Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 2041–2049
Carbohydrate RESEARCH

Minireview

# The N-glycosylation pattern of *Caenorhabditis elegans*

Katharina Paschinger,* Martin Gutternigg, Dubravko Rendić and Iain B. H. Wilson

*Department für Chemie, Universität für Bodenkultur, Muthgasse 18, A-1190 Wien, Austria*

Received 1 October 2007; received in revised form 7 December 2007; accepted 17 December 2007
Available online 28 December 2007

Presented at Eurocarb 14th Lübeck, Germany, September 2007

**Abstract**—Determining the exact nature of N-glycosylation in *Caenorhabditis elegans*, a nematode worm and genetic model organism, has proved to have been an unexpected challenge in recent years; a wide range of modifications of its N-linked oligosaccharides have been proposed on the basis of structural and genomic analysis. Particularly mass spectrometric studies by a number of groups, as well as the characterisation of recombinant enzymes, have highlighted those aspects of N-glycosylation that are conserved in animals, those which are seemingly unique to this species and those which are shared with parasitic nematodes. These data, of importance for therapeutic developments, are reviewed.


*Keywords:* *Caenorhabditis*; Biosynthesis; Fucose; N-Glycans; Mass spectrometry

**Contents**



## 1. Introduction

Although *Caenorhabditis elegans* was the first multicellular organism to have its genome sequenced, it is only in the past few years that information regarding its glycome has been accumulated. Despite the simple anatomy of *C. elegans*, the complexity of the glycans existing in the worm is striking—especially when compared to the glycan pattern present in the fruitfly *Drosophila melanogaster*, another model organism. Since the structural variation is so high, a potential classification of the worm's N-glycans into several classes has been proposed. Thereby, oligomannosidic, paucimannosidic, truncated complex, phosphorylcholine (PC)-containing and fucose-rich N-glycans were defined (see Fig. 1).[1] Furthermore, the O-glycans and glycolipids of *C. elegans* also show structural peculiarities, with 'pseudo-core 2′ O-glycans of the form Galβ1,3(Glcβ1,6)-GalNAc, some modified with fucose, and *arthro*-series glycosphingolipids with a GalNAcβ1,4GlcNAcβ1,

*Abbreviations*: CCM, core chitobiose modification; GnGn, MM, etc., biantennary N-glycans with either terminal GlcNAc or Man residues for symbolic structures see Figure. 4; GnT, *N*-acetylglucosaminyltransferase (GlcNAc-T); LacdiNAc, GalNAcβ1,4GlcNAcβ-R; Lewis x, Galβ1,4(Fucα1,3)GlcNAcβ-R; *fut*-, fucosyltransferase gene or mutant; FUT-, fucosyltransferase protein; PC, phosphorylcholine

* Corresponding author. Tel.: +43 1 36006 6541; fax: +43 1 36006 6076; e-mail: katharina.paschinger@boku.ac.at

0008-6215/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2007.12.018

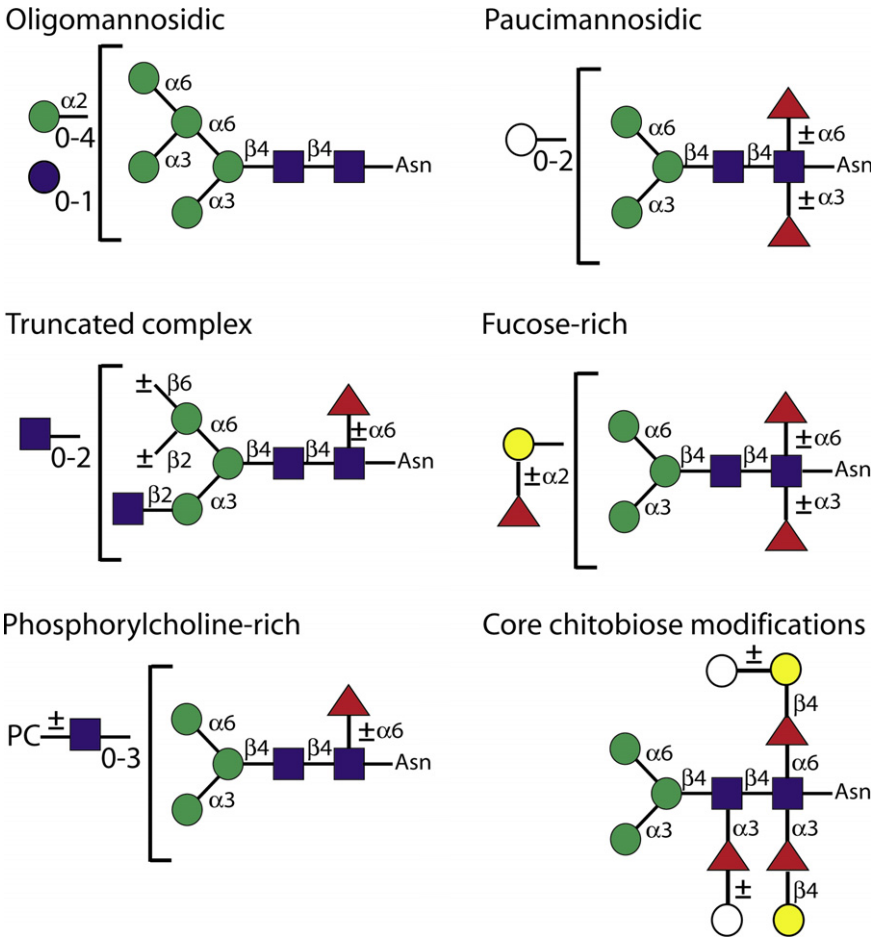


**Figure 1.** Putative N-glycan structures reported as being present in the worm. The symbols used are blue squares for GlcNAc, green circles for Man, yellow circles for Gal, open circles for undefined hexose residues and red triangles for Fuc residues (nomenclature of the Consortium for Functional Glycomics). The figure is based on that of Haslam and Dell,[1] with the exception of the depiction of the core chitobiose modifications (CCMs) with up to probably three Galβ1,4Fuc disaccharides.[10] Potential methylations of, for example, fucose are not shown.

3Manβ1,4Glc-Cer core, some containing also fucose and phosphorylcholine.[2–4]

As a harmless member of a family of organisms (nematodes), which are often causing parasitic diseases in plants, farm animals and man, *C. elegans* is a potent model for studying the genetic basis for glycomic variation in its relatives. In this context, particularly the PC-containing glycans of nematodes have attracted attention due to their immune-modulating function;[5] just as the Tarnhelm made Siegfried invisible in the Nibelungen legend, phosphorylcholine may make nematodes immunologically 'camouflaged'.

## 2. A summary of the N-glycomic data

In total, six laboratories have been involved in the analysis of *C. elegans* glycans with mostly overlapping results.[6–11] An analysis of all published data (summarised in Table 1) demonstrates that *C. elegans* has an extraordinarily diverse N-glycosylation repertoire. Considering only the *m*/*z* values obtained, it appears that nearly 150 structures are present in wild-type (N2) and/or mutant worms (see Table 1[3,6,12–14]).

Since a hexose residue can either be a mannose or a galactose or, in the case of Hex10, even a glucose and that a HexNAc can either be *N*-acetylglucosamine or *N*-acetylgalactosamine, the number of structures, including structural isomers, is enormous and in obvious contrast to the aforementioned relatively simple anatomy or the low position in the phylogenetic tree. Only a few attempts have been made to solve the actual structure of the individual glycans, which is mainly due to the low amounts of these compounds; of those examined by traditional HPLC-based techniques, in general only the most abundant, but less interesting, were subject to a more intensive analysis.[7,15]

The most abundant class of glycans in *C. elegans*, and at the same time the most conserved amongst eukaryotes, are the high mannose type glycans (i.e., the oligomannosidic $Man_{5–9}GlcNAc_2$ and $Glc_1Man_9GlcNAc_2$; see Fig. 1); these were found by all groups no matter which method was used. Only minor amounts of truncated hybrid and complex glycans with one or two HexNAc res-

**Table 1.** Summary of data published on the structural variation of N-glycans of the model organism *C. elegans*

| *m/z* | Additional to the core $GlcNAc_2$ $[M+Na]^+$ | Refs. 6,12,14 | Ref. 10 | Ref. 13 | Refs. 8,16,17 | Ref. 11 | Refs. 3,9 | Refs. 7,15 |
|---|---|---|---|---|---|---|---|---|
| 772 | Hex2 | W↓K nd X | nd | nd | nd | nd | + | nd |
| 813 | *HexNAc1 Hex1* | nd | nd | nd | 16 | nd | nd | nd |
| 918 | Hex2 Fuc1 | W K nd X | A H↓ | nd | nd | nd | + | nd |
| 932 | Hex2 Fuc1Me | W K nd nd | nd | nd | nd | nd | nd | nd |
| 934 | **Hex3** | W K M↓X | A H | W nd | 8,16,17 | ConA | + | MM |
| 948 | Hex3 Me1 | nd | nd | nd | 8 | nd | nd | nd |
| 975 | *HexNAc1 Hex2* | nd | nd | nd | 16 | nd | nd | nd |
| 1064 | Hex2 Fuc2 | W K nd X | nd H | W nd | nd | nd | + | $M0F^3F^6$ |
| 1078 | Hex2 Fuc2Me | W K nd nd | nd H | nd | nd | L | nd | nd |
| 1080 | Hex3 Fuc1 | W K nd X | A H | W nd | 8,16,17 | ConA, L | + | $MMF^6$ |
| 1094 | Hex3 Fuc1Me | W K nd nd | nd | nd | 8 | nd | nd | nd |
| 1096 | **Hex4** | W K M X↓ | A H | W G | 8,16,17 | ConA, L | + | nd |
| 1110 | Hex4 Me1 | nd | nd | nd | 8 | nd | nd | nd |
| 1137 | *HexNAc1 Hex3* | W K nd X | A H | W nd | 8,16,17 | ConA, L | + | MGn,GnM |
| 1178 | *HexNAc2 Hex2* | nd | nd | nd | 8 | nd | nd | nd |
| 1226 | Hex3 Fuc2 | W K nd X | A↓H | W nd | nd | ConA, L | + | $MMF^3F^6$ |
| 1229 | *Hex2 Fuc2 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 1240 | Hex3 Fuc2Me | W K M X | A↓H | W nd | nd | L | nd | nd |
| 1242 | Hex4 Fuc1 | W K M↓X↓ | A H | W G↓ | 8,16,17 | ConA, L | + | nd |
| 1244 | *Hex3 Fuc1 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 1256 | Hex4 Fuc1Me | W K M nd | A↓H | W nd | 8 | L | nd | nd |
| 1258 | **Hex5** | W K M X | A H | W G | 8,16,17 | ConA, L | + | Man5 |
| 1270 | Hex4 Fuc1Me2 | nd K nd nd | nd | nd | 8 | nd | nd | nd |
| 1283 | *HexNAc1 Hex3Fuc1* | W K nd X | A↓H | nd | 8,16 | ConA, L | + | nd |
| 1299 | *HexNAc1 Hex4* | nd | nd | nd | 16 | ConA | nd | nd |
| 1302 | *HexNAc1 Hex3 PC1* | W K nd X | + | nd | 8 | nd | + | nd |
| 1340 | *HexNAc2 Hex3* | W K nd X | A↓H↓ | W↓nd | 8,16,17 | L | + | GnGn |
| 1372 | Hex3 Fuc3 | W nd nd nd | A↓H↓ | W nd | nd | L | + | nd |
| 1386 | Hex3 Fuc3Me | W K nd nd | A↓H | W nd | nd | nd | nd | nd |
| 1388 | Hex4 Fuc2 | W K nd nd | A↓H | W nd | 17 | ConA, L | + | nd |
| 1401 | Hex4 Fuc2Me | W K M↓X↓ | A↓H | W nd | 8 | L | + | nd |
| 1403 | Hex5 Fuc1 | W K M X↓ | A↓H | W G↓ | 8,16,17 | ConA, L | + | nd |
| 1415 | Hex4 Fuc2Me2 | nd K nd nd | nd H↓ | nd | nd | nd | nd | nd |
| 1418 | Hex5 Fuc1Me | W↓K M X↓ | A↓H↓ | W↓nd | 8 | nd | nd | nd |
| 1420 | **Hex6** | W K M X | A H | W G | 8,16,17 | ConA | + | Man6 |
| 1429 | *HexNAc1 Hex3Fuc2* | nd nd nd X | nd | nd | nd | nd | nd | nd |
| 1432 | Hex5 Fuc1Me2 | nd K M↓nd | nd | nd | nd | nd | nd | nd |
| 1434 | Hex6 Me | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1445 | *HexNAc1 Hex4Fuc1* | W↓nd nd X | nd | nd | 16 | nd | nd | nd |
| 1448 | *HexNAc1 Hex3Fuc1 PC1* | W↓K nd nd | + | nd | nd | nd | + | nd |
| 1459 | *HexNAc1 Hex4Fuc1Me1* | nd nd M X | nd | nd | nd | nd | nd | nd |
| 1461 | *HexNAc1 Hex5* | nd nd M nd | nd | nd | 8,16 | ConA | nd | nd |
| 1464 | *HexNAc1 Hex4 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 1486 | *HexNAc2 Hex3Fuc1* | W K nd X | A↓H↓ | nd | nd | nd | + | nd |
| 1502 | *HexNAc2 Hex4* | nd | nd | nd | 8,16,17 A | nd | nd | nd |
| 1505 | *HexNAc2 Hex3 PC1* | W↓K nd X | + | nd | 8 | nd | + | nd |
| 1534 | Hex4 Fuc3 | W nd nd nd | A↓H | W nd | nd | ConA, L | + | nd |
| 1543 | *HexNAc3 Hex3* | nd | nd | nd | 8,17 | nd | + | nd |
| 1547 | Hex4 Fuc3Me | W K nd nd | A H | W nd | nd | L | + | nd |
| 1550 | Hex5 Fuc2 | W K nd nd | A H | W G↓ | 16,17 | ConA, L | + | nd |
| 1555 | *HexNAc3 Hex2Fuc1 Me2* | nd | nd | nd | 8 | nd | nd | nd |
| 1561 | Hex4 Fuc3Me2 | nd K nd nd | A↓H↓ | nd | nd | nd | + | nd |
| 1564 | Hex5 Fuc2Me | W↓K M X↓ | A↓H↓ | W nd | 8 | L | + | nd |
| 1566 | Hex6 Fuc1 | nd nd M nd | nd | nd G | 8 | L | + | nd |
| 1572 | *Hex4Fuc1 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 1578 | Hex5 Fuc2Me2 | nd K M↓nd | nd | nd | nd | nd | nd | nd |
| 1580 | Hex6 Fuc1Me | nd nd M nd | nd | nd | nd | L | nd | nd |
| 1582 | **Hex7** | W K M X | A H | W G | 8,16,17 | ConA | + | Man7 |
| 1591 | *HexNAc1 Hex4Fuc2* | nd nd nd X | nd | nd | nd | nd | nd | nd |
| 1594 | Hex6 Fuc1Me2 | nd K nd nd | nd | nd | nd | nd | nd | nd |
| 1605 | *HexNAc1 Hex4Fuc2 Me* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1607 | *HexNAc1 Hex5Fuc1* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1620 | *HexNAc1 Hex5Fuc1 Me* | nd nd M nd | nd | nd | nd | nd | nd | nd |

(*continued on next page*)

**Table 1** (*continued*)

| *m/z* | Additional to the core $GlcNAc_2$ $[M+Na]^+$ | Refs. 6,12,14 | Ref. 10 | Ref. 13 | Refs. 8,16,17 | Ref. 11 | Refs. 3,9 | Refs. 7,15 |
|---|---|---|---|---|---|---|---|---|
| 1623 | *HexNAc1 Hex6* | nd nd M nd | nd | nd | 8 | nd | nd | nd |
| 1626 | *HexNAc1 Hex5PC1* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1629 | *HexNAc1 Hex4 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 1635 | *HexNAc1 Hex5Fuc1 Me2* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1637 | *HexNAc1 Hex6 Me* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1647 | *HexNAc2 Hex4Fuc1* | nd | nd | nd | 8 | nd | nd | nd |
| 1650 | *HexNAc2 Hex3Fuc1 PC1* | nd nd nd X↓ | + | nd | nd | nd | + | nd |
| 1670 | *HexNAc2 Hex3 PC2* | nd nd nd X↓ | nd | nd | 8 | nd | nd | nd |
| 1680 | Hex4 Fuc4 | W↓nd nd nd | A↓H | W nd | nd | nd | + | nd |
| 1689 | *HexNAc3 Hex3Fuc1* | nd | nd | nd | nd | nd | + | nd |
| 1694 | Hex4 Fuc4Me | W nd nd nd | A↓H | W nd | nd | L | nd | nd |
| 1696 | Hex5 Fuc3 | W nd nd nd | A H | W nd | nd | ConA, L | + | nd |
| 1704 | *HexNAc3 Hex4* | nd | nd | nd | 17 A | nd | nd | nd |
| 1708 | *HexNAc3 Hex3 PC1* | nd | + | nd | 8 | nd | + | nd |
| 1710 | Hex5 Fuc3Me | W K nd X | A↓H | W nd | nd | L | nd | nd |
| 1712 | Hex6 Fuc2 | W↓K nd nd | A↓H | W G↓ | nd | L | + | nd |
| 1724 | Hex5 Fuc3Me2 | nd K nd nd | nd H↓ | nd | nd | nd | nd | nd |
| 1725 | Hex6 Fuc2Me | nd K M nd | nd H↓ | nd | nd | L | nd | nd |
| 1728 | Hex7 Fuc1 | nd nd M↓nd | nd | nd G | nd | nd | nd | nd |
| 1742 | Hex7 Fuc1Me | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1744 | **Hex8** | W K M X | A H | W G | 8,16,17 | ConA | + | Man8 |
| 1745 | *HexNAc4 Hex3* | nd | nd | nd | 16 | nd | + | Tetra-ant. |
| 1753 | *HexNAc1 Hex5Fuc2* | nd | nd | nd | 17 A | nd | nd | nd |
| 1753 | *Hex5 PC3* | nd | nd | nd | 16 | nd | nd | nd |
| 1769 | *HexNAc1 Hex6Fuc1* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1772 | *HexNAc1 Hex5Fuc1 PC1* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1781 | *HexNAc1 Hex5Fuc2 Me2* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1783 | *HexNAc1 Hex6Fuc1 Me* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1797 | *HexNAc1 Hex6Fuc1 Me2* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1816 | *HexNAc2 Hex3Fuc1 PC2* | nd nd nd X↓ | nd | nd | nd | nd | nd | nd |
| 1832 | *HexNAc2 Hex4 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 1842 | Hex5 Fuc4 | W↓nd nd nd | A↓H | W nd | 17 A | nd | + | nd |
| 1851 | *HexNAc3 Hex4Fuc1* | nd | nd | nd | 8 | nd | nd | nd |
| 1854 | *HexNAc3 Hex3Fuc1 PC1* | nd | + | nd | 8 | nd | + | nd |
| 1855 | Hex5 Fuc4Me | W nd nd X | A H | W nd | nd | L | + | nd |
| 1857 | Hex6 Fuc3 | W↓nd nd X↓ | A↓H | W nd | nd | ConA, L | + | nd |
| 1866 | *HexNAc3 Hex5* | nd | nd | nd | 8,16,17 A | nd | nd | nd |
| 1869 | Hex5 Fuc4Me2 | nd | nd H↓ | nd | nd | L | nd | nd |
| 1872 | Hex6 Fuc3Me | W↓K nd X↓ | A↓H | W nd | nd | L | nd | nd |
| 1873 | *HexNAc3 Hex3 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 1874 | Hex7 Fuc2 | W↓nd nd X↓ | nd | W↓G↓ | nd | nd | nd | nd |
| 1888 | Hex7 Fuc2Me | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1890 | Hex8 Fuc1 | nd | nd | nd G↓ | nd | nd | nd | nd |
| 1904 | Hex8 Fuc1Me | W↓K M nd | nd | nd | nd | nd | nd | nd |
| 1906 | **Hex9** | W K M X | A H | W G | 8,16,17 | ConA | + | Man9 |
| 1924 | *HexNAc2 Hex3Fuc4* | nd | nd | nd | 8 | nd | nd | nd |
| 1929 | *HexNAc1 Hex6Fuc2 Me* | nd nd M nd | nd | nd | nd | nd | nd | nd |
| 1931 | *HexNAc1 Hex7Fuc1* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 1948 | *HexNAc5 Hex3* | nd | nd | nd | nd | nd | + | nd |
| 1975 | *HexNAc2 Hex5Fuc1 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 1997 | *HexNAc3 Hex4Fuc2* | nd | nd | nd | 8 | nd | nd | nd |
| 2000 | *HexNAc3 Hex3Fuc2 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2004 | Hex6 Fuc4 | nd | nd H↓ | W nd | nd | nd | nd | nd |
| 2016 | *HexNAc3 Hex4Fuc1 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2018 | Hex6 Fuc4Me | W↓nd nd X↓ | A↓H | W nd | nd | L | nd | nd |
| 2020 | Hex7 Fuc3 | nd | nd H | W nd | nd | L | nd | nd |
| 2032 | Hex6 Fuc4Me2 | nd | nd H↓ | nd | nd | nd | nd | nd |
| 2034 | Hex7 Fuc3Me | nd | nd H↓ | W↓nd | nd | nd | nd | nd |
| 2036 | Hex8 Fuc2 | nd | nd | nd G↓ | nd | nd | nd | nd |
| 2051 | Hex9 Fuc1 | nd | nd | nd G↓ | nd | nd | nd | nd |
| 2069 | **Hex10** | W↓K M↓X↓ | A↓H↓ | W↓G↓ | 8,16 | nd | + | GlcMan9 |
| 2076 | *HexNAc4 Hex3 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 2089 | *HexNAc1 Hex6Fuc3 Me2* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |
| 2091 | *HexNAc1 Hex7Fuc2 Me* | nd nd M↓nd | nd | nd | nd | nd | nd | nd |

**Table 1** (*continued*)

| *m/z* | Additional to the core $GlcNAc_2$ $[M+Na]^+$ | Refs. 6,12,14 | Ref. 10 | Ref. 13 | Refs. 8,16,17 | Ref. 11 | Refs. 3,9 | Refs. 7,15 |
|---|---|---|---|---|---|---|---|---|
| 2159 | *HexNAc3 Hex5Fuc2* | nd | nd | nd | 8 | nd | nd | nd |
| 2162 | *HexNAc3 Hex4Fuc2 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2166 | Hex7 Fuc4 | nd | nd H↓ | W↓nd | nd | nd | nd | nd |
| 2180 | Hex7 Fuc4Me | nd | nd H↓ | W↓nd | nd | nd | nd | nd |
| 2182 | Hex8 Fuc3 | nd | nd H↓ | W↓nd | nd | nd | nd | nd |
| 2200 | *HexNAc4 Hex4Fuc2* | nd | nd | nd | 8 | nd | nd | nd |
| 2219 | *HexNAc4 Hex4Fuc1 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2346 | *HexNAc4 Hex4Fuc3* | nd | nd | nd | 8 | nd | nd | nd |
| 2365 | *HexNAc4 Hex4Fuc2 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2384 | *HexNAc4 Hex4Fuc1 PC2* | nd | nd | nd | 8 | nd | nd | nd |
| 2403 | *HexNAc5 Hex4Fuc2* | nd | nd | nd | 8 | nd | nd | nd |
| 2479 | *HexNAc6 Hex4 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2527 | *HexNAc4 Hex5Fuc2 PC1* | nd | nd | nd | 8 | nd | nd | nd |
| 2530 | *HexNAc4 Hex4Fuc2 PC2* | nd | nd | nd | 8 | nd | nd | nd |

To summarise the glycosylation potential, data published in the literature so far on *C. elegans* from various laboratories, generally based on mass spectrometry of the wild-type N2 strain, are combined into this table. In general mixed-stages of wild-type or deletion/knock out mutants of *C. elegans* were used except in one study analysing the glycan profile of developmental stages.[8]

*Notes*: **Hex** for hexose (mannose, glucose and galactose), **HexNAc** either *N*-acetylglucosamine or *N*-acetylgalactosamine, **Fuc** for fucose, **PC** for phosphorylcholine, **Me** for methyl group; To simplify the table, the inner two GlcNAc residues are not included in the structural formulae. High mannosidic N-glycans are highlighted in bold, complex and hybrid types with PC and/or additional HexNAc residues are in italics. The glycome of the wild-type N2 worm (**W**) is compared, for the Refs. 6, 12 and 14, to the glycomes of various mutants (**K** for Ce *fut-1* deletion mutant (VC 378), **M** for the mannosidase II mutant (*aman-2*, tm1078), **X** for the hexosaminidase mutant (*hex-2*, tm2350)) and, for Ref. 13, to the triple GnTI knock-out mutant (**G**). A structure annotated with [W nd] is, therefore, found in the wild-type but not detected (**nd**) in the mutant. Ref. 10 compared the occurrence of certain structures when either hydrazine (**H**) or PNGase A (**A**) was used to release the glycans. In this case, [nd H↓] means that the structure was observed after release with hydrazine but not with PNGase A; a relative abundance of $\leqslant$5% is marked with an ↓. Refs. 8,16 and 17 generally used PNGase F to release the glycans, only a few more structures were found when using PNGase A (**A**) instead. Ref. 11 analysed the N-glycans of soluble glycoproteins separated through their binding to the lectin ConA (**ConA**) or to galectin (**L**), In this case, [L] means that the structure was found on galectin-bound glycoproteins. For Refs. 3,9 and 10 a + marks a detected structure. Refer to Figure 4 for the glycan abbreviations used for Refs. 7 and 15.

idues on the core pentasaccharide were found by all, while multiple non-reducing terminal 6-phosphorylcholine-substituted GlcNAc residues ('PC-rich' glycans) were only observed in some studies (see also Fig. 1). Highly-fucosylated truncated structures and methyl groups closely associated to the fucosylated structures were sometimes overlooked although they represent an important subset of the features present on *C. elegans* glycans.

The observed differences in the results are partly dependent on the method used to release the glycans or are due to the overall sensitivity of the analysis method itself. In the cases where the analysis is more or less exclusively based on mass spectrometric methods, careful interpretation is the key especially when being confronted with unusual and partially unknown structures such as the special 'core chitobiose modifications' (CCMs; see also Fig. 1).[10]

## 3. Unexpected modifications in the light of the glycogenomic potential

The core modifications found by Reinhold's group included a galactose residue linked β1,4 to the core α1,6-fucose with another hexose present on the galactose and another Galβ1,4Fuc disaccharide present on the 3-hydroxyl group of the protein linked GlcNAc.[10] A third Hex-Fuc modification, attached via the 3-hydroxyl of the second core GlcNAc, was not further analysed due to the low amount. The relevant 4-substituted fucose observed in the CCMs was neither found nor perhaps looked for, by GC–MS, in either *C. elegans*[18] or in any other closely-related nematodes like *Dictyocaulus viviparus*,[19] *Haemonchus contortus*[20] and *Parelaphostrongylus tenuis* that were examined.[21] Such a Hex-Fuc motif was previously found only in octopus rhodopsin[22] and keyhole limpet haemocyanin.[23] On the other hand, the presence of these core modifications could be an explanation for the multiply fucosylated paucimannosidic structures found by nearly all groups studying *C. elegans* glycans. In a recent analysis, galactose-substituted core α1,6-fucose was found to be a rather prominent modification of the N-glycans of a hexosaminidase (*hex-2*) mutant.[12]

Other than 1,4-substituted fucose, terminal GalNAc was not found in *C. elegans* by GC–MS, although in our studies a functional recombinant *C. elegans* GalNAc transferase (encoded by the *bre-4*/Y73E7A.7 gene), a member of the glycosyltransferase family 7, was able to add two GalNAcs to an asialoagalactoglycopeptide (the so-called GnGn) forming two LacdiNAc motifs in vitro. Furthermore, the same functional enzyme also transferred GalNAc to O-glycan-like substrates in vitro[24] and is presumed to be also involved in glycolipid biosynthesis in the worm in vivo (indeed, the name *bre-4* refers to *Bacillus* toxin resistance; one of these *Bacillus* toxins uses worm glycolipids as receptors).

When BRE-4 (the name in capitals refers to the protein, rather than the gene) was expressed in CHO cells, the enzyme was also able to participate in the synthesis of poly LacdiNAc motifs.[25]

Nevertheless, the *anti*-LacdiNAc antibody 100 2H5-A raised against *Schistosoma* glycans[26] showed a relatively weak binding to a wild-type N2 *C. elegans* extract (unpublished data), which might reflect the absence of LacdiNAc units in the worm, which are usually present in *Schistosoma* glycans.[27,28] The apparent lack of GalNAc residues on *C. elegans* N-glycans might be due to a restricted supply of the nucleotide sugar and may be influenced by the metabolism of the worms when growing in, for example, liquid culture. Other mouse IgM sera tested against mono-fucosylated LacdiNAc, difucosylated LacdiNAc and Lewis x did not crossreact with the *C. elegans* extracts, suggesting the absence of such modifications in this organism (unpublished data); thus it is unclear whether the BRE-4 GalNAc transferase has a role in N-glycan biosynthesis in the worm. On the other hand, in parasitic nematodes, LacdiNAc has been found on *Trichinella* N-glycans,[29] whereas Lewis x moieties have been found in *Dictyocaulus*.[19] In its fucosylated form, LacdiNAc is present not just in schistosomes, but also on bee venom glycoproteins.[30]

In general, glycans are a result of the interplay of glycan substrates, glycosidases and glycosyltransferases and, as such, are obviously related to the 'glycogenomic' potential of a species. The genome of *C. elegans*, published in 1998, can be searched for the glycosyltransferases relevant to N-glycan biosynthesis. The presence or absence of such genes can therefore be taken into consideration when proposing glycan structures.

Several glycosyltransferases like GnTIII, GnTIV and GnTVI (all three forming β1,4-GlcNAc linkages) and α2,3/6/8-sialyltransferases, leading to the formation of the 'normal' complex type N-glycans with tetra- or pentaantennary, sialylated structures in vertebrates, have no obvious homologues in the worm. The β1,2-xylosyltransferase, responsible for the xylosylation of the N-glycan core in plants and trematodes, is absent too. Structures originating from these enzymatic activities are, therefore, very likely not present in the worm.

On the other hand, the three homologues of GnTI, one of GnTII and one of GnTV only give rise to a restricted number of complex glycans (see Fig. 2 for an explanation of their activities); the low amount of bi- and triantennary glycans can be explained by a high endogenous activity of a Golgi hexosaminidase removing exclusively the GlcNAc residue from the α1,3-antennae, similar to that in insect cells.[31]

The high fucose content of the N-glycans is reflected by the finding of one core α1,6-fucosyltransferase homologue (FUT-8), one characterised α1,2- (out of twenty-five putative) and five α1,3/4-fucosyltransferases in the genome. One of the α1,3-fucosyltransferases (FUT-1; also known as CEFT1) was proven to modify the N-glycan core, but is inhibited in the presence of the GlcNAc residue transferred by GnTI.[6] Of the remaining four α1,3-fucosyltransferases, two were found to produce Lewis x structures using Galβ1,4GlcNAc-pNP as a substrate, whereas a third one was using a LacdiNAc structure as its acceptor; however, FUT-4 (otherwise known as CEFT5) was not active with any tested substrate.[32] The presence of enzymes with Lewis-type activities contrasts with the aforementioned lack of any obvious expression of Lewis-type epitopes; thus, the actual role of such enzymes in vivo still needs to be elucidated. The activity of one of the 25 homologues of α1,2-fucosyltransferases present could be, in theory, responsible for the in vitro or in vivo synthesis of the 2-linked fucosylated glycans similar to the O-glycan or glycolipid modifications in *Schistosoma* eggs or cercarial glycocalyx, which possess Fucα1,2+/-Fucα1,2Fucα1,3 on a backbone of HexNAc-containing units (see Fig. 3).[33–35] Certainly, up to three fucose residues, including α1,2-linked 2-*O*-methylfucose,[3] have been observed on O-glycans of wild-type N2 worms.

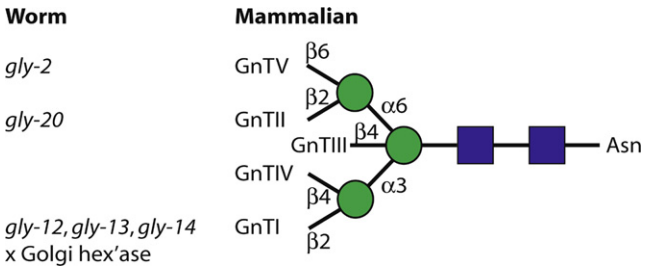


**Figure 2.** *N*-Acetylglucosaminyltransferase homologues in *C. elegans*. The GnTI activities transferring to the α1,3-antenna, encoded by the *gly-12*, *gly-13* and *gly-14* genes, are counteracted by the Golgi hexosaminidase(s). The activities of the GnTII and GnTV homologues (encoded, respectively, by the *gly-20* and *gly-2* genes) are presumed to be dependent on the prior action of GnTI. For GnTIII and GnTIV no obvious homologues have been found.

Finally, the galactosyltransferase families capable of transferring Gal or GalNAc either β1,3 to GlcNAc, β1,3 to GalNAc or β1,4 to GlcNAc have about 20 members in the worm[36] and are, therefore, candidates for being responsible for the unusual hexose modifications of fucose-containing structures.

## 4. Use of mutants in worm glycomics

Since the final definition of the N-glycan structures and the relevant biosynthetic pathways in the worm is still

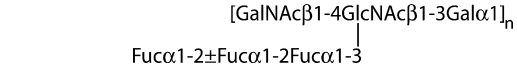


**Figure 3.** An example product of fucosyltransferases in schistosomes modifying LacdiNAc units of O-glycans; whether similar reactions occur in *C. elegans* has yet to be clarified.

awaited, the use of mutants is invaluable. Mutants with deletions in genes with proven or putative function are available from the *Caenorhabditis* Genetics Centre or other laboratories and analysing their glycan profiles aids definition of the biochemical functions of these genes with respect to the substrate specificities and their involvement in the glycosylation pathways leading to the enormous range of observed structures.

One example is the triple knock out of all genes encoding forms of GnTI; these enzymes are responsible for the β1,2-GlcNAc linkage to the α1,3-antenna and ablation of these genes results in a distinct change in the glycan pattern, which is, however, not accompanied by an apparent change of the phenotype under laboratory conditions,[13] other than altered susceptibility to bacterial infection.[37] As expected, the complex and hybrid structures depending on the prior action of GnTI as well as the paucimannose glycans are absent in the triple knock out mutant; our own unpublished data also indicates a loss of PC-modified N-glycans. The methylated structures are missing, which is mainly due to the decrease of fucosylated structures; indeed, the N-glycans of this mutant carry maximally only two fucose residues. Glycans resulting from the action of GnTI-independent fucosyltransferase(s) acting on $Hex_{7-9}GlcNAc_2$ can be observed, whereas the small amount of Man4 is possibly the product of mannosidase III.[13]

The presence of anti-horseradish peroxidase epitopes in the triple GnTI knockout is one piece of evidence demonstrating that the worm's core α1,3-fucosyltransferase (FUT-1) is rather unique. Whereas the overall 'result' of possessing core α1,3-fucose (an epitope of anti-horseradish peroxidase) is the same in plants, insects, trematodes and nematodes, it is obvious that the specificity of FUT-1 is a result of convergent evolution as compared to the other species, since it is the only one described to date, which is GnTI-independent; furthermore, as mentioned above, the enzyme is prevented from acting when a GlcNAc residue is present on the α1,3-antenna.[6] This residue is, however, normally a prerequisite for the action of core fucosyltransferases. The *C. elegans* core α1,6-fucosyltransferase, encoded by the *fut-8* gene, though, still requires the prior action of GnTI and is thus not differing from the other analysed invertebrate or vertebrate core fucosyltransferases.[38] Thus, to form difucosylated N-glycans in *C. elegans*, transfer by GnTI is followed by the consecutive actions of FUT-8, Golgi hexosaminidase and, finally, FUT-1.

Another key mutant, analysed in our laboratory, was the Golgi mannosidase II mutant, which was shown to

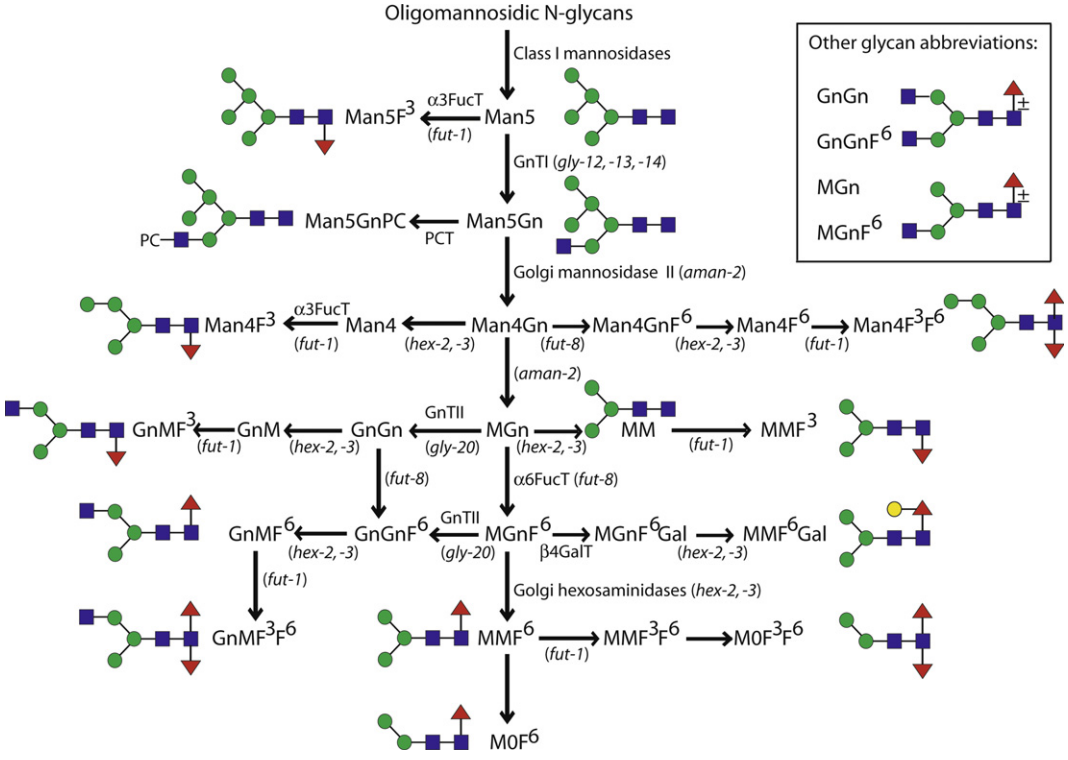


**Figure 4.** Pathways in *C. elegans* leading to fucosylated paucimannosidic structures. This summary is not intended to be exhaustive, but shows a number of reactions proven in vitro or postulated to exist in vivo (particularly, phosphorylcholinyltransferases can probably modify any non-reducing-terminal GlcNAc); pathways to core chitobiose modifications have not been elucidated to date. Names of genes relevant for certain biosynthetic steps are indicated in brackets. Glycan nomenclature is based upon that introduced by Schachter with selected structures also shown in graphical forms.

lack, as expected, the paucimannosidic truncated structures, but had a whole subset of structures not found in the wild-type.[14] When analysing the N-glycans, evidence for products of the transfer by core α1,6 fucosyltransferase to Man5Gn was found (reflecting that this enzyme utilises MGn and Man5Gn and not only GnGn), whereas it was also obvious that FUT-1 modifies Man5 and that the phosphorylcholinyltransferase(s) modified Man5Gn. In the *hex-2* mutant, lacking a potential Golgi hexosaminidase, structures containing non-reducing terminal GlcNAc were enriched as well as structures with galactose-modified core α1,6-fucose.[12] Overall, based on the knowledge we gained from analysing mutants and the substrate specificity of the relevant glycosyltransferases as well as previous summaries of the glycosylation capabilities of the worm,[14,37,38] we can postulate the pathways shown in Figure 4.

Although, as stated above the pathway to the paucimannosidic fucosylated structures differs from the other invertebrates, the outcome of paucimannosidic fucosylated structures is similar. So why did *C. elegans* and probably the other nematodes undergo the effort to develop that difference? The clue may be to consider the highly-fucosylated structures such as the core chitobiose modifications which are probably processed from the paucimannosidic fucosylated structures; during their biosynthesis, the high hexosaminidase activity, which results in an absence of 'complex' substrates, is perhaps being functionally bypassed via the 'GlcNAc-independent' core α1,3-fucosyltransferase FUT-1. These multifucosylated structures presumably represent an important set of modifications and, in other nematodes, are probably serving a special aim such as supporting the parasitic life-style as in *Haemonchus contortus*. However, under laboratory conditions *C. elegans* mutants lacking highly-fucosylated structures are still seemingly quite happy, although those lacking more-or-less all fucose (the *bre-1*/*gmd-1* mutant) do have lower brood sizes.[39] On the other hand, agar plates with *Escherichia coli* are not reflecting the normal nematodal life conditions. Indeed, as nematodes have specialised in synthesising such a range of unusual fucosylated structures, in comparison to other higher organisms in which both fucose and sialic acid are involved in recognition, one may expect some biological function also in the non-parasitic species, which requires alternative 'stresses' (e.g., infection) to be shown to be significant.

## 5. Conclusion

Finally, to recapitulate the possible and partly proven N-glycan modifications in *C. elegans* (see also Fig. 5): the core of the N-glycans can indeed carry α1,3 and α1,6 core fucose which in turn can be modified by galactose residues. The complex glycans carrying up to three antennae can be capped with PC, whereas either an H-type modification or linkage of both galactose and fucose directly to the trimannosyl core (potentially methylated) have been also proposed; on the other hand, probably neither Lewis x, fucosylated LacdiNAc, bisecting GlcNAc nor tetraantennary glycans are expressed in *C. elegans*. Many of the N-glycan modifications are unique to the worm, but the study of those shared with parasitic worms, such as PC, which was found in *Ascaris suum*,[40] *Trichinella spiralis*[29] and filarial nematodes[41] or trifucosylated core shared with *Haemonchus contortus*,[42] may yield interesting therapeutic avenues.

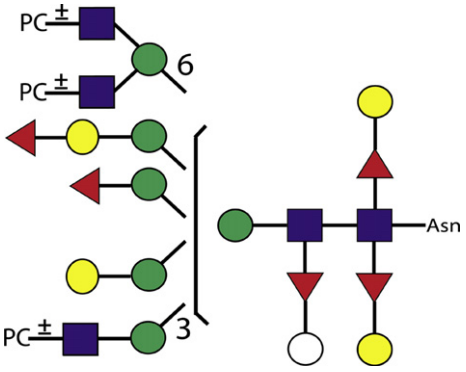


**Figure 5.** Putative modifications of the core and the antennae present in *C. elegans*. Based on studies by various groups, predominantly using mass spectrometry, many of these features are, to date, unique to the worm. The typical oligomannosidic structures are not shown. The symbols used are blue squares for GlcNAc, green circles for Man, yellow circles for Gal, open circles for undefined hexose residues and red triangles for Fuc residues (nomenclature of the Consortium for Functional Glycomics).

## Acknowledgements

Our work on *C. elegans* has been funded by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (Grants P15475 and P18447 to I.B.H.W.) We also thank other members of the group, especially Denise Kerner, and our collaborators Günter Lochnit (Justus-Liebig-Universität Gießen) and Verena Jantsch (Universität Wien) for their participation in these projects.

## References

1. Haslam, S. M.; Dell, A. *Biochimie* **2003**, *85*, 25–32.
2. Gerdt, S.; Dennis, R. D.; Borgonie, G.; Schnabel, R.; Geyer, R. *Eur. J. Biochem.* **1999**, *266*, 952–963.
3. Griffitts, J. S.; Haslam, S. M.; Yang, T.; Garczynski, S. F.; Mulloy, B.; Morris, H.; Cremer, P. S.; Dell, A.; Adang, M. J.; Aroian, R. V. *Science* **2005**, *307*, 922–925.
4. Guerardel, Y.; Balanzino, L.; Maes, E.; Leroy, Y.; Coddeville, B.; Oriol, R.; Strecker, G. *Biochem. J.* **2001**, *357*, 167–182.

5. Harnett, W.; Harnett, M. M. *Biochim. Biophys. Acta* **2001**, *1539*, 7–15.
6. Paschinger, K.; Rendić, D.; Lochnit, G.; Jantsch, V.; Wilson, I. B. H. *J. Biol. Chem.* **2004**, *279*, 49588–49598.
7. Natsuka, S.; Kawaguchi, M.; Wada, Y.; Ichikawa, A.; Ikura, K.; Hase, S. *J. Biochem. (Tokyo)* **2005**, *138*, 209–213.
8. Cipollo, J. F.; Awad, A. M.; Costello, C. E.; Hirschberg, C. B. *J. Biol. Chem.* **2005**, *280*, 26063–26072.
9. Haslam, S. M.; Gems, D.; Morris, H. R.; Dell, A. *Biochem. Soc. Symp.* **2002**, *69*, 117–134.
10. Hanneman, A. J.; Rosa, J. C.; Ashline, D.; Reinhold, V. N. *Glycobiology* **2006**, *16*, 874–890.
11. Hirabayashi, J.; Hayama, K.; Kaji, H.; Isobe, T.; Kasai, K. *J. Biochem. (Tokyo)* **2002**, *132*, 103–114.
12. Gutternigg, M.; Kretschmer-Lubich, D.; Paschinger, K.; Rendić, D.; Hader, J.; Geier, P.; Ranftl, R.; Jantsch, V.; Lochnit, G.; Wilson, I. B. H. *J. Biol. Chem.* **2007**, *282*, 27825–27840.
13. Zhu, S.; Hanneman, A.; Reinhold, V.; Spence, A.; Schachter, H. *Biochem. J.* **2004**, *382*, 995–1001.
14. Paschinger, K.; Hackl, M.; Gutternigg, M.; Kretschmer-Lubich, D.; Stemmer, U.; Jantsch, V.; Lochnit, G.; Wilson, I. B. H. *J. Biol. Chem.* **2006**, *281*, 28265–28277.
15. Natsuka, S.; Adachi, J.; Kawaguchi, M.; Nakakita, S.; Hase, S.; Ichikawa, A.; Ikura, K. *J. Biochem. (Tokyo)* **2002**, *131*, 807–813.
16. Cipollo, J. F.; Costello, C. E.; Hirschberg, C. B. *J. Biol. Chem.* **2002**, *277*, 49143–49157.
17. Cipollo, J. F.; Awad, A. M.; Costello, C. E.; Hirschberg, C. B. *J. Biol. Chem.* **2004**, *279*, 52893–52903.
18. *Glycogenomics: The Impact of Genomics and Informatics in Glycobiology*; Drickamer, K., Dell, A., Eds.; Portland Press: London, 2002.
19. Haslam, S. M.; Coles, G. C.; Morris, H. R.; Dell, A. *Glycobiology* **2000**, *10*, 223–229.
20. Haslam, S. M.; Coles, G. C.; Munn, E. A.; Smith, T. S.; Smith, H. F.; Morris, H. R.; Dell, A. *J. Biol. Chem.* **1996**, *271*, 30561–30570.
21. Duffy, M. S.; Morris, H. R.; Dell, A.; Appleton, J. A.; Haslam, S. M. *Glycobiology* **2006**, *16*, 854–862.
22. Zhang, Y.; Iwasa, T.; Tsuda, M.; Kobata, A.; Takasaki, S. *Glycobiology* **1997**, *7*, 1153–1158.
23. Wuhrer, M.; Robijn, M. L.; Koeleman, C. A.; Balog, C. I.; Geyer, R.; Deelder, A. M.; Hokke, C. H. *Biochem. J.* **2004**, *378*, 625–632.
24. Kawar, Z.; van Die, I.; Cummings, R. D. *J. Biol. Chem.* **2002**, *277*, 34924–34932.
25. Kawar, Z. S.; Haslam, S. M.; Morris, H. R.; Dell, A.; Cummings, R. D. *J. Biol. Chem.* **2005**, *280*, 12810–12819.
26. Robijn, M. L.; Wuhrer, M.; Kornelis, D.; Deelder, A. M.; Geyer, R.; Hokke, C. H. *Parasitology* **2005**, *130*, 67–77.
27. Wuhrer, M.; Koeleman, C. A.; Deelder, A. M.; Hokke, C. H. *FEBS J.* **2006**, *273*, 347–361.
28. Wuhrer, M.; Koeleman, C. A.; Fitzpatrick, J. M.; Hoffmann, K. F.; Deelder, A. M.; Hokke, C. H. *Glycobiology* **2006**, *16*, 991–1006.
29. Morelle, W.; Haslam, S. M.; Olivier, V.; Appleton, J. A.; Morris, H. R.; Dell, A. *Glycobiology* **2000**, *10*, 941–950.
30. Kubelka, V.; Altmann, F.; Staudacher, E.; Tretter, V.; März, L.; Hård, K.; Kamerling, J. P.; Vliegenthart, J. F. G. *Eur. J. Biochem.* **1993**, *213*, 1193–1204.
31. Altmann, F.; Schwihla, H.; Staudacher, E.; Glössl, J.; März, L. *J. Biol. Chem.* **1995**, *270*, 17344–17349.
32. Nguyen, K.; Van Die, I.; Grundahl, K. M.; Kawar, Z. S.; Cummings, R. D. *Glycobiology* **2007**, *17*, 586–599.
33. Khoo, K.-H.; Sarda, S.; Xu, X.; Caulfield, J. P.; McNeil, M. R.; Homans, S. W.; Morris, H. R.; Dell, A. *J. Biol. Chem.* **1995**, *270*, 17114–17123.
34. Khoo, K. H.; Chatterjee, D.; Caulfield, J. P.; Morris, H. R.; Dell, A. *Glycobiology* **1997**, *7*, 653–661.
35. Wuhrer, M.; Kantelhardt, S. R.; Dennis, R. D.; Doenhoff, M. J.; Lochnit, G.; Geyer, R. *Eur. J. Biochem.* **2002**, *269*, 481–493.
36. Altmann, F.; Fabini, G.; Ahorn, H.; Wilson, I. B. H. *Biochimie* **2001**, *83*, 703–712.
37. Shi, H.; Tan, J.; Schachter, H. *Methods Enzymol.* **2006**, *417*, 359–389.
38. Paschinger, K.; Staudacher, E.; Stemmer, U.; Fabini, G.; Wilson, I. B. H. *Glycobiology* **2005**, *15*, 463–474.
39. Barrows, B. D.; Haslam, S. M.; Bischof, L. J.; Morris, H. R.; Dell, A.; Aroian, R. V. *J. Biol. Chem.* **2007**, *282*, 3302–3311.
40. Pöltl, G.; Kerner, D.; Paschinger, K.; Wilson, I. B. H. *FEBS J.* **2007**, *274*, 714–726.
41. Haslam, S. M.; Houston, K. M.; Harnett, W.; Reason, A. J.; Morris, H. R.; Dell, A. *J. Biol. Chem.* **1999**, *274*, 20953–20960.
42. Haslam, S. M.; Coles, G. C.; Reason, A. J.; Morris, H. R.; Dell, A. *Mol. Biochem. Parasitol.* **1998**, *93*, 143–147.