Published in final edited form as:
Biochim Biophys Acta. 2017 April ; 1861(4): 699–714. doi:10.1016/j.bbagen.2017.01.009.

The underestimated N-glycomes of lepidopteran species

Rhiannon Stanton[a], **Alba Hykollari**[a], **Barbara Eckmair**[a], **Daniel Malzl**[a], **Martin Dragosits**[a], **Dieter Palmberger**[b], **Ping Wang**[c], **Iain B. H. Wilson**[a], and **Katharina Paschinger**[a,*]

[a]Department für Chemie, Universität für Bodenkultur, 1190 Wien, Austria

[b]Department für Biotechnologie, Universität für Bodenkultur, 1190 Wien, Austria

[c]Department of Entomology, Cornell University, Geneva, NY 14456, USA

Abstract

Background—Insects are significant to the environment, agriculture, health and biotechnology. Many of these aspects display some relationship to glycosylation, e.g., in case of pathogen binding or production of humanised antibodies; for a long time, it has been considered that insect N-glycosylation potentials are rather similar and simple, but as more species are glycomically analysed in depth, it is becoming obvious that there is indeed a large structural diversity and interspecies variability.

Methods—Using an off-line LC-MALDI-TOF MS approach, we have analysed the N-glycomes of two lepidopteran species (the cabbage looper *Trichoplusia ni* and the gypsy moth *Lymantria dispar*) as well as of the commonly-used *T. ni* High Five cell line.

Results—We detected not only sulphated, glucuronylated, core difucosylated and Lewis-like antennal fucosylated structures, but also the zwitterion phosphorylcholine on antennal GlcNAc residues, a modification otherwise familiar from nematodes; in *L. dispar*, N-glycans with glycolipid-like antennae containing α-linked *N*-acetylgalactosamine were also revealed.

Conclusion—The lepidopteran glycomes analysed not only display core α1,3-fucosylation, which is foreign to mammals, but also up to 5% anionic and/or zwitterionic glycans previously not found in these species.

Significance—The occurrence of anionic and zwitterionic glycans in the Lepidoptera data is not only of glycoanalytical and evolutionary interest, but is of biotechnological relevance as lepidopteran cell lines are potential factories for recombinant glycoprotein production.

Keywords

N-linked oligosaccharides; insect; glucuronic acid; sulphate; phosphorylcholine

1 Introduction

The first reports of N-glycan structures in insect cells date from the early 1980's and indicated a predominance of oligomannosidic and paucimannosidic forms (1,2). In the

* Corresponding author: Tel: +43-1-47654-77216; katharina.paschinger@boku.ac.at.

Europe PMC Funders Author Manuscripts

1990's, core α1,6-fucosylation was observed in *Drosophila melanogaster* larvae (3) and, around the same time, the unexpected finding of difucosylation of N-glycan cores (both α1,3- and α1,6-fucose on the reducing terminal GlcNAc residue) was reported for oligosaccharides released from honeybee venom phospholipase and hyaluronidase (4,5); later this feature was found on membrane glycoproteins of the *Mamestra brassicae* Mb-0503 cell line (6) as well as on recombinant glycoproteins (e.g., antibodies) expressed in *Trichoplusia ni* (High Five) cells (7). In 2001 it was then demonstrated that core α1,3-linked fucose is the molecular basis for the cross-reaction of some insect glycoproteins towards antibodies raised against plant glycoproteins (8). On the other hand, the matter of sialylation in insects was a topic of more controversy, but the identification of an active sialyltransferase gene and low amounts of sialylated N-glycans in *Drosophila melanogaster* demonstrated that insects indeed possess a sialylation capability (9,10). However, there were probably limited efforts directed specifically at discovery of anionic N-glycans in insect species, but a permethylation approach showed that glucuronylated structures were also present in *Drosophila* (11). Nevertheless, in general, insects and their cell lines have been considered to have a relatively limited and simple capacity to modify their N-glycans, except when re-engineered to produce mammalian-type structures (12).

We recently considered it of interest to apply a modified glycomic workflow, targeted to facilitate the analysis of anionic N-glycans, to the N-glycomes of *Anopheles gambiae* and *Aedes aegypti* as these mosquito species are vectors for both protist and viral pathogens; this resulted in the detection of a range of sulphated and/or glucuronylated N-glycan structures (13). As another recent study demonstrated sulphation of the N-glycans of a lepidopteran serine protease (14), it appeared timely to reappraise the N-glycomic potential of lepidopteran species. Thereby, we have examined the N-glycans of larvae of the cabbage looper *T. ni* and the gypsy moth *Lymantria dispar*; while these species are, respectively, agricultural and forestry pests (15,16), especially the former species is familiar to biotechnologists as the source of the aforementioned High Five cell line originally isolated as BTI-Tn-5B1-4 (17). Our results show an unprecedented variability in the antennal decorations of insect N-glycans with not only sulphate and glucuronic acid being present, but surprisingly also phosphorylcholine which is more familiar as a modification of nematode oligosaccharides.

2 Experimental Procedures

2.1 Biological material

Trichoplusia ni larvae (mid-fifth instar; Cornell strain (18)) reared on a high wheat germ diet (19) were dissected on ice to remove the entire gut, whereas *T. ni* High Five cells (BTI-TN-5B1-4) were cultivated in serum-free medium and then washed with phosphate-buffered saline; *Lymantria dispar* fifth or sixth instar larvae were also reared on the same high wheat germ diet. The entire larvae (excluding the gut in the case of *T. ni*) were lyophilised before grinding in liquid nitrogen, whereas the cells were lysed by heat treatment in water.

2.2 Enzymatic release and purification of N-glycans

Cellular or larval homogenates were proteolysed either with thermolysin in the case of the larvae or pepsin in the case of High Five cells (20–22), prior to cation exchange and gel filtration chromatography of the proteolysate. Thereafter, N-glycans were released from glycopeptides using peptide:N-glycosidase F (PNGase F; Roche) as previously described (21,22), with a subsequent digestion of the remaining glycopeptides using peptide:N-glycosidase A (PNGase A; Roche). After an initial purification by cation-exchange chromatography (Dowex AG50; flow-through), the glycans were subject to solid-phase extraction on non-porous graphitised carbon (SupelClean ENVICarb, Sigma-Aldrich) as described (22,23); the ‘neutral’ and ‘anionic-enriched’ fractions were subsequently eluted with 40% acetonitrile and 40% acetonitrile containing 0.1% trifluoroacetic acid respectively. The pools of glycans were then subjected to reductive amination using 2-aminopyridine (PA) (21). Refer to the Supplement for a Scheme depicting the workflow as well as for further explanations regarding the glycomic analyses and assignments (see also Ref. 22).

2.3 HPLC fractionation

Complete pyridylaminated N-glycomes were fractionated by reversed-phase HPLC (Ascentis Express RP-amide from Sigma-Aldrich; 150 x 4.6 mm, 2.7 μm) and a gradient of 30% (v/v) methanol (buffer B) in 100 mM ammonium acetate, pH 4 (buffer A) was applied at a flow rate of 0.8 ml/min (Shimadzu LC-30 AD pumps) as follows: 0-4 min, 0% B; 4-14 min, 0-5% B; 14-24 min, 5-15% B; 24-34 min, 15-35% B; 34-35 min, return to starting conditions (24). Glycan fractions were collected manually based on observation of the fluorescence intensity (excitation/emission at 320/400 nm; Shimadzu RF 20 AXS detector) and analysed by MALDI-TOF MS and MS/MS. The RP-HPLC column was calibrated daily in terms of glucose units using a pyridylaminated dextran hydrolysate and the degree of polymerisation of single standards was verified by MALDI-TOF MS (22).

2.4 Mass spectrometry

Monoisotopic MALDI-TOF MS was performed using an Autoflex Speed (Bruker Daltonics, Bremen) instrument in either positive or negative reflectron modes with 6-aza-2-thiothymine (ATT; Sigma-Aldrich) as matrix; samples (0.8 μl) were vacuum dried on ground or polished steel plates before addition of matrix (0.8 μl of 3 mg/ml ATT in 50% ethanol) and crystallisation again under vacuum (22). The matrix region was suppressed and so MS spectra were normally recorded in the range *m/z* 700-3500. MS/MS was in general performed by laser-induced dissociation of the $[M+H]^+$ or $[M-H]^-$ pseudomolecular ions; typically 1000 shots were summed for MS and 3000 for MS/MS. Spectra were processed with the manufacturer’s software (Bruker Flexanalysis 3.3.80) using the SNAP algorithm with a signal/noise threshold of 6 for MS (unsmoothed) and 3 for MS/MS (four-times smoothed). Glycan spectra were manually interpreted on the basis of the masses of the predicted component monosaccharides, differences of mass in glycan series, fragmentation pattern, comparison with co-eluting structures from other insects and nematodes and chemical or exoglycosidase digestions. The symbolic annotations of spectra and chromatograms are according to the standard nomenclature (25).

2.5 Enzymatic and chemical treatments

Glycans were treated, prior to re-analysis by MALDI-TOF MS, with α-fucosidases (bovine kidney from Sigma-Aldrich or almond α1,3-specific from Prozyme), α-mannosidases (jack bean from Sigma, *Aspergillus* α1,2-specific from Prozyme and *Xanthomonas* α1,2/3-specific or α1,6-specific from New England Biolabs), β-glucuronidase (Megazyme; desalted and concentrated ten-fold before use), β-*N*-acetylhexosaminidases (jack bean from Sigma-Aldrich or in-house produced recombinant forms of *Caenorhabditis elegans* HEX-4 specific for GalNAc residues or *Apis mellifera* FDL specific for the product of GlcNAc-transferase I (26)) or α-*N*-acetylgalactosaminidase (chicken liver; Sigma-Aldrich) in 50 mM ammonium acetate, pH 5, at 37 °C overnight (except for pH 6.5 in the case of HEX-4, pH 7 in the case of *E. coli* β-glucuronidase or an incubation time of only 3 hours in the case of FDL). Hydrofluoric acid was used for removal of core or antennal α1,3-fucose or of phosphorylcholine (27). As appropriate, treated glycans were re-chromatographed by RP-HPLC to ascertain retention time shifts prior to MALDI-TOF-MS.

2.6 Western blotting

For detection of glycan epitopes, extracts of lepidopteran cells were subject to Western blotting with anti-horseradish peroxidase (anti-HRP; recognising core α1,3-fucose), C-reactive protein (recognising phosphorylcholine) or the TEPC15 anti-phosphorylcholine antibody basically as previously described (28,29). Specifically, cell pellets of *T. ni* High Five (BTI-Tn-5B1-4 (17)), Tnms42 (an alpha-nodavirus-free cell line generated as a subclone of High Five cells (30)) and HfVf, another virus-free High Five derivative, as well as *Spodoptera frugiperda* Sf9 cells (IPLB-SF21-AE subclone of Sf21 cells (31)), were solubilised with Triton X-100 as detergent, methanol precipitated (*i.e.*, mixed with a five-fold excess of methanol over extract, incubated at -80 °C for one hour and centrifuged at 4 °C, 21000 g) and subject to Western blotting. After blocking, nitrocellulose membranes were incubated with either 1:10000 diluted rabbit anti-HRP (Sigma-Aldrich), 1:200 diluted human C-reactive protein (MP Biochemicals; in the presence of 2.5 mM $CaCl_2$) or murine TEPC15 monoclonal antibody mouse IgA (Sigma-Aldrich); subsequently, the blots were respectively incubated with either 1:2000 peroxidase-conjugated anti-rabbit IgG, 1:1000 anti-C-reactive protein (followed by peroxidase-conjugated anti-rabbit IgG) or 1:1000 peroxidase-conjugated anti-mouse IgA. Chromogenic detection was with a solution of SigmaFAST 3,3'-diaminobenzidine tetrahydrochloride. Approximate equivalence of total protein loading was determined by Ponceau S staining (0.5% (w/v) in 1% acetic acid) of the membranes, prior to incubation with the antibodies (see also Supplementary Figure 2).

3 Results

3.1 General strategy for lepidopteran N-glycomics

In our previous study on dipteran N-glycans, with a focus on mosquito larvae, we demonstrated the presence of sulphated and glucuronylated antennae (13), which indicates that insect glycosylation is indeed more variable than previously thought and that these species have an inherent significant capacity to generate 'complex' oligosaccharides. Key to detecting these structures was fractionation of the free glycans by solid-phase extraction on non-porous graphitised carbon followed by fluorescent labelling, RP-HPLC using an RP-

amide fused core column and MALDI-TOF MS/MS in positive and negative ion modes. In this study, we analysed larvae of two moth species (larvae of the cabbage looper *T. ni* and the gypsy moth *L. dispar*) as well as the High Five cell line derived from *T. ni*, which is often used in research as a host for the baculovirus expression system. Each N-glycome was fractionated into three pools (PNGase F-released neutral, PNGase F-released anionic and PNGase A 'post F' released neutral) prior to fluorescent labelling by pyridylamination and HPLC (see Scheme in supplement). The structural proposals for each pool are summarised in Figures 1, 2 and 3 as well as in Supplementary Figures 1 and 2(see also the Supplementary Table) and are based on manually-interpreted MS/MS data as well as on the results of chemical and enzymatic digests.

3.2 Unusual oligomannosidic and fucosylated structures

The most abundant and simplest glycans are the oligomannosidic ($Hex_{5-9}HexNAc_2$) and paucimannosidic forms ($Hex_{2-3}HexNAc_2$ with and without core α1,6-fucose), whereby the relative order of elution of typical oligomannosidic glycans (also verified by MS/MS; e.g., Man8B > Man8A > Man8C) is basically the same as for a standard C18 column (32,33). As these structures co-eluted with those from other organisms whose glycans were analysed with the RP-amide fused core HPLC column and displayed similar MS/MS fragmentation patterns (24), they are annotated on the chromatograms without further discussion. Thereby, the core α1,6-fucosylated form of $Man_3GlcNAc_2$(*m/z* 1135 as $[M+H]^+$) was the major glycan in *T. ni* larvae, whereas the neutral glycomes of High Five cells and *L. dispar* larvae were dominated, respectively, by $Man_6GlcNAc_2$ and $Man_9GlcNAc_2$ (*m/z* 1475 and 1961). However, related fucosylated oligomannosidic, fucosylated hybrid and some oligomannosidic structures required enzymatic digests to aid their identification.

In the case of the detected $Hex_7HexNAc_2Fuc_1$ (*m/z* 1783; 7.7 g.u.), the sensitivity to bovine α-fucosidase and fungal α1,2-mannosidase treatments (Figure 4 A-C), which were accompanied by relevant changes in the MS/MS spectrum (Figure 4 I-K), indicated that the structure is an α1,6-fucosylated form of $Man_7GlcNAc_2$ with two α1,2-mannose residues on the lower arm; such a structure 'breaks' the normally-assumed rules for core fucosylation in insects as normally a GlcNAc residue on the lower arm (in the context of $Man_{3-5}GlcNAc_3$) is expected to be a pre-requisite for a potential core α1,6-fucosyltransferase substrate. Furthermore, there were three isoforms of $Hex_5HexNAc_2Fuc_1$ (*m/z* 1459) varying in terms of elution time (8.4, 9.5 and 11.0 g.u.) but also in their fragmentation patterns and α1,2-mannosidase sensitivity (Figure 4 L-P), which are concluded to be α1,6-fucosylated forms of the three detected *m/z* 1313 isomers. Indeed, in addition to the 'normal' biosynthetic $Man_5GlcNAc_2$ (*m/z* 1313) eluting at 7.2. g.u., two less familiar $Man_5GlcNAc_2$ isomers (6.4 and 8.4 g.u.) were detected with either long lower or middle arms (major fragment ions at *m/z* 989 or 827) and α1,2-mannosidase digestion properties analogous to those of the unusual $Man_5GlcNAc_2Fuc_1$ isomers. An example hybrid fucosylated glycan type is represented by a $Hex_4HexNAc_3Fuc_1$ glycan (*m/z* 1500; 8.4 g.u.), which was sensitive to FDL β-hexosaminidase as well as jack bean and *Xanthomonas* α1,6-specific mannosidases (Figure 4 D-F, Q and R).

The more unusual oligomannosidic glycans are exemplified by a fraction containing $Hex_{6-7}HexNAc_2$ structures as well as $Hex_{10}HexNAc_2$. Due to the loss of two mannose residues each upon α1,2-mannosidase digestion, it is concluded that the *m/z* 1475 glycan with the intense *m/z* 665 fragment is a $Man_6GlcNAc_2$ structure lacking a 'lower' α1,3-mannose residue, whereas the *m/z* 1637 and 2123 structures are respectively a $Man_7GlcNAc_2$ isomer (major *m/z* 827 fragment) and a terminally monoglucosylated glycan (major *m/z* 1313 fragment; Figure 4 G, H and S-X and Supplementary Figure 3 C and D). In general, the structures of the neutral oligomannosidic-based glycans suggest a broad variation in ER/Golgi mannosidase-mediated processing with some degree of incomplete removal of the 'middle' B-mannose.

3.3 Core difucosylated N-glycans

A glycan epitope typically associated with insects is the core difucosylated structure whereby the proximal reducing-terminal GlcNAc is both α1,3- and α1,6-fucosylated (34), which gives rise to characteristic *m/z* 592 $HexNAc_1Fuc_2$ fragment ions in the case of pyridylaminated derivatives (Supplementary Figure 2 I-M). Due to the presence of the α1,3-fucose, such structures can only be released with PNGase A. The relatively low amounts of PNGase A-released glycans contained also some oligomannosidic and monofucosylated species, which is due to a slightly incomplete release with PNGase F (Supplementary Figure 1 and 2A). Glycans with *m/z* 1281 ($Hex_3HexNAc_2Fuc_2$) were present in all samples and co-eluted with proven $Man_3GlcNAc_2Fuc_2$ structures from other insects and nematodes (13,24); particularly, the HighFive cell line was found to contain a number of difucosylated glycans of the form $Hex_{2-3}HexNAc_{2-4}Fuc_2$ with two different elution times for $Hex_2HexNAc_2Fuc_2$ (*m/z* 1119; 6.2 and 6.7 g.u.) and $Hex_3HexNAc_3Fuc_2$ (*m/z* 1484; 6.2 and 7.8 g.u.) indicative of antennal isomers as deduced from exoglycosidase sensitivities and comparison to previously-published data (Supplementary Figure 2 B-G). A couple of early-eluting monofucosylated glycans (4-5 g.u.; $Hex_{2-3}HexNAc_{2-3}Fuc_1$), in addition to the late-eluting forms (9-10 g.u.), were detected in the High Five and *L. dispar* PNGase A-released samples. Consistent with previous observations that core α1,3-fucose results in lower retention on RP-HPLC than core α1,6-fucose (24), such early-eluting forms are concluded to be core α1,3-fucosylated (Supplementary Figure 2H). The presence of core α1,3-fucose was also indicated by cross-reactivity of *T. ni* cell extracts towards anti-HRP (Supplementary Figure 2A); the absence of staining with Sf9 extracts is in accordance with previous observations (35).

3.4 Antennal LacdiNAc and type I LacNAc

Fucosylated LacdiNAc (GalNAcβ1,4[Fucα1,3]GlcNAc) has been previously observed as a feature of bee venom glycans as well as of mosquitoes (5,13); therefore, the presence of this epitope was not unexpected in lepidopteran species. To prove this, almond α1,3-fucosidase treatment was performed on $Hex_3HexNAc_4Fuc_2$ (*m/z* 1687; 8.0 g.u.) isolated from *L. dispar* and co-eluting with glycans of the same composition derived from nematodes or mosquitoes (13,27); this resulted in a loss of the *m/z* 553 B fragment ions ($HexNAc_2Fuc_1$) and appearance of ones at *m/z* 407 (two HexNAc residues in series; Figure 5 A and B as well as Supplementary Figure 3 G and H). The later-eluting but less abundant isomer (10.5 g.u.) was concluded, due to its later elution time and lack of the *m/z* 1525 Y fragment (associated with

loss of α1,6-mannose), to have the fucosylated LacdiNAc motif on the upper arm (Figure 5C and Supplementary Figure 3 J).

In addition and as discussed below, but to our surprise for an insect glycome, there were also lepidopteran glycans whose fragmentation patterns were indicative of the presence of phosphorylcholine (*Δm/z* 165). For instance, MS/MS of the *m/z* 1852 glycan resulted in detection of *m/z* 718 and 880 ($Hex_{0-1}HexNAc_2Fuc_1PC_1$) B fragment ions reminiscent of a glycan with the same elution and compositional properties from *Trichuris suis* (27). Treatment with *C. elegans* GalNAc-specific HEX-4 and almond α1,3-fucosidase altered the fragmentation pattern giving rise to ones at *m/z* 515 or 572 which correspond to $HexNAc_1Fuc_1PC_1$ and $HexNAc_2PC_1$; the presence of Y-fragments resulting from loss of the α1,6-mannose (*m/z* 1690, 1487 and 1544) are indicative of lower arm antennae (Figure 5 D-F). These data indicate not only that the zwitterion substitutes a β1,2-linked GlcNAc with the basic structure being the same as the aforementioned 8 g.u. isomer of *m/z* 1687, but actually are a proof of a C6 substitution: as (i) the fucose is in an α1,3-linkage as judged by its high sensitivity to hydrofluoric acid or almond α1,3-fucosidase, (ii) the second HexNAc can be removed with HEX-4 (thus showing a β1,4-linkage) and (iii) the C2 is occupied by the *N*-acetyl moiety, only the C6 of the GlcNAc can form a bond with the phosphorylcholine, which is analogous to the specifically proven linkage on zwitterionic LacdiNAc-containing nematode glycolipids (36).

Other antennal variations were exemplified by unmodified $Hex_1HexNAc_1$ and $HexNAc_2$ motifs; one *m/z* 1500 glycan ($Hex_4HexNAc_3Fuc_1$) was sensitive to both α1,2/3-specific mannosidase and β1,3-specific galactosidase treatments due to the presence of a type I LacNAc (Galβ1,3GlcNAc) modification of the upper arm (Supplementary Figure 4 A-C). For the composition $Hex_3HexNAc_5Fuc_1$ (*m/z* 1744), two elution positions (11 and 13.5 g.u.) were observed in the *T. ni* glycome, as previous seen in mosquitoes (13); the contrasting MS/MS spectra indicated the presence of a LacdiNAc motif on the former (the B fragment of *m/z* 407 being diagnostic), whereas a triantennary glycan is proposed for the latter (Supplementary Figure 4 D and E), which co-elutes with an asialoagalacto-form of the major fetuin N-glycan known to have β1,2- and β1,4-GlcNAc linked to the lower arm (37).

3.5 Phosphorylcholine-modified N-glycans

In addition to the aforementioned phosphorylcholine-modified *m/z* 1852 structure, distinct B fragment ions at *m/z* 369 ($HexNAc_1PC_1$) were detected for a range of glycans with elution times and MS/MS spectra similar to nematode glycans (see annotations on Figures 1-3). Three isomeric variants were observed for, e.g., *m/z* 1706 ($Hex_3HexNAc_4Fuc_1PC_1$), the most abundant of which (9.0 g.u.) was sensitive to bovine α-fucosidase, *C. elegans* HEX-4 and hydrofluoric acid treatments; MS/MS of the respective products of *m/z* 1560, 1503 and 1541 verified loss of either core fucose, a terminal GalNAc or a phosphorylcholine, thereby revealing the zwitterionic modification of a lower arm LacdiNAc antenna (Figure 5 G-J). Whereas the *m/z* 1706 structure eluting at 14 g.u. (Supplementary Figure 4H) was concluded on the basis of its late retention to possess the phosphorylcholine-modified LacdiNAc on the upper arm, the third variant (9.5 g.u.) is a biantennary glycan with no LacdiNAc as shown by the lack of an *m/z* 572 ($HexNAc_2PC_1$) MS/MS B fragment and presence of one at *m/z* 531

($Hex_1HexNAcPC_1$) instead (Supplementary Figure 4G). Other hybrid and biantennnary phosphorylcholine-modified glycans ($Hex_{3-4}HexNAc_{4-6}Fuc_1PC_{1-2}$; Supplementary Figure 4 I-L) were also detected and the sub-terminal position of the phosphorylcholine groups in the $HexNAc_2PC_1$ motifs confirmed by sensitivity towards HEX-4 (inset in Supplementary Figure 4L). As shown by Western blotting, we could demonstrate also that the phosphorylcholine on insect cell line glycoproteins is recognised by an anti-phosphorylcholine monoclonal as well as by human C-reactive protein with some differences in relative binding patterns being observed (Supplementary Figure 4F).

3.6 Sulphated N-glycans

Based on fluorescence intensities, it can be estimated that about 1-5% of the three N-glycomes examined were accounted for by anionic glycans. Our previous study on mosquito glycans revealed the presence of sulphation of mannose or of fucose residues (13) and this was also observed here for the lepidopteran species. In general, sulphation of mannose results in relatively early elution times (as compared to the related non-sulphated structure) and in-source loss of one sulphate (either for the monosulphated structures in positive mode or the disulphated structures in negative mode). Thus, in negative mode analyses, disulphated glycans are detected as traces of the true parent $[M-2H+Na]^-$ ions, whereas monosulphated glycans showed only the $[M-H]^-$ ion. This differential effect can be observed for $Hex_3HexNAc_{2-3}Fuc_1S_{1-2}$ glycans in two different RP-amide fractions of 7.2 and 7.5 g.u. (*m/z* 1213 and 1416; Figure 6 A and D). Other than the detectable trace of disulphation (*m/z* 1315/1518; $\Delta m/z$ = 102), the inability to remove any mannose residues from the disulphated forms after the hexosaminidase digestion contrasted with α-mannosidase sensitivity of the related monosulphated glycans (Figures 6 B, C, E and F).

Similarly, a disulphated but non-fucosylated glycan ($Hex_3HexNAc_3S_2$; *m/z* 1372 as $[M-2H+Na]^-$) was also resistant to α-mannosidase treatment after β-hexosaminidase digestion (5.4 g.u.; Figure 6 G-I). The disulphation of the trimannosyl core could be concluded also from the negative mode fragments at *m/z* 870 and 667 absent upon fragmentation of the $[M-SO_3H]^-$ 'pseudomonosulphated' ion (Figure 6 M-O). Related monosulphated glycans (*m/z* 1067 and 1270) were digested down to a structure of *m/z* 905 ($Hex_2HexNAc_2S_1$) using β-hexosaminidase and α1,2/3-specific mannosidase digestion in series thereby proving that monosulphation of mannose tends to be associated with the α1,6-mannose residue (Figure 6 J-L); the MS/MS B fragment ions at *m/z* 241 and 403 also localise the sulphate on a hexose (Figure 6 Q-R). Verification of the structural relationship between the monosulphated core fucosylated and non-fucosylated glycans was also based on retention time shifts (Supplementary Figure 5) upon rechromatography following glycosidase digestion: after α1,2/3-mannosidase treatment, the majority of the *m/z* 1213 and 1416 glycans lost one hexose, yielding products of *m/z* 1051 and 1254 which eluted slightly later (i.e., where in the original 'anionic' chromatogram glycans were observed with the same *m/z*). Bovine α-fucosidase digestion then resulted in loss of the core fucose (with products of *m/z* 905 and 1108) and a shift to earlier elution, thereby again co-eluting with sulphated glycans found in the original anionic chromatogram. Consistent with the low quantities of sulphated glycans, it was possible to detect monosulphated core difucosylated glycans ($Hex_3HexNAc_{2-4}Fuc_2S_1$;

m/z 1359, 1562 and 1765) only in a PNGase A digest of High Five cells where higher amounts of sulphated structures were found *per se* (Supplementary Figure 6).

Even in the case of structures carrying longer antennae, e.g., with either type I LacNAc (*m/z* 1578 and 1781) or LacdiNAc units, the sulphate is still located on the α1,6-mannose as proven by digests with the 'lower-arm-specific' FDL β-hexosaminidase, α1,2/3-mannosidase and β1,3-specific galactosidase (Figure 7 A-D) with shifts in the negative mode MS/MS spectra culminating in B fragment ions at *m/z* 768 ($Hex_3HexNAc_1S_1$; Figure 7 H-K). A minor glycan with *m/z* 1987 in the negative ion mode carrying both phosphorylcholine and sulphate lost the phosphorylcholine and a terminal HexNAc when treated with hydrofluoric acid or *C. elegans* HEX-4 (Figure 7 E-G); due to the ease of the ionisation of phosphorylcholine in the positive mode, MS/MS of the 'in source' $[M-SO_3+H]^+$ ion before and after HEX-4 digestion revealed shifts in the *m/z* 572 and 734 B fragments to *m/z* 369 and 531 (i.e., from $Hex_{0-1}HexNAc_2PC_1$ to $Hex_{0-1}HexNAc_1PC_1$; Figure 7 M and N). Note that sulphate is resistant to hydrofluoric acid (unlike phosphate) and so the data verify that the 80 Da modification of insect glycans is sulphate; indeed, sulphated glycans modified with phosphorylcholine ($Hex_3HexNAc_4Fuc_{0-1}PC_1S_1$; *m/z* 1638 and its core fucosylated form of *m/z* 1784) only lost the zwitterion upon hydrofluoric acid treatment. Subsequent rechomatography resulted in later retention times and co-elution with the mannose sulphated structures of *m/z* 1473 and 1619.

A small subset of rather late eluting sulphated glycans yielded MS/MS fragments of around *m/z* 225, reminiscent of the sulphated fucose occurring in mosquitoes (13). As expected the *m/z* 1213 glycan (9.5 g.u. isomer) was resistant to bovine α-fucosidase, whereas it lost two hexose residues when incubated with jack bean mannosidase (Figure 7 O-Q); the MS/MS spectra of the glycan before and after mannosidase treatment showed a change in the predicted $^{O,2}A$ cross-ring cleavage fragment ions (at *m/z* 1034 and 710; Figure 7 R and S) and contrasted with those of the mannose sulphated version (Figure 7 T and U), thereby confirming sulphation of core fucose.

3.7 Glucuronylated N-glycans

We previously identified glucuronic acid on N-glycans from mosquitoes (13), a result which was compatible with the analysis of permethylated N-glycans from *Drosophila* embryos and S2 cells (11,38). Thus, we anticipated that lepidopteran species may also express glucuronylated N-glycans; our previous experience with off-line MALDI-TOF-MS suggested that these were detectable in negative mode, but fragmented best in positive mode. Thereby, the MS/MS spectra of monoglucuronylated N-glycans with *m/z* 1674/1676, 1877/1879 and 2080/2082 ($HexA_1Hex_4HexNAc_{3-5}Fuc_1$; as $[M-H]^-$ and $[M+H]^+$, see Figure 8 D, E, G and H) had signature B fragment ions at *m/z* 542 and 745 in positive mode ($HexA_1Hex_1HexNAc_{1-2}$) as well as at *m/z* 743 and 946 in negative mode ($HexA_1Hex_1HexNAc_{2-3}$) indicating the presence of a $HexA_1Hex_1$ unit linked to either one, two or three antennal HexNAc residues in series. Defining the modification on either the 'upper' or 'lower' arm was possible by determining the α1,2/3-mannosidase sensitivity or resistance of the *m/z* 1674 ($[M-H]^-$) glycans (Figure 8 A and B). As in mosquitoes and in accordance with their earlier elution, some glucuronylated glycans were also sulphated as

exemplified by the *m/z* 1754 glycan and losses of 80, 176 or 162 observed upon MS/MS (Figure 8F).

Interestingly, a number of glycans were modified with both phosphorylcholine and glucuronic acid; the presence of *m/z* 369 B fragment ions was considered indicative for the modification of a HexNAc with phosphorylcholine, whereas *m/z* 910 or 1113 B fragments were interpreted as being $HexA_1Hex_1HexNAc_{2\text{-}3}PC_1$ (Figure 8 I-O and Q) and so showing the modification with both moieties on the same antenna. The presence of *m/z* 531 fragments means that the phosphorylcholine is attached to a HexNAc directly linked to a hexose and so can be interpreted as a $PC_1GlcNAc_1Man_1$ motif; however, as *m/z* 540 and 908 are present together in negative mode MS/MS spectra (Figure 8M), the minimal fragment containing both phosphorylcholine and glucuronic acid is based on the Galβ1,3GalNAcβ1,4GlcNAc motif.

In two cases, an *m/z* 1113 fragment suggested that another HexNAc was present on some antenna (*m/z* 2101 and 2247; Figure 8K and N), thus these structures contain phosphorylcholine-modified variants of the elongated antennae present in the *m/z* 2080/2082 glycan (Figure 8 G and H) for which an *m/z* 946 negative B fragment ion ($HexA_1Hex_1HexNAc_3$) was observed. A further antennal variation was detected for the $HexA_1Hex_4HexNAc_4Fuc_2PC_1$ glycan (*m/z* 2190); in this case, *m/z* 515, 910 and 1056 fragments indicate that, as in the aforementioned *m/z* 1852 glycan, fucose and phosphorylcholine substitute the same HexNAc also in the context of a glucuronylated antenna (Figure 8 L). In case of a second isomer of $HexA_1Hex_4HexNAc_5Fuc_1PC_1$ (*m/z* 2247) and, in contrast to the aforementioned form with a long arm (Figure 8N), MS/MS resulted in an *m/z* 910 B fragment (Figure 8O). This glycan possessed on one arm an unsubstituted GlcNAc which was lost upon jack bean β-hexosaminidase treatment (Figure 8 C and P). Another biantennary glycan ($HexA_1Hex_4HexNAc_6Fuc_1PC_2$; *m/z* 2616), which is the largest glycan detected in this study, was sensitive to *C. elegans* HEX-4 treatment which resulted in a loss of the *m/z* 572 B fragment; together with the retention of the *m/z* 910 fragment, the data supported the proposal of a biantennary structure with two phosphorylcholine-modified LacdiNAc motifs, one of which carries a GlcAGal cap (Figures 8 Q and R).

MS/MS data alone could not solve the question regarding the third HexNAc in series on the *m/z* 2082, 2101 and 2247 structures described above. However, the *L. dispar* glycome contains a significant amount of an *m/z* 2247 isomer with an *m/z* 1113 MS/MS fragment (Figure 9B). This was subject to chemical and enzymatic treatments in combination with HPLC and MS/MS. As a first step the relevant fraction was incubated with hydrofluoric acid to remove the phosphorylcholine moiety. This resulted on HPLC in a shift of the majority of the fraction from 6.1 to 7.2 g.u. (Figure 9A), containing an *m/z* 2080/2082 species with *m/z* 946/948 B fragments detected in negative and positive modes (Figure 9C); this structure was in turn treated with β-glucuronidase and β1,3-specific galactosidase resulting in a further shift to 10 g.u. and the appearance of an *m/z* 610 fragment compatible with three HexNAc residues in series (MS/MS of *m/z* 1744; Figure 9D). Considering that insect glycolipids with a GlcAβ1,3Galβ1,3GalNAcα1,4GalNAcβ1,4GlcNAcβ-R motif are known (39), the 10 g.u. product was then treated with chicken liver α-*N*-acetylgalactosaminidase and loss of one

HexNAc residue was observed together with a shift to 9.5 g.u. (*m/z* 1541) and the replacement of the *m/z* 610 B fragment by one of *m/z* 407 (HexdiNAc; Figure 9E). As the first intermediate product co-eluted with the *m/z* 2082 glycan found in the original glycome, these data also can be taken as fixing its structure; as the final product of *m/z* 1541 of these digestions is known to have a 'lower' arm due to its elution time (27), the complex antenna is concluded to be on the α1,3-mannose, while the later-eluting *m/z* 2247 isomer (Figure 8N) is carrying this antenna probably on the α1,6-mannose as 'upper' arm modifications tend to result in increased retention on RP-type columns (40).

A similar approach was used to resolve the *L. dispar m/z* 2044 glycan eluting at 7.6 g.u. ($HexA_1Hex_4HexNAc_4Fuc_1PC_1$; see Figures 8M and 9G for negative and positive mode MS/MS). Hydrofluoric acid treatment followed by combined glucuronidase/galactosidase digestion resulted in increased HPLC elution time and a shift first to *m/z* 1879 and then to *m/z* 1541 (Figure 9 F, H and I); subsequent HEX-4 and arm-specific FDL treatments resulted in products of *m/z* 1338 and 1135 (Figure 9 J and K) indicative that the *m/z* 2044 glycan is carrying a lower arm antenna. As the intermediate *m/z* 1879 product in this experiment co-eluted with the $HexA_1Hex_4HexNAc_4Fuc_1$ in the original *L. dispar* glycome, it is concluded that the primary *m/z* 1879 structures from this and dipteran species possess a lower arm GlcAβ1,3Galβ1,3GalNAcβ1,4GlcNAc modification. The fucose residue of the *m/z* 2044 glycan is core α1,6-linked as judged by the shift to lower retention time upon bovine α-fucosidase digestion (Supplementary Figure 5) and by the presence of the *m/z* 446 Y1 fragment ($Fuc_1GlcNAc_1$-PA) in the MS/MS spectra obtained after removal of the phosphorylcholine residue (Figure 9H).

4 Discussion

4.1 Insect glycomics

The various N-glycomic features of the two lepidopteran species (*T. ni* and *L. dispar*) as well as the High Five cell line analysed here vastly expand the repertoire of proven N-glycan structures in insect larvae and insect cells. The richest N-glycome was probably that of *T. ni* larvae (in terms of the total number of glycans), but the *L. dispar* larval N-glycome contained structures present in neither the *T. ni* larvae nor High Five samples. The newly-discovered features include sulphation, glucuronylation, antennal fucosylation, phosphorylcholination and α-*N*-acetylgalactosaminylation (see Figure 10 for a summary). 'Charged' features previously noted on insect glycans such as sulphation of mannose or core fucose and glucuronylation of β1,3-galactose were found in both lepidopteran species (11,13,14); however, unlike *Drosophila* embryos (10), we detected no sialylation. In terms of the neutral N-glycans, the usual 'suspects', i.e., paucimannosidic structures with zero, one and two core fucose residues as well as oligomannosidic oligosaccharides are present in all three samples as were variants with non-reducing terminal GlcNAc residues. Noteworthy is also the presence of core α1,6-fucose on oligomannosidic structures with a substituted 'lower' or 'middle' arm α1,3-mannose, which reflects a presumed biosynthetic flexibility in these species. Their glycan antennae can be modified with 'type I' galactose (Galβ1,3GlcNAc) or with LacdiNAc (GalNAcβ1,4GlcNAc), in contrast to the 'type II' β1,4Gal often found in mammalian glycomes; a combination of both the type I and

LacdiNAc motifs (as a variant of Galβ1,3GalNAc or 'T antigen') has been found on royal jelly glycoproteins (41), whereas here we also found type I, T antigen and LacdiNAc motifs as the basic antennae for modification with fucose, glucuronic acid or phosphorylcholine.

The two modifications previously not found on insect N-glycans (phosphorylcholine and α-*N*-acetylgalactosamine) are not without precedent in the context of either (i) N-glycans of other invertebrates or (ii) other glycoconjugates from insects. Specifically, phosphorylcholine is a known component on the antennae of nematode and cestode N-glycans (29,42), nematode and annelid glycolipids (36,43), but also on O-glycans of a recombinant protein produced in *Sf9* cells (44); the related zwitterionic modifications of phosphoethanolamine and aminoethylphosphonate (lacking the three methyl groups associated with the amine) have been detected on insect glycolipids, O-glycans and glycoproteins (39,45–47). On the other hand, α1,4-linked *N*-acetylgalactosamine is present on insect glycolipids (39,45) and so we assume the same linkage (susceptible to chicken α-*N*-acetylgalactosaminidase) is occurring on the extended N-glycan antennae of *L. dispar.*

As mentioned in the introduction (1,2), the very first analyses of insect N-glycans were of mosquito cell lines, followed by examination of the glycans in *Drosophila* larvae, on bee venom glycoproteins as well as of three insect cell lines (not including High Five cells) (3–6). However, the vast majority of structural studies on insect glycosylation have been on individual recombinant glycoproteins expressed using the baculovirus system. Thus, the overall glycomic potential of many of the host cell lines, as well as of organisms from which they are originally derived, is underestimated; to our knowledge, other than mere monosaccharide compositions (48), only the glycans present on recombinant proteins produced from *T. ni* High Five cells or derivatives thereof have been analysed to date (e.g., antibodies, alkaline phosphatase or transferrin (7,34,49–51)) and not those of the endogenous proteins. There are only two reports on glycosylation of *L. dispar*, which merely indicate the presence of pauci- and oligomannosidic glycans also on recombinant glycoproteins produced by the LdEIta and Ld652Y cell lines (52,53). In terms of other lepidopteran larvae, a recent HPLC-based study of *Bombyx mori* (silkworm) fifth instar larvae also only demonstrated the presence of neutral N-glycans (54). Certainly, most such glycomic analyses have been based on the expectation of finding neutral N-glycans and neither anionic nor zwitterionic structures. Therefore, the presence of such N-glycans on heterologous glycoproteins produced by High Five cells cannot be ruled out, which may have repercussions for the biological activity of these recombinant products.

4.2 Analytical considerations

A combination of RP-amide HPLC, MALDI-TOF MS/MS as well as chemical and enzymatic digestions enabled an in-depth determination of the N-glycans of the two lepidopteran species. RP-HPLC offers isomeric information as length and position of the antennae or of the core fucose affects the retention time (55) and this has recently been verified for the RP-amide column used in the current study (24). Various trends could be observed which are either in keeping with the literature or can be newly-defined here on the basis of the current study: (i) in the case of oligomannosidic variants, the position of an α1,2-mannose on the middle 'B' arm was associated with higher retention times, while

α1,2-mannose on the lower A arm (α1,3-antenna) results in shift forward (see the isomers of *m/z* 1459, 1313, 1516 and 1662; Figure 1) as does α1,6-mannose as opposed to α1,3-mannose on the upper C arm (*m/z* 1151, 1297 and 1500); (ii) retarding retention was the presence of a GlcNAc on the α1,6-antenna ('upper' arm) for biantennary $Hex_3HexNAc_4Fuc_{0-1}PC_{0-1}$ isomers as opposed to those with LacdiNAc on the 'lower' α1,3-antenna (*m/z* 1395, 1541, 1560 and 1706; Figure 1), while (iii) an upper arm LacdiNAc modification results in even greater retardation of the elution as exemplified for *m/z* 1687 and 1706 isomers ($Hex_3HexNAc_4Fuc_2$ or $Hex_3HexNAc_4Fuc_1PC_1$; Figure 3). On the other hand, (iv) the addition of a phosphorylcholine residue reduces retention time on the RP-amide column (compare glycans of *m/z* 1687 or 1879 with 1852 or 2044; Figure 3); this latter effect was previously shown for nematode glycans on two fused-core columns, but contrasts with the behaviour of phosphorylcholine-modified glycans on a classical C18 material (56).

The shift due to phosphorylcholine can also be seen for sulphated glycans with compositions of $Hex_3HexNAc_4Fuc_{0-1}PC_{1-2}S_1$ (*m/z* 1638 and 1803 or 1784 and 1949 as compared to *m/z* 1473 and 1619; Figure 2). Also the length and position of the antenna has a major effect on elution as shown by the glycans of *m/z* 2247 or 2044 as well as their digestion products with seemingly the long antenna containing α-linked GalNAc resulting in reduced retention time as compared to the biantennary isomer (Figure 3). Finally, as previously observed for mosquito and oyster glycans (13,57), the presence of either a glucuronic acid or sulphate residue also reduces retention on the RP-amide column, whereas the retardation resulting from core α1,6-fucosylation is verified by the RP-HPLC elution behaviour after fucosidase treatment (Supplementary Figure 5).

Key also to the structural definitions were the data from MS/MS before and after treatments with glycosidases or with hydrofluoric acid resulting in gain and loss of diagnostic fragments. Another aspect is the pre-fractionation based on sequential PNGase F/A-release and solid-phase extraction with non-porous graphitised carbon into neutral or anionic PNGase-F and neutral PNGase-A-released pools; this meant that the low abundance anionic and difucosylated species were separately applied to the HPLC and so their MS signals were not suppressed by the more abundant neutral glycans. Thus, yet again and similar to our other recent studies (13,20,24,27,56,57), the off-line LC-MALDI-TOF-MS/MS approach we have applied has yielded a very rich data set which enables the definition of an N-glycome in its whole variety and indicative of a broader-than-expected glycomic potential. Indeed, when combining the data for all three pools (neutral or anionic F-released and A-released), about 100-120 N-glycans representing up to 80 different compositions are proposed for each original larval or cell line sample, rather than the 10-20 glycan structures commonly reported for individual insect-produced glycoproteins.

4.3 Insect glycogenomic potential

In order to generate the structures we observe in *T. ni* and *L. dispar*, a set of relevant glycoenzymes is required, some of which have been characterised previously from insects. Recent transcriptomics studies indicate that a range of relevant glycan-modifying enzymes is indeed expressed in High Five cells (30,58), but (other than two of the following examples)

most of these are not proven homologues in terms of activity. For modifications of the core region, α1,3- and α1,6-fucosyltransferases are described from *Drosophila melanogaster*, *Apis mellifera,* Sf9 cells*, Bombyx mori* and *Anopheles gambiae* (8,59–64). For basic antennal processing, *N*-acetylglucosaminyltransferases I and II from *Drosophila* and Sf9 cells (65,66) and Golgi hexosaminidases (known as FDL and specifically removing the GlcNAc transferred by *N*-acetylglucosaminyltransferase I) from *Drosophila* and lepidopteran sources (67–71), as well as Golgi mannosidases II and III from *Drosophila* and Sf9 cells (72,73), are known. Potential homologues of 'capping' enzymes include a β1,4-*N*-acetylgalactosaminyltransferase from *T. ni* (74,75), Lewis-type LacdiNAc-modifying α1,3-fucosyltransferases from *Apis mellifera* and *Anopheles gambiae* (60,64), GalNAc-modifying β1,3-galactosyltransferases from *Apis* (76), a LacdiNAc-modifying α1,4-*N*-acetylgalactosaminyltransferase from *Drosophila* (77) and β1,3-glucuronyltransferases from *Drosophila* (78). On the other hand, no mannose- or fucose-modifying sulphotransferases are known from any source and the only reported *in vitro* assay for a recombinant phosphorylcholinyltransferase is for the *mf1* protein from *Mycoplasma fermentans* (79). Thus, there are a number of unknown genes required to generate the full range of insect N-glycans; our glycomic analyses on dipteran and lepidopteran species indicate, however, that the antennal modifications do differ between orders and species within the class of the Insecta. Analyses of further orders are required in order to make general conclusions about the extent and evolution of anionic/zwitterionic glycan modifications in insects.

The major neutral $Man_{5\text{-}9}GlcNAc_2$ structures are typical oligomannosidic species expected as intracellular ER/Golgi biosynthetic intermediates; indeed, high levels of oligomannosidic glycans are known from analyses of whole mammalian cells (as opposed to the more complex cell surface glycomes (80)). Whereas more untypical $Man_{5,6}GlcNAc_2$ isomers retaining the B α1,2-mannose may also be intermediates captured during the analyses, structures lacking the A arm entirely are probably final products as they cannot be further processed to hybrid or complex structures due to the lack of the acceptor residue for *N*-acetylglucosaminyltransferase I. Nevertheless, some oligomannosidic glycans (even $Man_7GlcNAc_2$) detected in our analyses are core α1,6-fucosylated; thus, it can be expected that lepidopteran core α1,6-fucosyltransferases may have a relaxed substrate specificity – i.e., they do not absolutely require a GlcNAc residue to be present on the core α1,3-mannose (specifically the GlcNAc transferred by *N*-acetylglucosaminyltransferase I). *In vitro* studies to date have consistently indicated that such a GlcNAc residue is a pre-requisite for transfer of fucose by FUT-8 homologues (81); however, recently knock-out studies have shown that FUT-8 is responsible for the core α1,6-fucosylation of $Man_5GlcNAc_2$ in mammalian cells lacking *N*-acetylglucosaminyltransferase I (82). On the other hand, we find a relatively low degree of core α1,3/α1,6-difucosylation (insect-type anti-HRP epitope) in all three samples, which is comparable to the 1% level found in *Drosophila* (8); this contrasts with the higher levels (even 50%) found on secreted recombinant proteins expressed in High Five cells (51,83) or in the whole N-glycome (≈20%) of the *Drosophila* BG2-c6 neural cell line (35).

The demonstration that even a commonly-used cell line (High Five) is carrying unusual forms of glycosylation may have both positive and negative repercussions beyond the glycoanalytical ones, especially for biotechnological applications. On one hand, care must be taken if biological functions are being studied or if an insect 'factory' is used to prepare

therapeutic products as our results indicate that lepidopteran glycoproteins are recognised by human C-reactive protein; the variations in relative CRP binding between Sf9 and *T. ni* cell lines as compared to the TEPC15 antibody (Supplementary Figure 4) may be due to differences in epitope recognition. On the other hand, our study shows that there is potential to engineer nematode-like antennal glycosylation if required for the mimicking of parasitic immunomodulatory proteins. In any case, the adequate identification of the genes encoding glycan-modifying enzymes in insects is not just a matter of curiosity, but can enable production of interesting glycoforms of a range of proteins.

4.4 Biological considerations

The presence of such a diverse range of glycans raises questions regarding their function. Obviously, in contrast to parasitic helminths (84), the phosphorylcholine is not present in insects as a means of immunomodulation of mammalian host organisms (as insects are not parasites of mammalian species even if they are vectors of some human diseases), but interactions with the insects' own pathogens may well be modulated through such a zwitterionic modification. The extended antennae share motifs with glycolipids and it is known that both insect glycoproteins and glycolipids (as well as nematode glycolipids) are 'hijacked' as targets for *Bacillus thuringiensis* crystal toxins (85). Indeed, one study concluded that terminal α-GalNAc (as here unambiguously proven to be found on *L. dispar* N-glycans as a subterminal substitution) may be an epitope for the Cry1Ac toxin (86).

In a therapeutic context, certainly insect-derived glycoproteins are far from being as widely used as those from, e.g., Chinese hamster ovary cells, but some approved vaccines, such as FluBlok®, are produced using the baculovirus expression system (87); however, in terms of widespread use, it may rather be as *in vitro* tools that the glycosylation status of insect-produced proteins is currently relevant. Certainly, glycan-dependent false positive reactivity of a recombinant tapeworm antigen expressed in High Five cells has been reported (88).

Another question is as to why the dipteran and lepidopteran species examined to date share some glycan modifications, but not others. Is this due to environmental or life-style differences? Being related (in terms of species) is not associated with possessing an identical glycome: even humans and chimpanzees differ in terms of the absence or present of a type of sialic acid (89). However, deleting or supplementing aspects of the glycome may well reflect speciation and avoidance of pathogens or other selective advantage during evolution. Certainly, our identification of a wider range of glycan structures may enable a fresh look at the role of glycans in many processes in insects and other species.

Supplementary Data

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Dr. Axel Schopf (Universität für Bodenkultur Wien) for respectively cultivating the gypsy moth larvae. This work was supported by the Austrian Science Fund (FWF; grants P26662 to A.H., P25092 to D.P., P23922 to I.B.H.W and P25058 to K.P.).

References

1. Butters TD, Hughes RC. Isolation and characterisation of mosquito cell membrane glycoproteins. Biochim Biophys Acta. 1981; 640:655–671. [PubMed: 7213698]
2. Hsieh P, Robbins PW. Regulation of asparagine-linked oligosaccharide processing. Oligosaccharide processing in *Aedes albopictus* mosquito cells. J Biol Chem. 1984; 259:2375–2382. [PubMed: 6698972]
3. Williams PJ, Wormald MR, Dwek RA, Rademacher TW, Parker GF, Roberts DR. Characterisation of oligosaccharides from *Drosophila melanogaster* glycoproteins. Biochim Biophys Acta Gen Subj. 1991; 1075:146–153.
4. Kubelka V, Altmann F, März L. The asparagine-linked carbohydrate of honeybee venom hyaluronidase. Glycoconjugate J. 1995; 12:77–83.
5. Kubelka V, Altmann F, Staudacher E, Tretter V, März L, Hård K, Kamerling JP, Vliegenthart JFG. Primary structures of the N-linked carbohydrate chains from honeybee venom phospholipase A_2. Eur J Biochem. 1993; 213:1193–1204. [PubMed: 8504812]
6. Kubelka V, Altmann F, Kornfeld G, März L. Structures of the N-linked oligosaccharides of the membrane glycoproteins from three lepidopteran cell lines (Sf-21, IZD-Mb-0503, Bm-N). Arch Biochem Biophys. 1994; 308:148–157. [PubMed: 8311447]
7. Hsu TA, Takahashi N, Tsukamoto Y, Kato K, Shimada I, Masuda K, Whiteley EM, Fan JQ, Lee YC, Betenbaugh MJ. Differential *N*-glycan patterns of secreted and intracellular IgG produced in *Trichoplusia ni* cells. J Biol Chem. 1997; 272:9062–9070. [PubMed: 9083032]
8. Fabini G, Freilinger A, Altmann F, Wilson IBH. Identification of core α1,3-fucosylated glycans and the requisite fucosyltransferase in *Drosophila melanogaster*. Potential basis of the neural anti-horseradish peroxidase epitope. J Biol Chem. 2001; 276:28058–28067. [PubMed: 11382750]
9. Koles K, Irvine KD, Panin VM. Functional characterization of *Drosophila* sialyltransferase. J Biol Chem. 2004; 279:4346–4357. [PubMed: 14612445]
10. Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M. Dynamic developmental elaboration of N-linked glycan complexity in the *Drosophila melanogaster* embryo. J Biol Chem. 2007; 282:9127–9142. [PubMed: 17264077]
11. Aoki K, Tiemeyer M. The glycomics of glycan glucuronylation in *Drosophila melanogaster*. Methods Enzymol. 2010; 480:297–321. [PubMed: 20816215]
12. Geisler, C., Jarvis, D. Insect Cell Glycosylation Patterns in the Context of Biopharmaceuticals. Post-translational Modification of Protein Biopharmaceuticals. Walsh, G., editor. Wiley-VCH Verlag GmbH & Co; Weinheim, Germany: 2009. p. 165-191.
13. Kurz S, Aoki K, Jin C, Karlsson NG, Tiemeyer M, Wilson IBH, Paschinger K. Targetted release and fractionation reveal glucuronylated and sulphated N- and O-glycans in larvae of dipteran insects. J Proteomics. 2015; 126:172–188. [PubMed: 26047717]
14. Cabrera G, Salazar V, Montesino R, Tambara Y, Struwe WB, Lugo EL, Harvey DJ, Antoine L, Rincon M, Domon B, Mendez Martinez MD, et al. Structural characterization and biological implications of sulfated N-glycans in a serine protease from the neotropical moth *Hylesia metabus* (Cramer [1775]) (Lepidoptera: Saturniidae). Glycobiology. 2015; 26:230–250. [PubMed: 26537504]
15. Shorey HH, Andres LA, Hale RL. The biology of *Trichoplusia ni* (Lepidoptera: Noctuidae). I. Life history and behavior. Ann Entomol Soc Amer. 1962; 55:591–597.
16. Pogue, MG., Schaefer, PW. A review of selected species of Lymantria Hübner (1819) (Lepidoptera: Noctuidae: Lymantriinae) from subtropical and temperate regions of Asia, including the descriptions of three new species, some potentially invasive to North America. Forest Health Technology Enterprise Team; Morgantown, WV USA: 2007.
17. Wickham TJ, Davis T, Granados RR, Shuler ML, Wood HA. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. Biotechnol Prog. 1992; 8:391–396. [PubMed: 1369220]
18. Zhang X, Tiewsiri K, Kain W, Huang L, Wang P. Resistance of *Trichoplusia ni* to *Bacillus thuringiensis* Toxin Cry1Ac Is Independent of Alteration of the Cadherin-Like Receptor for Cry Toxins. PLoS ONE. 2012; 7:e35991. [PubMed: 22606242]

19. Bell, RA., Owens, CD., Shapiro, M., Tardif, JR. Development of mass-rearing technology. The gypsy moth: research toward integrated pest management. Doane, CC., McManus, ML., editors. US Department of Agriculture; Washington, DC: 1981. p. 599-633.
20. Eckmair B, Jin C, Abed-Navandi D, Paschinger K. Multi-step fractionation and mass spectrometry reveals zwitterionic and anionic modifications of the N- and O-glycans of a marine snail. Mol Cell Proteomics. 2016; 15:573–597. [PubMed: 26598642]
21. Paschinger K, Hykollari A, Razzazi-Fazeli E, Greenwell P, Leitsch D, Walochnik J, Wilson IBH. The N-glycans of *Trichomonas vaginalis* contain variable core and antennal modifications. Glycobiology. 2012; 22:300–313. [PubMed: 21983210]
22. Hykollari A, Paschinger K, Eckmair B, Wilson IBH. Analysis of Invertebrate and Protist N-Glycans. Methods Mol Biol. 2017; 1503:167–184. [PubMed: 27743366]
23. Hykollari A, Balog CI, Rendić D, Braulke T, Wilson IBH, Paschinger K. Mass spectrometric analysis of neutral and anionic N-glycans from a *Dictyostelium discoideum* model for human congenital disorder of glycosylation CDG IL. J Proteome Res. 2013; 12:1173–1187. [PubMed: 23320427]
24. Yan S, Wilson IBH, Paschinger K. Comparison of RP-HPLC modes to analyse the N-glycome of the free-living nematode *Pristionchus pacificus*. Electrophoresis. 2015; 36:1314–1329. [PubMed: 25639343]
25. Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, Prestegard JJ, et al. Symbol Nomenclature for Graphical Representations of Glycans. Glycobiology. 2015; 25:1323–1324. [PubMed: 26543186]
26. Dragosits M, Yan S, Razzazi-Fazeli E, Wilson IBH, Rendić D. Enzymatic properties and subtle differences in the substrate specificity of phylogenetically distinct invertebrate N-glycan processing hexosaminidases. Glycobiology. 2015; 25:448–464. [PubMed: 25488985]
27. Wilson IBH, Paschinger K. Sweet secrets of a therapeutic worm: Mass spectrometric N-glycomic analysis of *Trichuris suis*. Anal Bioanal Chem. 2016; 408:461–471. [PubMed: 26650734]
28. Paschinger K, Rendić D, Lochnit G, Jantsch V, Wilson IBH. Molecular basis of anti-horseradish peroxidase staining in *Caenorhabditis elegans*. J Biol Chem. 2004; 279:49588–49598. [PubMed: 15364955]
29. Paschinger K, Gonzalez-Sapienza GG, Wilson IBH. Mass spectrometric analysis of the immunodominant glycan epitope of *Echinococcus granulosus* antigen Ag5. Int J Parasitol. 2012; 42:279–285. [PubMed: 22342524]
30. Chen YR, Zhong S, Fei Z, Gao S, Zhang S, Li Z, Wang P, Blissard GW. Transcriptome responses of the host *Trichoplusia ni* to infection by the baculovirus Autographa californica multiple nucleopolyhedrovirus. J Virol. 2014; 88:13781–13797. [PubMed: 25231311]
31. Vaughn JL, Goodwin RH, Tompkins GJ, McCawley P. The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). In Vitro. 1977; 13:213–217. [PubMed: 68913]
32. Tomiya N, Lee YC, Yoshida T, Wada Y, Awaya J, Kurono M, Takahashi N. Calculated two-dimensional sugar map of pyridylaminated oligosaccharides: Elucidation of the jack bean α-mannosidase digestion pathway of $Man_9GlcNAc_2$. Analytical Biochemistry. 1991; 193:90–100. [PubMed: 2042746]
33. Hykollari A, Eckmair B, Voglmeir J, Jin C, Yan S, Vanbeselaere J, Razzazi-Fazeli E, Wilson IBH, Paschinger K. More than just oligomannose: An N-glycomic comparison of *Penicillium* species. Mol Cell Proteomics. 2016; 15:73–92. [PubMed: 26515459]
34. Rendić D, Wilson IBH, Paschinger K. The glycosylation capacity of insect cells. Croatica Chimica Acta. 2008; 8:7–21.
35. Rendić D, Linder A, Paschinger K, Borth N, Wilson IBH, Fabini G. Modulation of neural carbohydrate epitope expression in *Drosophila melanogaster* cells. J Biol Chem. 2006; 281:3343–3353. [PubMed: 16314421]
36. Lochnit G, Dennis RD, Ulmer AJ, Geyer R. Structural elucidation and monokine-inducing activity of two biologically active zwitterionic glycosphingolipids derived from the porcine parasitic nematode *Ascaris suum*. J Biol Chem. 1998; 273:466–474. [PubMed: 9417105]

37. Green ED, Adelt G, Baenziger JU, Wilson S, Van Halbeek H. The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz ^{1}H NMR spectroscopy. J Biol Chem. 1988; 263:18253–18268. [PubMed: 2461366]
38. Mabashi-Asazuma H, Kuo CW, Khoo KH, Jarvis DL. Modifying an Insect Cell N-Glycan Processing Pathway Using CRISPR-Cas Technology. ACS Chem Biol. 2015; 10:2199–2208. [PubMed: 26241388]
39. Dabrowski U, Dabrowski J, Helling F, Wiegandt H. Novel phosphorus-containing glycosphingolipids from the blowfly *Calliphora vicina* Meigen. Structural analysis by ^{1}H and ^{1}H[^{31}P]-edited NMR spectroscopy at 600 and 500 megahertz. J Biol Chem. 1990; 265:9737–9743. [PubMed: 2351669]
40. Tomiya N, Kurono M, Ishihara H, Tejima S, Endo S, Arata Y, Takahashi N. Structural analysis of N-linked oligosaccharides by a combination of glycopeptidase, exoglycosidases, and high-performance liquid chromatography. Anal Biochem. 1987; 163:489–499. [PubMed: 3661998]
41. Kimura M, Kimura Y, Tsumura K, Okihara K, Sugimoto H, Yamada H, Yonekura M. 350-kDa royal jelly glycoprotein (apisin), which stimulates proliferation of human monocytes, bears the β1-3galactosylated N-glycan: analysis of the N-glycosylation site. Biosci Biotechnol Biochem. 2003; 67:2055–2058. [PubMed: 14520005]
42. Paschinger K, Gutternigg M, Rendić D, Wilson IBH. The N-glycosylation pattern of *Caenorhabditis elegans*. Carbohydr Res. 2008; 343:2041–2049. [PubMed: 18226806]
43. Sugita M, Fujii H, Dulaney JT, Inagaki F, Suzuki M, Suzuki A, Ohta S. Structural elucidation of two novel amphoteric glycosphingolipids from the earthworm, *Pheretima hilgendorfi*. Biochim Biophys Acta. 1995; 1259:220–226. [PubMed: 8541328]
44. Gaunitz S, Jin C, Nilsson A, Liu J, Karlsson NG, Holgersson J. Mucin-type proteins produced in the *Trichoplusia ni* and *Spodoptera frugiperda* insect cell lines carry novel O-glycans with phosphocholine and sulfate substitutions. Glycobiology. 2013; 23:778–796. [PubMed: 23463814]
45. Seppo A, Moreland M, Schweingruber H, Tiemeyer M. Zwitterionic and acidic glycosphingolipids of the *Drosophila melanogaster* embryo. Eur J Biochem. 2000; 267:3549–3558. [PubMed: 10848971]
46. Maes E, Garenaux E, Strecker G, Leroy Y, Wieruszeski JM, Brassart C, Guerardel Y. Major O-glycans from the nest of *Vespula germanica* contain phospho-ethanolamine. Carbohydr Res. 2005; 340:1852–1858. [PubMed: 15963963]
47. Hård K, Van Doorn JM, Thomas-Oates JE, Kamerling JP, Van der Horst DJ. Structure of the Asn-linked oligosaccharides of apolipophorin III from the insect *Locusta migratoria*. Carbohydrate-linked 2-aminoethylphosphonate as a constituent of a glycoprotein. Biochemistry. 1993; 32:766–775. [PubMed: 8422381]
48. Park YI, Wood HA, Lee YC. Monosaccharide compositions of *Danaus plexippus* (monarch butterfly) and *Trichoplusia ni* (cabbage looper) egg glycoproteins. Glycoconj J. 1999; 16:629–638. [PubMed: 10972141]
49. Joshi L, Davis TR, Mattu TS, Rudd PM, Dwek RA, Shuler ML, Wood HA. Influence of baculovirus-host cell interactions on complex N-linked glycosylation of a recombinant human protein. Biotechnol Prog. 2000; 16:650–656. [PubMed: 10933841]
50. Ailor E, Takahashi N, Tsukamoto Y, Masuda K, Rahman BA, Jarvis DL, Lee YC, Betenbaugh MJ. N-glycan patterns of human transferrin produced in *Trichoplusia ni* insect cells: effects of mammalian galactosyltransferase. Glycobiology. 2000; 10:837–847. [PubMed: 10929010]
51. Palmberger D, Wilson IB, Berger I, Grabherr R, Rendic D. SweetBac: a new approach for the production of mammalianised glycoproteins in insect cells. PLoS One. 2012; 7:e34226. [PubMed: 22485160]
52. Kulakosky PC, Hughes PR, Wood HA. N-linked glycosylation of a baculovirus-expressed recombinant glycoprotein in insect larvae and tissue culture cells. Glycobiology. 1998; 8:741–745. [PubMed: 9621115]
53. Choi O, Tomiya N, Kim JH, Slavicek JM, Betenbaugh MJ, Lee YC. N-glycan structures of human transferrin produced by *Lymantria dispar* (gypsy moth) cells using the LdMNPV expression system. Glycobiology. 2003; 13:539–548. [PubMed: 12672704]

54. Kajiura H, Hamaguchi Y, Mizushima H, Misaki R, Fujiyama K. Sialylation potentials of the silkworm, *Bombyx mori; B. mori* possesses an active α2,6-sialyltransferase. Glycobiology. 2015; 25:1441–1453. [PubMed: 26306633]
55. Tomiya N, Awaya J, Kurono M, Endo S, Arata Y, Takahashi N. Analyses of N-linked oligosaccharides using a two-dimensional mapping technique. Anal Biochem. 1988; 171:73–90. [PubMed: 3407923]
56. Yan S, Jin C, Wilson IBH, Paschinger K. Comparisons of *Caenorhabditis* fucosyltransferase mutants reveal a multiplicity of isomeric N-glycan structures. J Proteome Res. 2015; 14:5291–5305. [PubMed: 26538210]
57. Kurz S, Jin C, Hykollari A, Gregorich D, Giomarelli B, Vasta GR, Wilson IBH, Paschinger K. Haemocytes and plasma of the eastern oyster (*Crassostrea virginica*) display a diverse repertoire of sulphated and blood group A-modified N-glycans. J Biol Chem. 2013; 288:24410–24428. [PubMed: 23824194]
58. Yu K, Yu Y, Tang X, Chen H, Xiao J, Su XD. Transcriptome analyses of insect cells to facilitate baculovirus-insect expression. Protein Cell. 2016; 7:373–382. [PubMed: 27017378]
59. Paschinger K, Staudacher E, Stemmer U, Fabini G, Wilson IBH. Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates. Glycobiology. 2005; 15:463–474. [PubMed: 15604090]
60. Rendić D, Klaudiny J, Stemmer U, Schmidt J, Paschinger K, Wilson IBH. Towards abolition of immunogenic structures in insect cells: characterization of a honey-bee (*Apis mellifera*) multi-gene family reveals both an allergy-related core α1,3-fucosyltransferase and the first insect Lewis-histo-blood-group-related antigen-synthesizing enzyme. Biochem J. 2007; 402:105–115. [PubMed: 17029591]
61. Juliant S, Harduin-Lepers A, Monjaret F, Catieau B, Violet ML, Cerutti P, Ozil A, Duonor-Cerutti M. The α1,6-fucosyltransferase gene (fut8) from the Sf9 lepidopteran insect cell line: insights into fut8 evolution. PLoS One. 2014; 9:e110422. [PubMed: 25333276]
62. Ihara H, Okada T, Ikeda Y. Cloning, expression and characterization of *Bombyx mori* α1,6-fucosyltransferase. Biochem Biophys Res Commun. 2014; 450:953–960. [PubMed: 24973708]
63. Minagawa S, Sekiguchi S, Nakaso Y, Tomita M, Takahisa M, Yasuda H. Identification of Core α1,3-Fucosyltransferase Gene From Silkworm: An Insect Popularly Used to Express Mammalian Proteins. J Insect Sci. 2015; 15
64. Kurz S, King JG, Dinglasan RR, Paschinger K, Wilson IBH. The fucomic potential of mosquitoes: Fucosylated N-glycan epitopes and their cognate fucosyltransferases. Insect Biochem Mol Biol. 2016; 68:52–63. [PubMed: 26617287]
65. Sarkar M, Schachter H. Cloning and expression of *Drosophila melanogaster* UDP-*N*-acetylglucosamine: α-6-D-mannoside-β-1,2-*N*-acetylglucosaminyltransferase I. Biol Chem. 2001; 382:209–217. [PubMed: 11308019]
66. Geisler C, Jarvis DL. Substrate specificities and intracellular distributions of three N-glycan processing enzymes functioning at a key branch point in the insect N-glycosylation pathway. J Biol Chem. 2012; 287:7084–7097. [PubMed: 22238347]
67. Léonard R, Rendić D, Rabouille C, Wilson IBH, Préat T, Altmann F. The *Drosophila fused lobes* gene encodes an *N*-acetylglucosaminidase involved in N-glycan processing. J Biol Chem. 2006; 281:4867–4875. [PubMed: 16339150]
68. Geisler C, Aumiller JJ, Jarvis DL. A *fused lobes* gene encodes the processing β-N-acetylglucosaminidase in Sf9 cells. J Biol Chem. 2008; 283:11330–11339. [PubMed: 18303021]
69. Geisler C, Jarvis DL. Identification of genes encoding N-glycan processing β-N-acetylglucosaminidases in *Trichoplusia ni* and *Bombyx mori*: Implications for glycoengineering of baculovirus expression systems. Biotechnol Prog. 2010; 26:34–44. [PubMed: 19882694]
70. Nomura T, Ikeda M, Ishiyama S, Mita K, Tamura T, Okada T, Fujiyama K, Usami A. Cloning and characterization of a β-N-acetylglucosaminidase (BmFDL) from silkworm *Bombyx mori*. J Biosci Bioeng. 2010; 110:386–391. [PubMed: 20547376]
71. Huo Y, Chen L, Qu M, Chen Q, Yang Q. Biochemical characterization of a novel β-N-acetylhexosaminidase from the insect *Ostrinia furnacalis*. Arch Insect Biochem Physiol. 2013; 83:115–126. [PubMed: 23703967]

72. Rabouille C, Kuntz DA, Lockyer A, Watson R, Signorelli T, Rose DR, van den Heuvel M, Roberts DB. The *Drosophila GMII* gene encodes a Golgi α-mannosidase II. J Cell Sci. 1999; 112:3319–3330. [PubMed: 10504337]
73. Kawar Z, Karaveg K, Moremen KW, Jarvis DL. Insect cells encode a class II α-mannosidase with unique properties. J Biol Chem. 2001; 276:16335–16340. [PubMed: 11279010]
74. van Die I, van Tetering A, Bakker H, van den Eijnden DH, Joziasse DH. Glycosylation in lepidopteran insect cells: identification of a β 1→4-N-acetylgalactosaminyltransferase involved in the synthesis of complex-type oligosaccharide chains. Glycobiology. 1996; 6:157–164. [PubMed: 8727788]
75. Vadaie N, Jarvis DL. Molecular cloning and functional characterization of a Lepidopteran insect β4-*N*-acetylgalactosaminyltransferase with broad substrate specificity, a functional role in glycoprotein biosynthesis, and a potential functional role in glycolipid biosynthesis. J Biol Chem. 2004; 279:33501–33508. [PubMed: 15173167]
76. Ichimiya T, Maeda M, Sakamura S, Kanazawa M, Nishihara S, Kimura Y. Identification of β1,3-galactosyltransferases responsible for biosynthesis of insect complex-type N-glycans containing a T-antigen unit in the honeybee. Glycoconj J. 2015; 32:141–151. [PubMed: 25931033]
77. Mucha J, Domlatil J, Lochnit G, Rendić D, Paschinger K, Hinterkörner G, Hofinger A, Kosma P, Wilson IBH. The *Drosophila melanogaster* homologue of the humna histo-blood group P^k gene encodes a glycolipid-modifying α1,4-*N*-acetylgalactosaminyltransferase. Biochem J. 2004; 382:67–74. [PubMed: 15130086]
78. Kim BT, Tsuchida K, Lincecum J, Kitagawa H, Bernfield M, Sugahara K. Identification and characterization of three *Drosophila melanogaster* glucuronyltransferases responsible for the synthesis of the conserved glycosaminoglycan-protein linkage region of proteoglycans. Two novel homologs exhibit broad specificity toward oligosaccharides from proteoglycans, glycoproteins, and glycosphingolipids. J Biol Chem. 2003; 278:9116–9124. [PubMed: 12511570]
79. Ishida N, Irikura D, Matsuda K, Sato S, Sone T, Tanaka M, Asano K. Molecular cloning and expression of a novel cholinephosphotransferase involved in glycoglycerophospholipid biosynthesis of Mycoplasma fermentans. Curr Microbiol. 2009; 58:535–540. [PubMed: 19219498]
80. Hamouda H, Kaup M, Ullah M, Berger M, Sandig V, Tauber R, Blanchard V. Rapid analysis of cell surface N-glycosylation from living cells using mass spectrometry. J Proteome Res. 2014; 13:6144–6151. [PubMed: 25348702]
81. Wilson JR, Williams D, Schachter H. The control of glycoprotein synthesis *N*-acetylglucosamine linkage to a mannose residue as a signal for the attachment of L-fucose to the asparagine-linked *N*-acetylglucosamine of glycopeptide from α_1-acid glycoprotein. Biochem Biophys Res Commun. 1976; 72:909–916. [PubMed: 985526]
82. Yang Q, Wang LX. Mammalian α1,6-fucosyltransferase (FUT8) is the sole enzyme responsible for the N-acetylglucosaminyltransferase I-independent core fucosylation of high-mannose N-glycans. J Biol Chem. 2016; 291:11064–11071. [PubMed: 27008861]
83. Palmberger D, Rendic D, Tauber P, Krammer F, Wilson IBH, Grabherr R. Insect cells for antibody production: evaluation of an efficient alternative. J Biotechnol. 2011; 153:160–166. [PubMed: 21477625]
84. Goodridge HS, Deehan MR, Harnett W, Harnett MM. Subversion of immunological signalling by a filarial nematode phosphorylcholine-containing secreted product. Cell Signal. 2005; 17:11–16. [PubMed: 15451020]
85. van Frankenhuyze K. Cross-order and cross-phylum activity of *Bacillus thuringiensis* pesticidal proteins. J Invertebr Pathol. 2013; 114:76–85. [PubMed: 23747826]
86. Sarkar A, Hess D, Mondal HA, Banerjee S, Sharma HC, Das S. Homodimeric alkaline phosphatase located at *Helicoverpa armigera* midgut, a putative receptor of Cry1Ac contains α-GalNAc in terminal glycan structure as interactive epitope. J Proteome Res. 2009; 8:1838–1848. [PubMed: 19714811]
87. Yang LP. Recombinant trivalent influenza vaccine (FluBlok®): a review of its use in the prevention of seasonal influenza in adults. Drugs. 2013; 73:1357–1366. [PubMed: 23928902]
88. Hancock K, Narang S, Pattabhi S, Yushak ML, Khan A, Lin SC, Plemons R, Betenbaugh MJ, Tsang VC. False positive reactivity of recombinant, diagnostic, glycoproteins produced in High

Five insect cells: effect of glycosylation. J Immunol Methods. 2008; 330:130–136. [PubMed: 17868684]

89. Gagneux P, Varki A. Genetic differences between humans and great apes. Mol Phylogenet Evol. 2001; 18:2–13. [PubMed: 11161737]

Europe PMC Funders Author Manuscripts

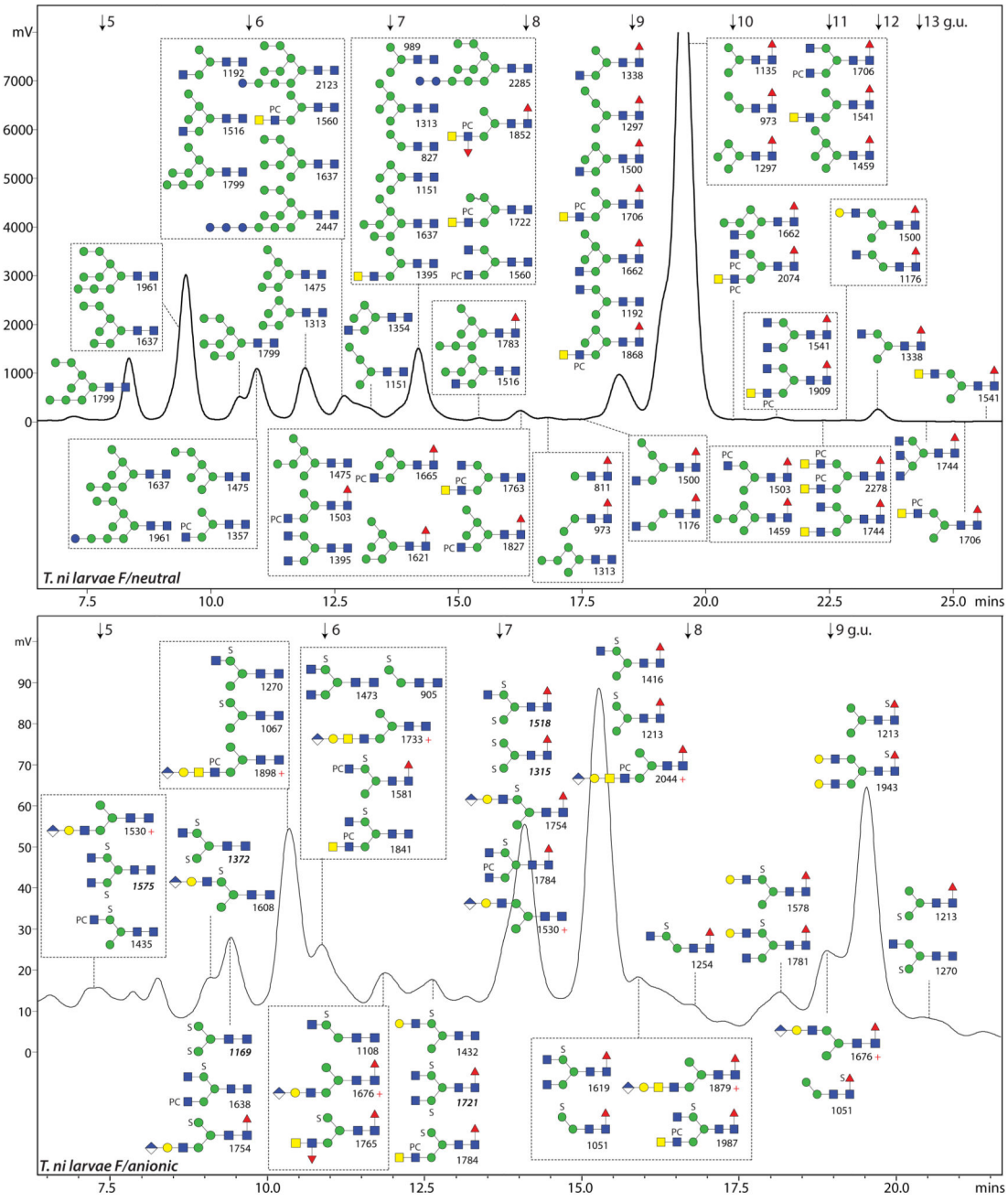

Figure 1. The N-glycome of *T. ni* fifth instar larvae.

RP-amide HPLC chromatograms of the PNGase F-released neutral and anionic pools (pyridylaminated) annotated with the structures concluded from co-elution, MS, MS/MS and digest data. Whereas about 30% of the neutral glycome was injected, 80% of the anionic glycome was applied, which is indicative of a 30-fold difference in amount; fluorescence intensity (320/400 nm) is given in mV. Structures are shown in the order of abundance in the relevant fraction (most abundant either shown uppermost or, in the case of boxes, at the top left) according to the Symbol Nomenclature for Glycans (25), whereby PC and S indicate

phosphorylcholine and sulphate (for glycosidic linkage information, refer to insets in Figure 2). The annotated *m/z* values in the upper panel are for protonated forms detected as $[M+H]^+$ in the positive ion mode (those in the lower panel are either positive (indicated with '+' for glucuronylated glycans lacking sulphate) or negative (either $[M-H]^-$ or, in italics for disulphated glycans, $[M-2H+Na]^-$). The elution positions of the isomaltose standards are indicated with arrows (5-13 glucose units; g.u.). The PNGase A-released *T. ni* glycome is shown in Supplementary Figure 1.

Europe PMC Funders Author Manuscripts

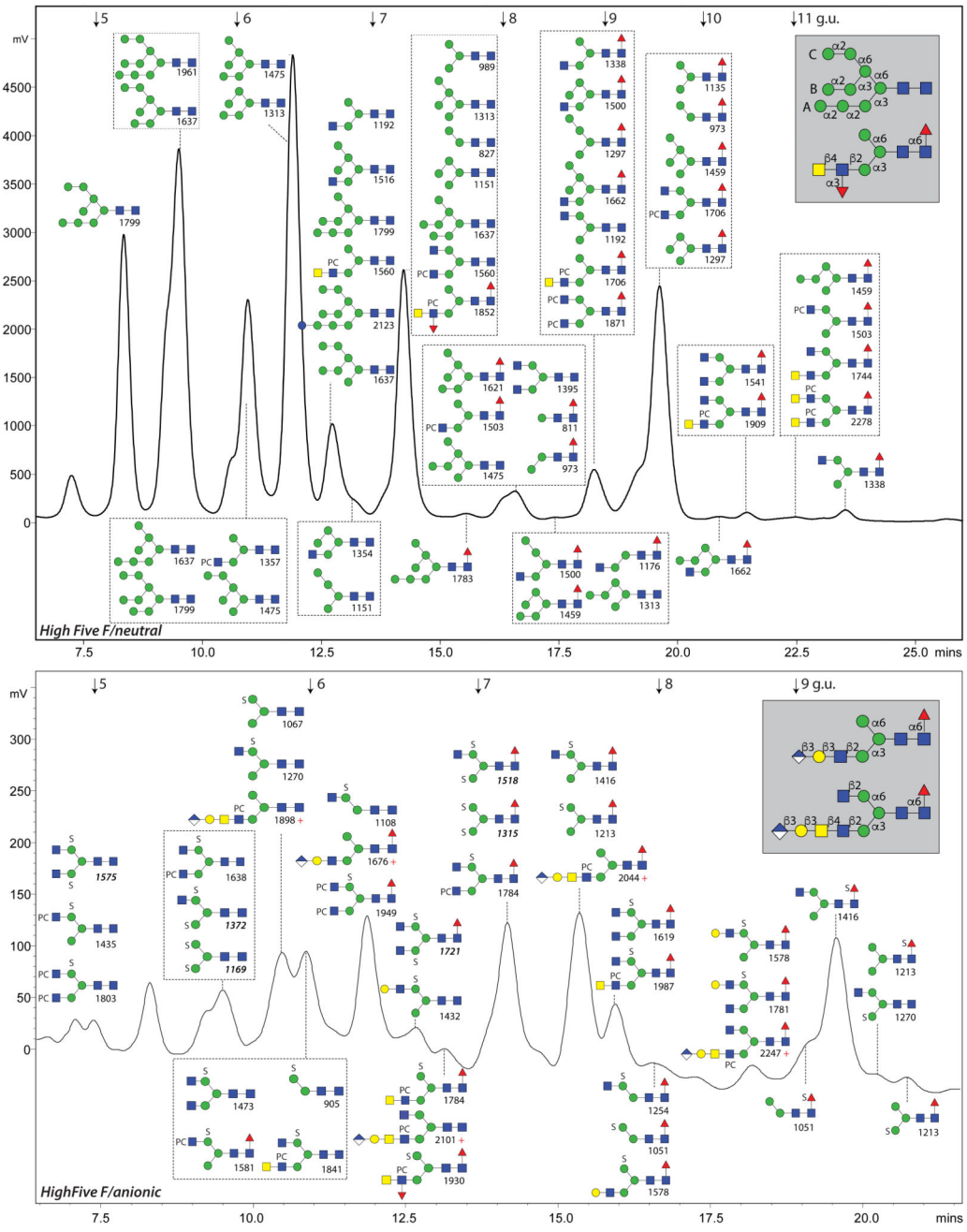

Figure 2. The N-glycome of High Five cells.

RP-amide HPLC chromatograms of the PNGase F-released neutral and anionic pools (pyridylaminated) annotated with the concluded structures and their *m/z* values (for further explanations, see legend to Figure 1). The PNGase A-released High Five glycome is shown in Supplementary Figure 2. Linkages for typical oligomannosidic (including the A, B and C branches), hybrid and complex glycans are shown in the grey boxes.

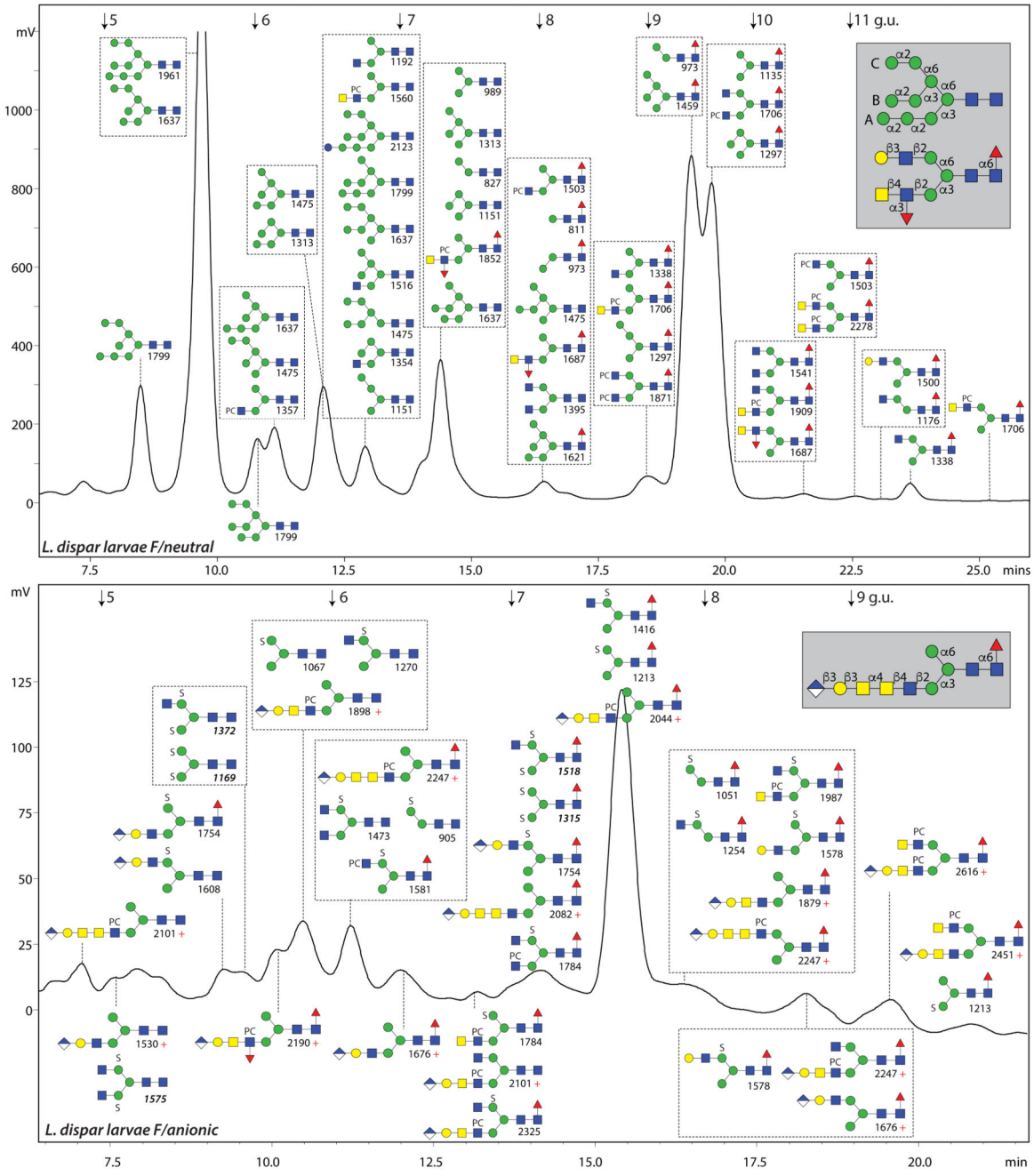

Figure 3. The N-glycome of *L. dispar* larvae.

RP-amide HPLC chromatograms of the PNGase F-released neutral and anionic pools (pyridylaminated) annotated with the concluded structures and their *m/z* values (for further explanations, see legend to Figure 1). The PNGase A-released *L. dispar* glycome is shown in Supplementary Figure 1. Whereas phosphorylcholine-modified glycans are more abundant in the *T. ni* neutral pool (see Figure 1), 'extra long' antennae with α-*N*-acetylgalactosamine were found with and without fucose and/or phosphorylcholine in the *L. dispar* anionic pool.

Linkages for example oligomannosidic, hybrid and complex glycans are shown in the grey boxes.

Europe PMC Funders Author Manuscripts

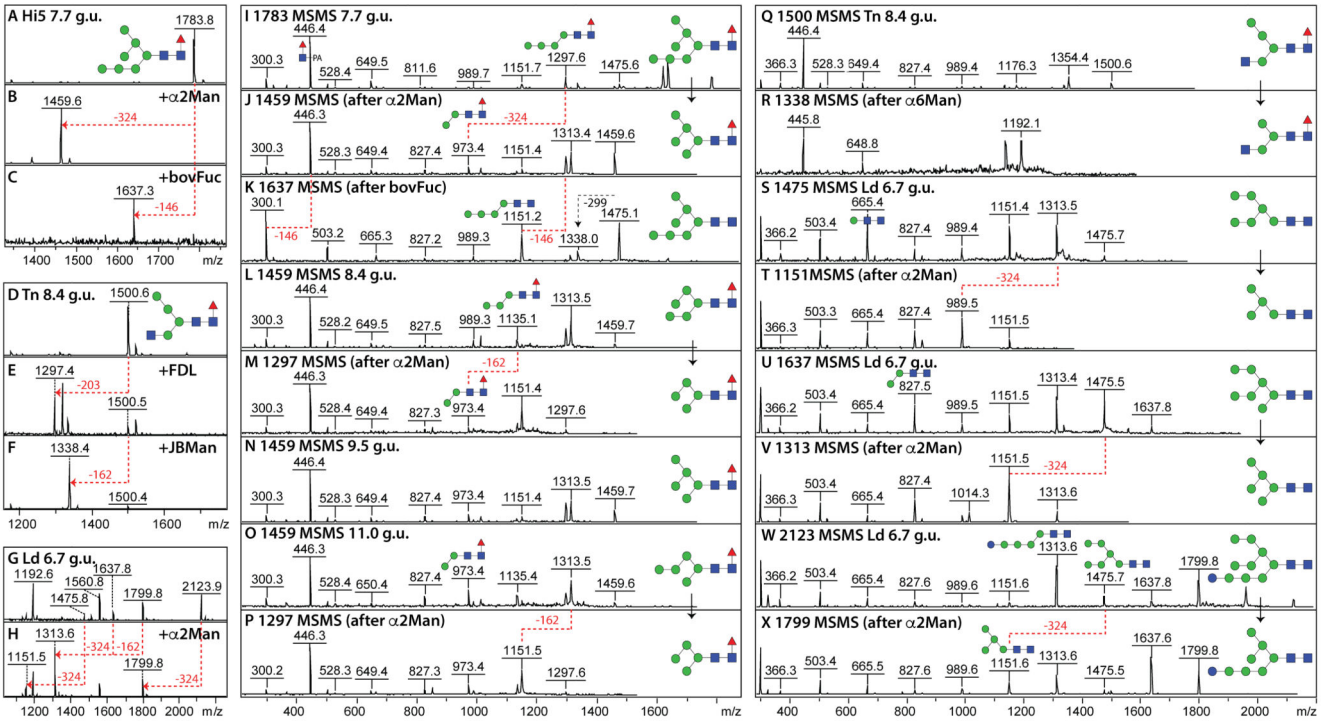

Figure 4. Structural determination of unusual fucosylated and oligomannosidic N-glycans.

(**A-H**) Pyridylaminated neutral N-glycans from different RP-amide fractions (indicated by Hi5, Tn and Ld as well as the g.u.) were subject to positive mode MALDI-TOF-MS before and after enzymatic digestion with either *Aspergillus* α1,2-specific mannosidase (α2Man), *Xanthomonas* α1,6-specific mannosidase (α6Man), jack bean α-mannosidase (JBMan), bovine α-fucosidase (bovFuc) or lower-arm-specific FDL β-hexosaminidase (FDL); the corresponding MS/MS spectra are shown in I-X. (**I-P**) MS/MS of fucosylated structures show an intense m/z 446 fragment due to $HexNAc_1Fuc_1$-PA, absent upon digestion with bovine α-fucosidase (see K); note that the *m/z* 1459 structure at 9.5 g.u. (see N) was not sensitive to α1,2-mannosidase as it is based on the 'Golgi' biosynthetic $Man_5GlcNAc_2$, whereas most of the presented structures are sensitive to this enzyme revealing thereby unusual fucosylated and mannosidic structures (see L and O). (**Q-X**) MS/MS showing the effect of selected mannosidase digestions on hybrid (*m/z* 1500) and oligomannosidic glycans (*m/z* 1475, 1637 and 2123). Key alterations in MS spectra of $[M+H]^+$ ions are shown with red dashed lines indicating loss of Fuc (146), Man (162 or 324) or GlcNAc (203); in the case of the MS/MS spectra, selected digestion-dependent shifts in significant fragments as well as their proposed structures are indicated. Fuller-range MS for panels G and H are shown in Supplementary Figure 3.

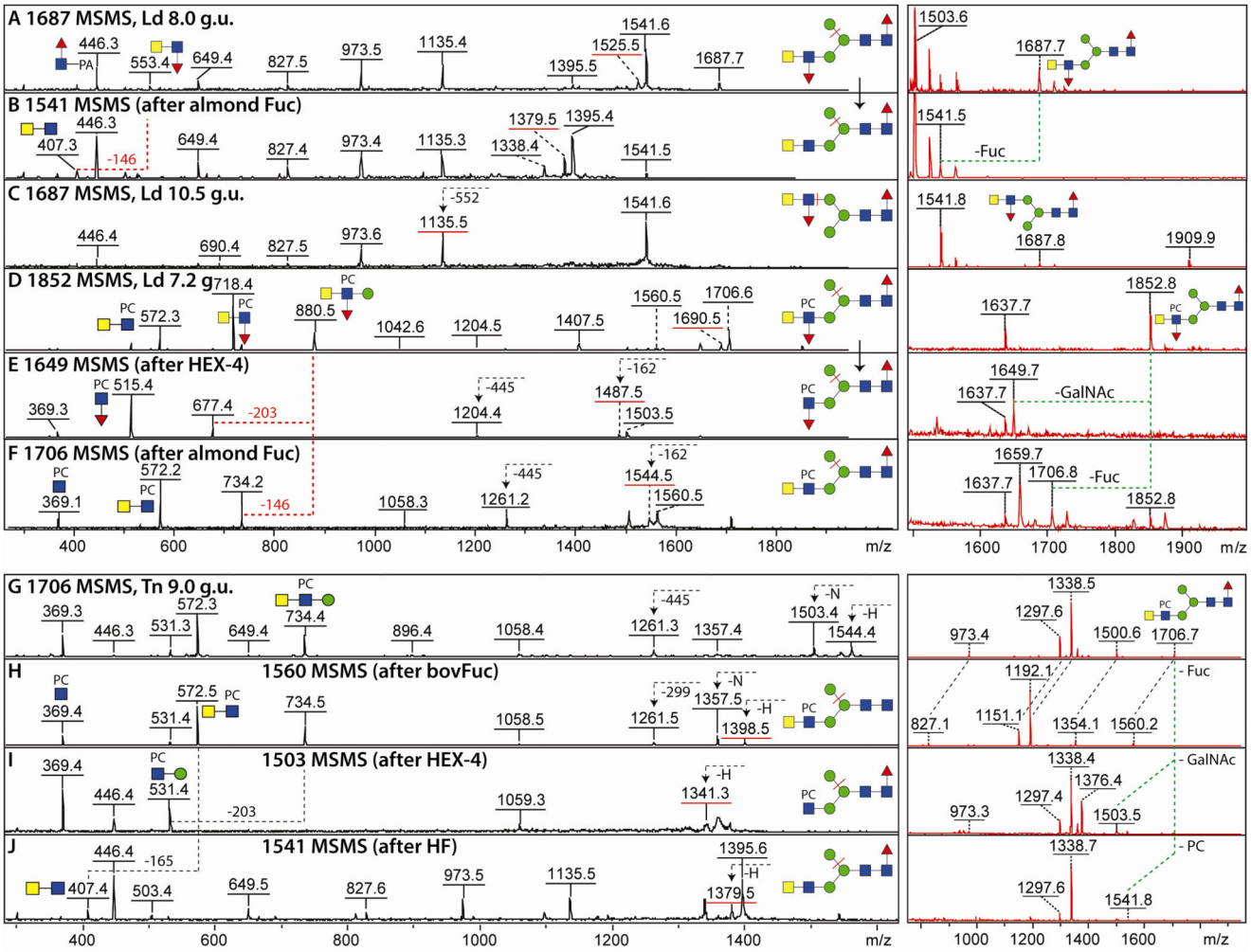

Figure 5. Characterisation of neutral and zwitterionic antennal modifications of *L. dispar* and *T. ni* glycans.

Selected positive mode MALDI-TOF MS and MS/MS spectra are shown for pyridylaminated N-glycans from different fractions (as indicated by the g.u.) before and after enzymatic digestions. Key B- and Y-fragments (all in the protonated form) are annotated with structures; those Y-ions predicted to result from loss of the α1,6-arm are underlined with a red bar, while *m/z* 369 or 572 fragments depend on the substitution of either a single GlcNAc or a LacdiNAc. The inset MS spectra in red (partly zoomed) are annotated with dashed lines indicating the relevant substrate-product shifts; only protonated adducts are annotated. (**A-C**) Two isomers of $Hex_3HexNAc_4Fuc_2$ (*m/z* 1687) with different elution times display MS/MS spectra both diverge from that of another isomer found in the PNGase A-digest of High Five glycoproteins (Supplementary Figure 2M). (**D-F**) MS/MS spectra of $Hex_3HexNAc_4Fuc_2PC_1$ (*m/z* 1852) before and after specific *C. elegans* HEX-4 β-*N*-acetylgalactosaminidase or almond α1,3-fucosidase treatment. (**G-J**) MS/MS of a $Hex_3HexNAc_4Fuc_1PC_1$ glycan (*m/z* 1706) before and after bovine α-fucosidase (removing the core fucose), *C. elegans* HEX-4 β-*N*-acetylgalactosaminidase (removing the terminal GalNAc) or hydrofluoric acid (removing phosphorylcholine). Buffers in some enzyme

preparations tend to result in shifts to sodiated or potassiated adducts ($\Delta m/z$ 22 or 38). Fuller-range MS for five of the right hand MS panels are shown in Supplementary Figure 3 and further examples of neutral and zwitterionic antennal modifications are shown in Supplementary Figure 4.

Europe PMC Funders Author Manuscripts

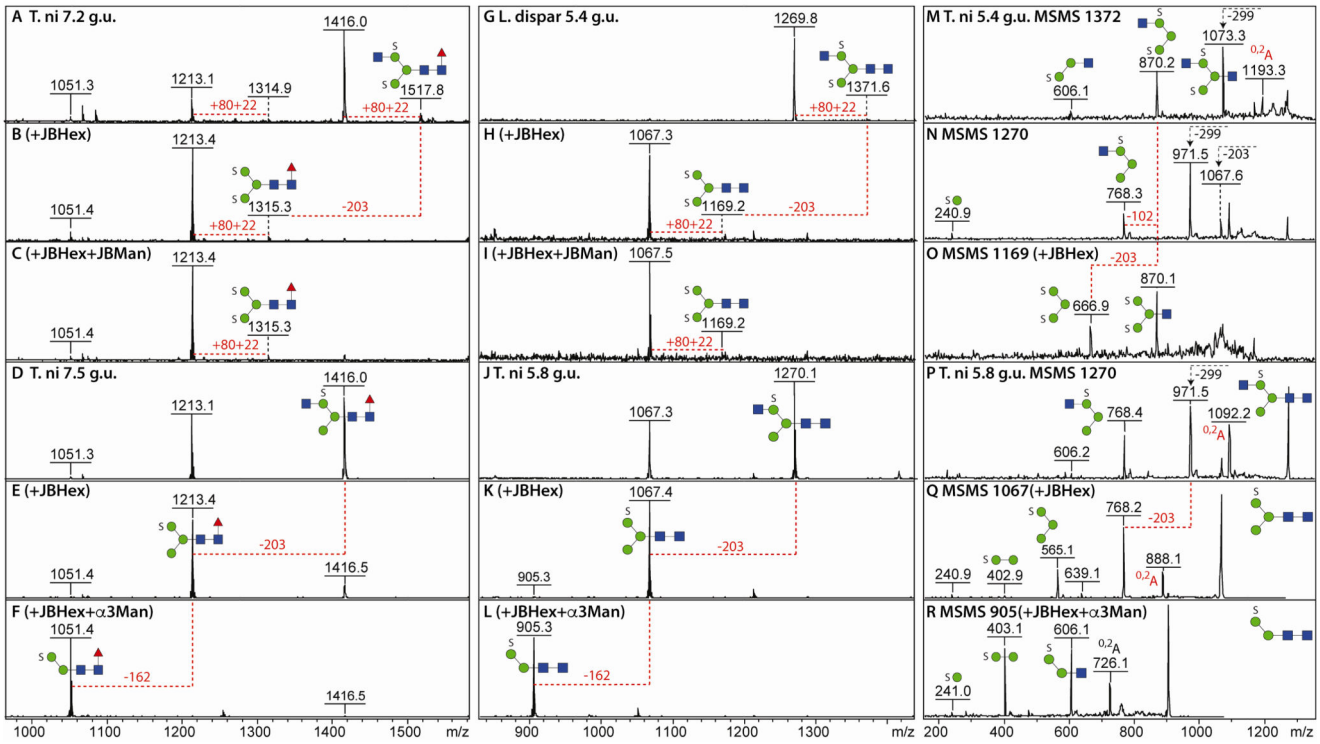

Figure 6. Characterisation of sulphated N-glycans from *T. ni* and *L. dispar* larvae.

Negative ion mode MALDI-TOF MS of fractions from the anionic pools before and after α-mannosidase and β-hexosaminidase digests of disulphated (**A-C and G-I**) and monosulphated (**D-F and J-L**) N-glycans show the sensitivity of the latter (and resistance of the former) to mannosidase treatments; the monosulphated glycans were detected as $[M-H]^-$ ions, while the disulphated glycans which eluted earlier show a trace of the $[M-2H+Na]^-$ ion; the dominant detected form of a disulphated glycan was however an in-source $[M-SO_3H]^-$ fragment ion ($\Delta m/z$ = 102; i.e., loss of sodium and sulphate). (**M-O**) Negative mode MS/MS of a disulphated glycan before and after jack bean β-hexosaminidase treatment as well as of its in-source fragment. (**P-R**) Negative mode MS/MS of a monosulphated glycan before and after jack bean β-hexosaminidase and *Xanthomonas* α1,2/3-mannosidase treatments. Key alterations in MS and MS/MS spectra are shown with red dashed lines indicating loss of Man (162) or GlcNAc (203).

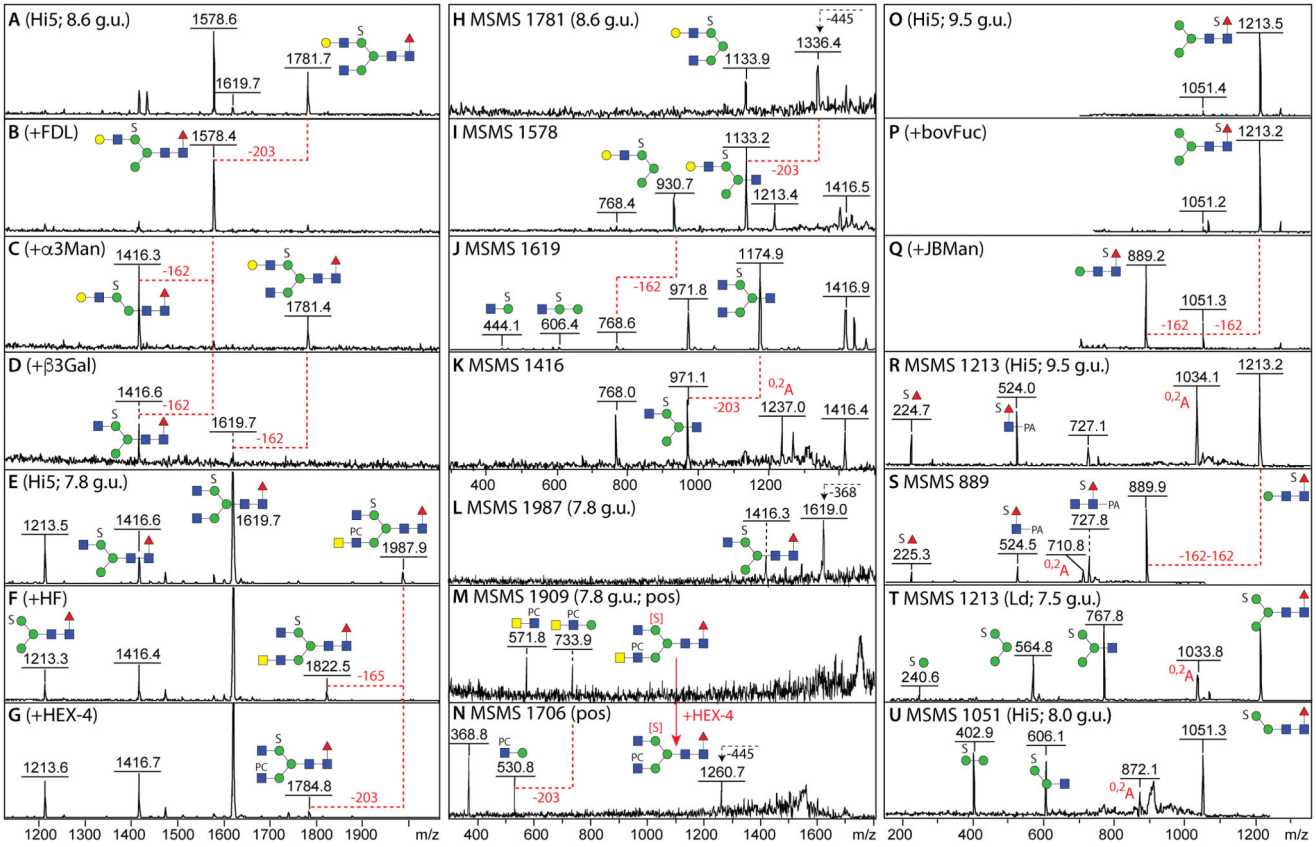

Figure 7. Characterisation of selected sulphated N-glycans from High Five cells.

(**A-D**) Negative mode MALDI-TOF MS of an anionic fraction before and after FDL β-hexosaminidase, *Xanthomonas* α1,2/3-mannosidase or *Xanthomonas* β1,3-galactosidase (the corresponding MS/MS are shown in **H-K**; the *m/z* 1237 ^{O,2}A ion results from cross-ring cleavage of the reducing-terminal GlcNAc). (**E-G**) Negative mode MALDI-TOF MS of an anionic fraction before and after hydrofluoric acid or *C. elegans* HEX-4 β-*N*-acetylgalactosaminidase. (**L-N**) The MS/MS of the *m/z* 1987 glycan [M-H]$^-$ as well as of the *m/z* 1909 and 1706 in-source fragment ions ([M-SO_3+H]$^+$; before and after HEX-4 treatment analysed in positive mode) show the presence of terminal GalNAc; the position of the sulphate lost during ionisation in positive mode is indicate by [S] in red. (**O-Q**) Negative mode MS of an *m/z* 1213 isomer before and after bovine fucosidase and jack bean mannosidase treatments, whereby the corresponding negative mode MS/MS (**R and S**) show the presence of an *m/z* 225 ion compatible with sulphation of the core fucose residue. (**T and U**) For comparative purposes, negative mode MS/MS of glycans with sulphation of mannose (*m/z* 1213 and 1051) are shown with similar $Hex_{1\text{-}3}S_1$ B fragment ions (*m/z* 241, 403 and 565) as related non-core-fucosylated glycans (see Figure 6Q and R).

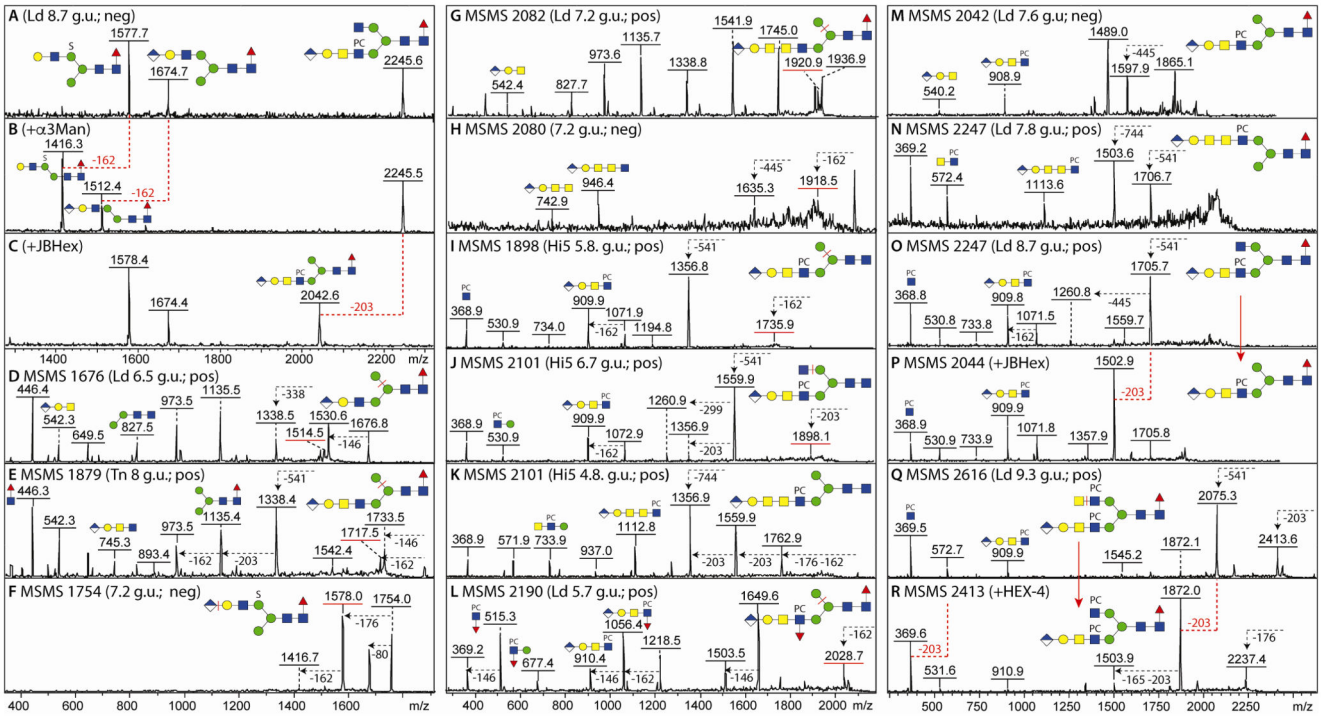

Figure 8. Characterisation of glucuronylated N-glycans.

(A-C) MALDI-TOF MS negative (neg) mode of an anionic fraction before and after α1,2/3-mannosidase or unspecific β-hexosaminidase treatments; the positive mode MS/MS of the *m/z* 2247 glycan as well as the *m/z* 2044 digestion product are shown in **O and P**. (**D and E**) Positive mode MS/MS of $HexA_1Hex_4HexNAc_{3-4}Fuc_1$ (*m/z* 1676 and 1879) shows the presence of the *m/z* 542 B-fragment characteristic of glucuronylated antennae. (**F**) Negative ion mode MS/MS of a sulphated/glucuronylated glycan (*m/z* 1754) shows only parental losses. (**G and H**) Positive and negative ion mode MS/MS spectra of $HexA_1Hex_4HexNAc_5Fuc_1$ are indicative of an unusually long lower arm antenna. (**I-L**) Positive mode MS/MS of phosphorylcholinylated/glucuronylated glycans showing the presence of the *m/z* 910 and 1113 $HexA_1Hex_1HexNAc_{2-3}PC_1$ fragments as compared to one carrying an antennal Lewis-like fucose residue as shown by the *m/z* 1056 fragment. (**M**) Negative mode MS/MS showing *m/z* 540 and 908 fragments indicative of modification of the same antennae with phosphorylcholine and glucuronic acid (see also Figure 9 for other analyses of this glycan). (**N-R**) Positive mode MS/MS of two isomers of $HexA_1Hex_4HexNAc_5Fuc_1PC_1$ and one of $HexA_1Hex_4HexNAc_6Fuc_1PC_2$ (*m/z* 2247 and 2616), also after hexosaminidase treatments (*m/z* 2044 and 2413), showing the presence of unsubstituted GlcNAc or GalNAc residues on one antenna. Key losses, alterations in spectra or fragment ions are annotated; Y-ions predicted to result from loss of mannose, HexNAc or glucuronic acid from the α1,6-arm are underlined with a red bar.

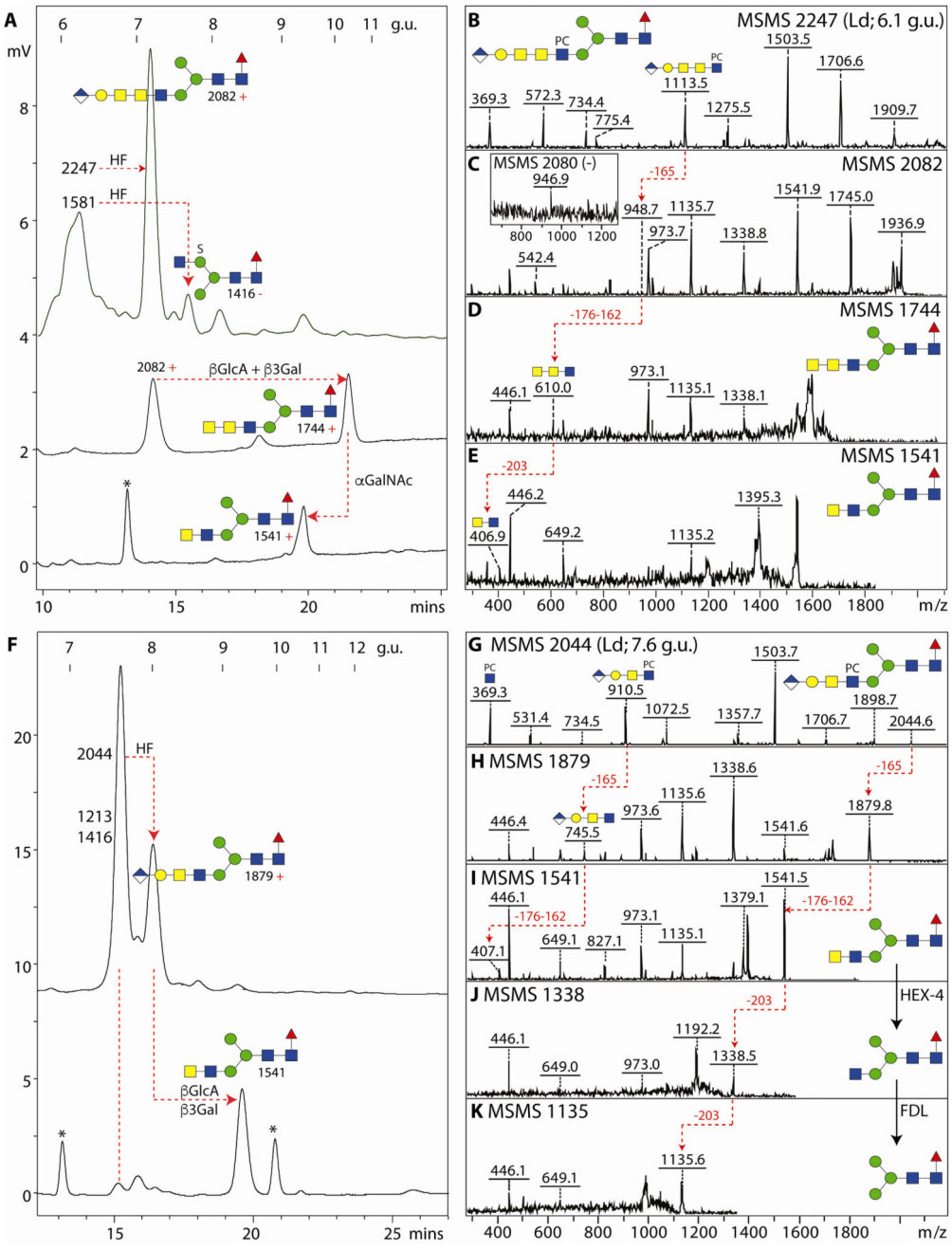

Figure 9. Determination of the isomeric structure of two glucuronylated zwitterionic N-glycans from *L. dispar*.

(A) RP-amide HPLC of the 6.1 g.u. fraction after hydrofluoric acid treatment and consecutive combined β-glucuronidase/β1,3-galactosidase and α-*N*-acetylgalactosaminidase digestions. (**B-E**) Positive mode MS/MS of the original *m/z* 2247 glycan (a third isomer of $HexA_1Hex_4HexNAc_5Fuc_1PC_1$; see also Figure 8 N and O) and the digestion products showing the progression from an antennal *m/z* 1113 fragment to one of *m/z* 407 (red dashed lines); additionally, the m/z 1581 glycan ($Hex_3HexNAc_3Fuc_1PC_1S_1$) present in the 6.1 g.u.

fraction was converted to one of 6.6 g.u. (*m/z* 1416 in negative mode) by the action of hydrofluoric acid (loss of phosphorylcholine). (**F**) RP-amide HPLC of the 7.7 g.u. fraction after hydrofluoric acid treatment and consecutive combined β-glucuronidase/β1,3-galactosidase digestion; the sulphated *m/z* 1213 and 1416 glycans ($Hex_3HexNAc_{2\text{-}3}Fuc_1S_1$) are resistant to hydrofluoric acid; contaminant non-glycan peaks in the digests are indicated with asterisks. (**G-K**) Positive mode MS/MS of the original *m/z* 2044 glycan (for the negative mode MS/MS of $HexA_1Hex_4HexNAc_4Fuc_1PC_1$ refer to Figure 8 M) and the digestion products showing the progression from an antennal *m/z* 910 B-fragment to one of *m/z* 407 as well as the subsequent loss of the 'lower' arm LacdiNAc upon treatment with the specific HEX-4 and FDL hexosaminidases (red dashed lines). The mannosidase and fucosidase sensitivities of the co-eluting sulphated glycans and fucosidase sensitivity of the *m/z* 2044 glycan are shown in Supplementary Figure 5.

Europe PMC Funders Author Manuscripts

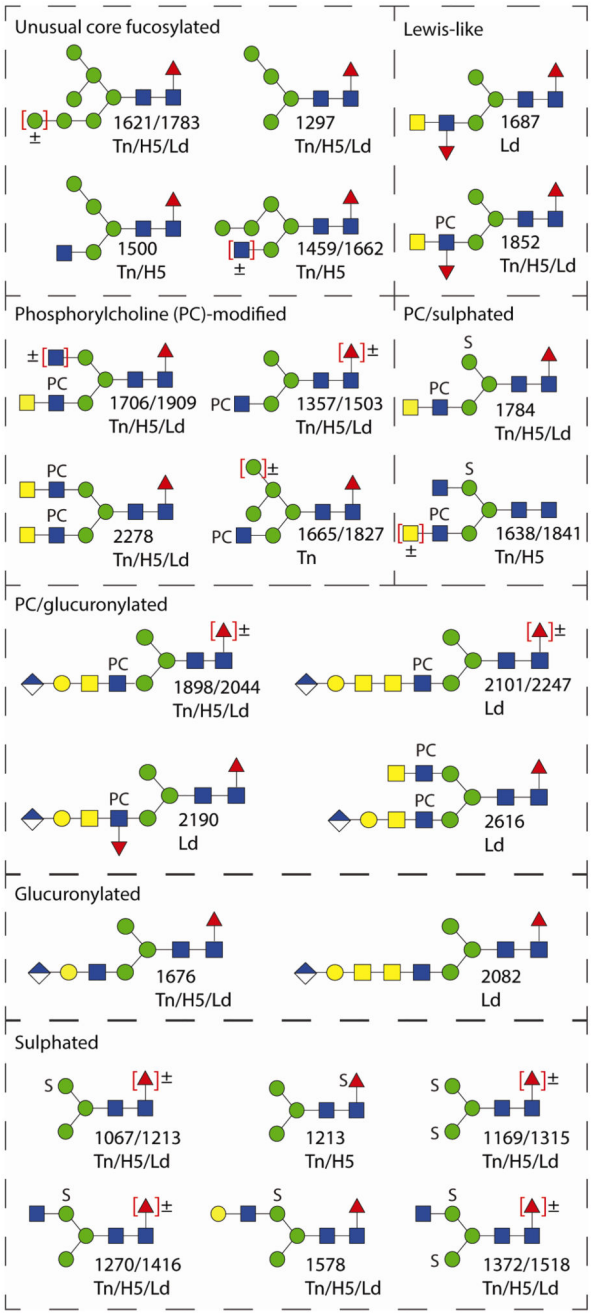

Figure 10. Summary of unusual N-glycans in lepidopteran species.

Examples of glycans with core α1,6-fucose, Lewis-like antennae or phosphorylcholine-modified, sulphated and glucuronylated antennae from *T. ni* larvae (Tn), High Five cells (H5) and *L. dispar* larvae (Ld) are shown together with the relevant positive or negative mode *m/z* value.