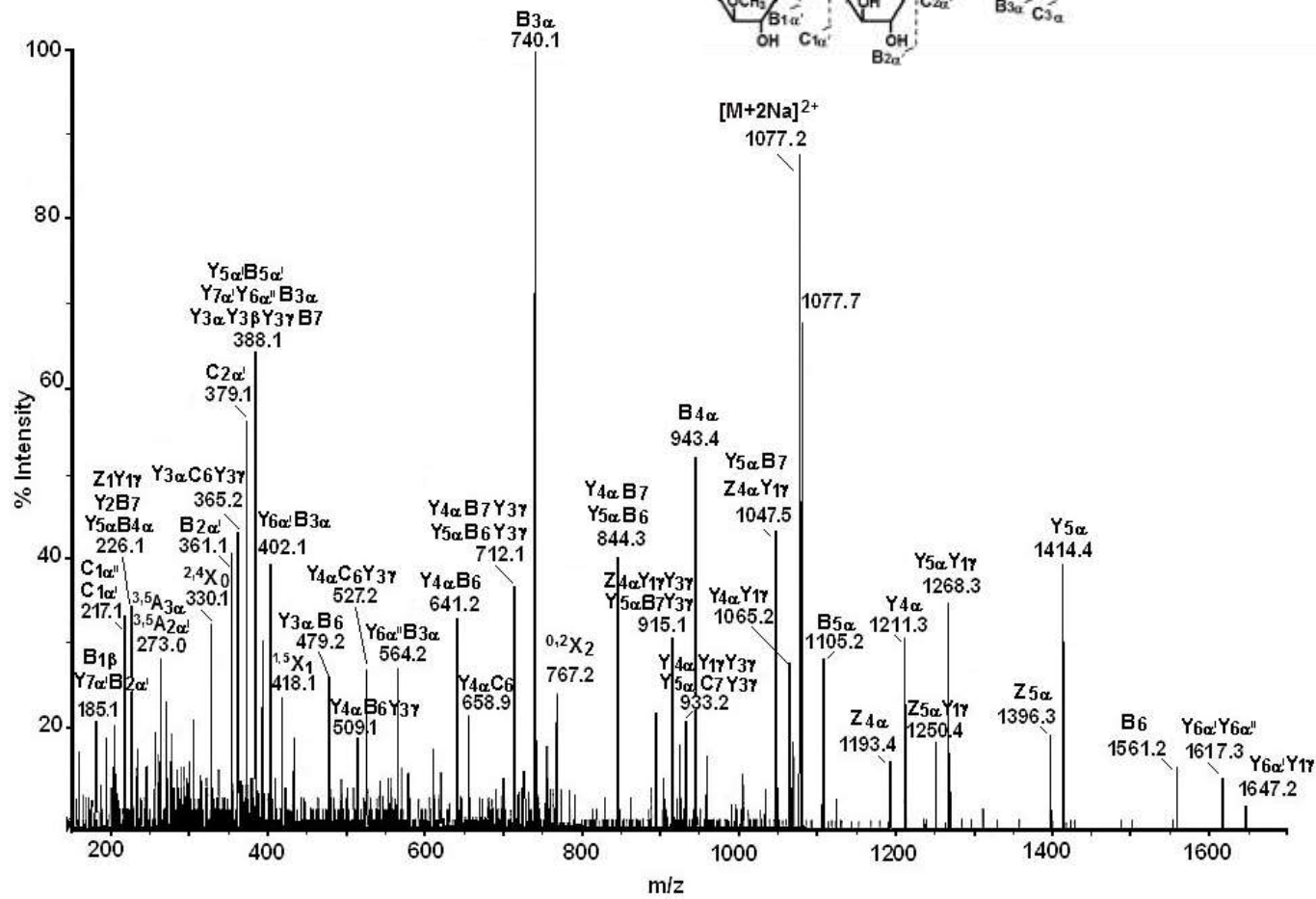
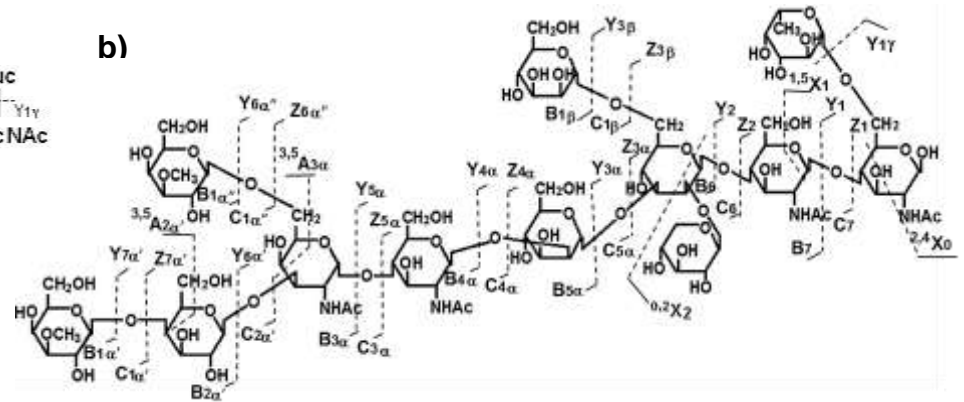


Figure 5.

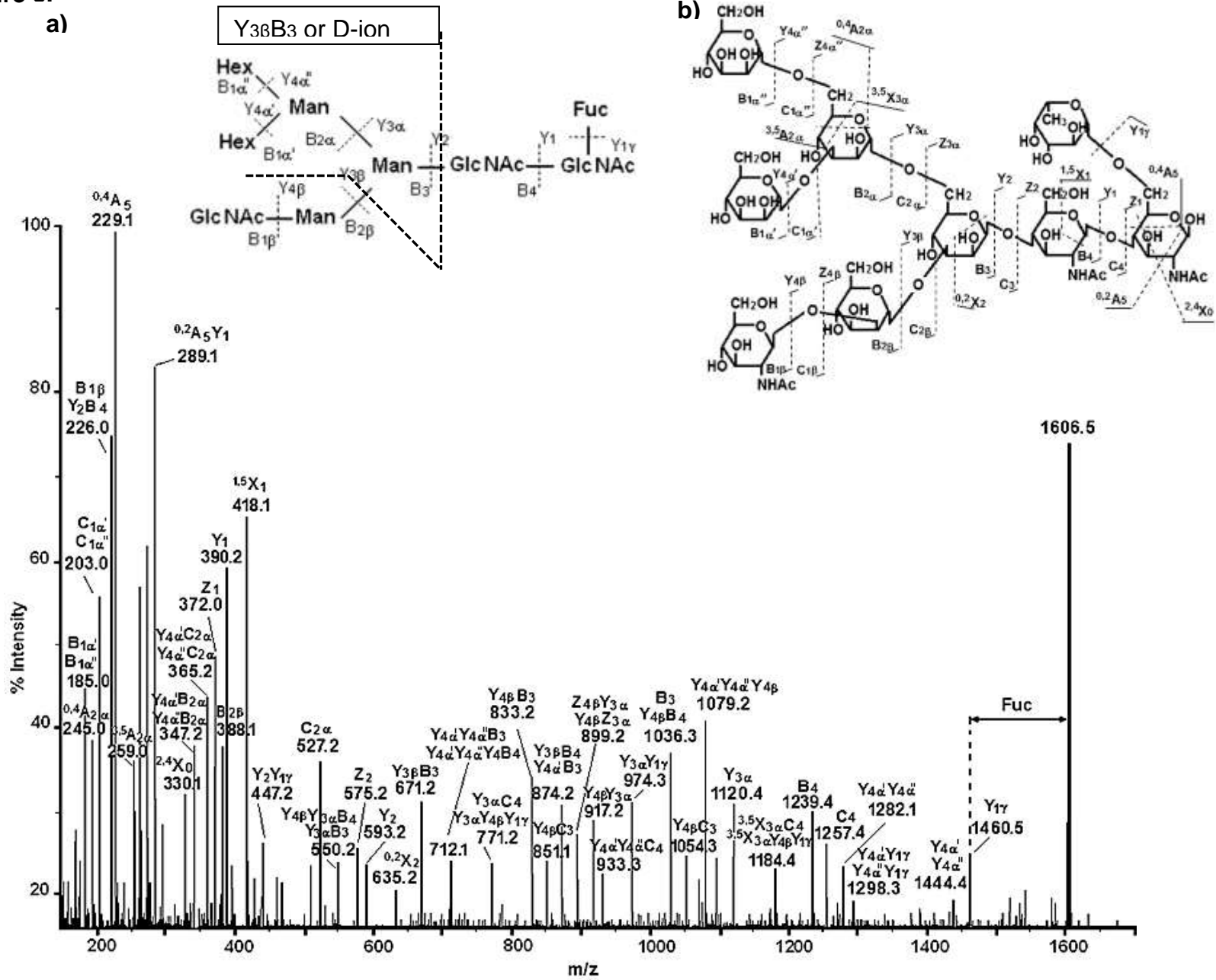


Figure 6.

