



Published in final edited form as:

J Proteomics. 2015 August 3; 126: 172–188. doi:10.1016/j.jprot.2015.05.030.

TARGETTED RELEASE AND FRACTIONATION REVEAL GLUCURONYLATED AND SULPHATED N- AND O-GLYCANS IN LARVAE OF DIPTERAN INSECTS

Simone Kurz¹, Kazuhiro Aoki², Chunsheng Jin³, Niclas G. Karlsson³, Michael Tiemeyer²,
Iain B.H. Wilson¹, and Katharina Paschinger¹

¹ Department für Chemie, Universität für Bodenkultur, 1190 Wien, Austria

² Complex Carbohydrate Research Centre, University of Georgia, Athens, Georgia 30602, USA

³ Department of Medical Biochemistry, University of Gothenburg, SE-405 30 Göteborg, Sweden

Abstract

Mosquitoes are important vectors of parasitic and viral diseases with *Anopheles gambiae* transmitting malaria and *Aedes aegypti* spreading yellow and Dengue fevers. Using two different approaches (solid-phase extraction and reversed-phase or hydrophilic interaction HPLC fractionation followed by MALDI-TOF MS or permethylation followed by NSI-MS), we examined the N-glycans of both *A. gambiae* and *A. aegypti* larvae and demonstrate the presence of a range of paucimannosidic glycans as well as bi- and tri-antennary glycans, some of which are modified with fucose or with sulphate or glucuronic acid residues; the latter anionic modifications were also found on N-glycans of larvae from another dipteran species (*Drosophila melanogaster*). The sulphate groups are attached primarily to core α -mannose residues (especially the α 1,6-linked mannose), whereas the glucuronic acid residues are linked to non-reducing β 1,3-galactose. Also, O-glycans were found to possess glucuronic acid and sulphate as well as phosphoethanolamine modifications. The presence of sulphated and glucuronylated N-glycans is a novel feature in dipteran glycomes; these structures have the potential to act as additional anionic glycan ligands involved in parasite interactions with the vector host.

Biological significance statement—Glycans of various types are relevant in many vector-parasite systems and it was previously shown that proteoglycans play roles in malaria transmission, whereas some fungal lectins are toxic towards mosquito larvae. The presence of sulphated and glucuronylated N- and O-glycans in mosquito larvae show that there are anionic glycans other than glycosaminoglycans; these could be ligands in interactions between pathogens (protozoan parasites and viruses) and the tissues of adult mosquitoes.

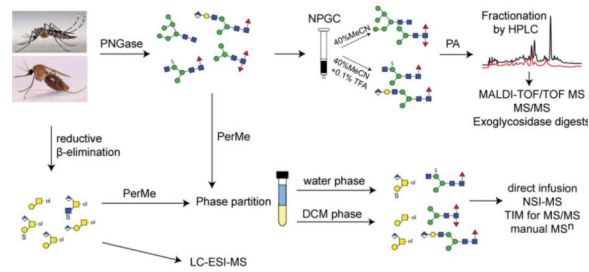
Abstract

* Corresponding author: iain.wilson@boku.ac.at.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest:

The authors declare no conflict of interest.



Keywords

glycomics; insects; mass spectrometric; glycans; oligosaccharides; HPLC

Mosquitoes are important vectors for a range of human and animal pathogens, including the malaria parasite *Plasmodium* [1] and the yellow fever, Dengue, Chikungunya and West Nile viruses [2-5]. As glycans coat both pathogen and host cell surfaces, many interactions between host and pathogen are glycan-mediated. In the case of malaria, attachment of the parasites to the gut and salivary glands of the mosquito vector are dependent on chondroitin and heparin sulphates respectively [6, 7], whereas heparin sulphate is also a 'receptor' for *Plasmodium* sporozoites in the mammalian liver [8]. However, mosquitoes are expected to have a range of cell surface glycoconjugates other than proteoglycans and these may also play roles in parasite transmission.

The basic repertoire of N- and O-glycans in insect species is well known with recent glycomic studies being centered on either the fruitfly *Drosophila melanogaster* or on the recombinant proteins expressed by insect cell lines [9, 10]. In general, so-called paucimannosidic N-glycans (with three or fewer mannose residues) with and without fucosylation as well as standard oligomannosidic N-glycans dominate, whereas the O-glycomes primarily consist of mono- and disaccharides. Nevertheless, studies over the past ten years have shown that low levels of more complex N- and O-glycans can be found in insect tissues. Triantennary and sialylated N-glycans have been found [11], whereas glucuronylation is a feature of both O-linked 'mucin-type' and 'O-fucose' oligosaccharides [12, 13]. Thus, it is apparent that procedures for analysing glycomes of insects should be 'open' to the possibility of finding more complicated neutral and anionic glycans.

We have developed different schemes suitable for analysing anionic glycans from either protist or insect sources [14, 15]. In the former case, solid-phase extraction to separate neutral from anionic glycans was found to be important to detect sulphated N-glycans from *Dictyostelium* by off-line LC-MALDI-TOF MS in their native state [14], whereas the use of phase extraction after permethylation enabled detection by NSI-MS of sialylated and glucuronylated glycans from *Drosophila* as well as sulphated O-glycans from mammalian mucins [15, 16]. In the present study, both approaches were employed in the analysis of the N- and O-glycans from the larvae of two mosquito species, the malaria vector *Anopheles gambiae* and the viral vector *Aedes aegypti*; thereby, we reveal a hitherto unknown complexity of the glycomes of both species.

EXPERIMENTAL PROCEDURES

Biological material

Anopheles gambiae (Keele) were reared and maintained in an environmental chamber at 27 °C with a relative humidity 80% and a 12h light/dark cycle. Eggs were hatched in distilled water, transferred to plastic pans and larvae were fed on a slurry of ground cat food (Purina cat chow). After 5 to 6 days of feeding, L3-L4 instar larvae were collected by centrifugation and stored at -80°C. *Aedes aegypti* (Rockefeller) larvae were reared, dependent on the source laboratory either using a slurry of Vitakraft Premium VITA 'Flockenfutter für alle Zierfische' (22 102) or as for *A. gambiae*. *Drosophila* larvae were prepared by standard cultivation techniques. In order to determine whether any of the glycans detected in the mosquito larval samples originated from the food source, N-glycans were prepared from the cat food and the fish food and were analysed by NSI-MS and off-line MALDI-TOF MS respectively; thereby fragmentation patterns and elution times could be compared in order to check for any potential contaminating glycans in the larval samples.

N-glycan purification and MALDI-TOF/TOF MS analysis

Frozen larvae (approximately 1.5-2.0 g wet weight) were boiled in deionised water for 10 minutes before grinding in a mortar. Thereafter, the slurry was made up to a total volume of 10 mL prior to the addition of formic acid (up to 5% (v/v)) and 1 mg porcine pepsin (Sigma-Aldrich). After proteolysis, cation exchange and gel filtration chromatography were performed as previously described [17] and the N-glycans released using PNGase F (recombinant; Roche) as well as by subsequent PNGase A (from almonds; Roche) digestion of remaining glycopeptides. De-N-glycosylated glycopeptides were gel filtrated and subject to β -elimination prior to LC-MS of native O-glycans (see below). The N-glycans were then subject to nonporous graphitized carbon (NPGC) chromatography for separation into neutral and anionic fractions [14]. As required, neutral N-glycans were further purified by a second solid-phase extraction using a Lichroprep RP18 cartridge column (25-40 μ m; Merck). Pyridylation was performed [17, 18] and the fluorescently-labelled N-glycans were fractionated by HPLC using either an Ascentis® Express RP-Amide column (150 \times 4.6 mm, 2.7 μ m; Supelco) with a gradient of 0.3% methanol per minute at a flow rate of 0.8 mL/min, or an IonPac AS11 column (HIAX, Dionex) as described previously [14, 19]. The control pyridylaminated triantennary N-glycan from foetal calf serum (the major component being fetuin) was purified by RP-HPLC as part of a previous study [17]. Collected fractions were lyophilized and reconstituted in water and analysed by MALDI-TOF/TOF MS in positive and negative ion modes (Bruker Autoflex Speed or Bruker UltrafleXtreme) with 6-aza-2-thiothymine (ATT) as matrix. The approximately 5500 MS and MS/MS spectra generated in this study were initially processed using the manufacturer's software (Bruker Daltonics FlexAnalysis 3.3.80) and then manually interpreted. Theoretical masses were calculated using the software GlycoWorkbench 2.0. The qualitative/semi-quantitative estimation of glycan amounts (from '+++' through 'trace') is based on HPLC fluorescent peak intensities; in case of multiple glycans in an HPLC fraction, a sub-estimation on the basis of MALDI-TOF MS data was made.

Exoglycosidase treatments

Further analysis of whole N-glycome pools or of selected HPLC fractions (see Results) by MALDI-TOF MS was performed after treatment overnight with either β -galactosidase (either *Xanthomonas manihotis* β 1,3-galactosidase from NEB or recombinant *Aspergillus niger* lacA β 1,4-galactosidase prepared in-house [20]), α -fucosidases (bovine kidney from Sigma-Aldrich or almond α 1,3-specific from Prozyme), α -mannosidases (jack bean from Sigma, *Xanthomonas* α 1,2/3-specific from NEB or *Xanthomonas* α 1,6-specific from NEB), *E. coli* β -glucuronidase (the kind gift of Megazyme; ultrafiltrated to remove some impurities before use) or β -hexosaminidase (either jack bean β -hexosaminidase from Sigma, β 1,3/4-specific *Streptomyces* chitinase from NEB or recombinant *Apis mellifera* FDL β 1,2-N-acetylglucosaminidase; prepared in-house [21]) in 25 mM ammonium acetate, pH 5.0 (pH 7.0 in the case of β -glucuronidase), at 37 °C overnight (three hours at 30 °C in the case of FDL, which under the conditions removes specifically the N-acetylglucosamine linked to the core α 1,3-mannose). Desulphation was performed by solvolysis as described below; a previously-studied sulphated N-glycan from *Dictyostelium* [14] was used as a positive control. Generally, chemically or enzymatically treated glycans were analysed by MALDI-TOF MS without further purification.

Glycan permethylation and NSI-MS analysis

Permethylated N- and O-glycans from insect larvae were prepared as described [15]. Frozen larvae were homogenised in ice-cold 50% (v/v) aqueous methanol and delipidated with chloroform/methanol/water (4:8:3, v/v/v). Insoluble proteins were precipitated by centrifugation and the resulting pellet was washed with acetone to produce a fine protein powder, a portion of which was subject to trypsinisation. Tryptic peptides were purified on C18 cartridges (Baker C18) and digested with either PNGase F (Prozyme) or PNGase A (Calbiochem) prior to another round of C18 chromatography to separate released N-glycans from residual glycopeptides. Separately, 2-3 mg of protein powder was subject to reductive β -elimination and released oligosaccharide alditols were purified on a C18 cartridge column as for the N-glycans. Permethylation of enzymatically- or chemically-released glycans was performed using iodomethane in a suspension of sodium hydroxide in dimethyl sulphoxide [22]. The permethylated glycans were then treated with water/dichloromethane (DCM; 1:1) to separate non-sulphated and sulphated glycans by phase partition; the lower organic phase contained non-sulphated permethylated glycans, whereas the upper aqueous phase contained sulphated glycans and both pools were subject to solid-phase extraction on C18 cartridges [16].

Solvolysis and re-permethylation of glycans with deuterated methyl iodide (CD₃I)

Sulphated glycans were dissolved in 100 μ L of 50 mM methanolic HCl (Supelco) and hydrolyzed for 4 hours at room temperature [16, 23, 24]. After drying under a gentle N₂ stream, resulting neutral glycans were re-permethylated with deuterated methyl iodide (CD₃I; Sigma-Aldrich) as described above. The lower organic (DCM) phase was extensively washed with water and dried under a N₂ stream prior to analysis with a nanospray ionization mass spectrometer (NSI-MSⁿ; Thermo Fisher Scientific) in positive ion mode [12].

Glycan analysis by Nanospray Ionization Mass Spectrometry (NSI-MSⁿ)

For MS analysis of non-sulphated glycans in positive ion mode, permethylated glycans were dissolved in 50 μ L of 1 mM sodium hydroxide in 50% (v/v) aqueous methanol for subsequent infusion. For MS analysis of sulphated glycans, permethylated glycans were reconstituted in 50 μ L of methanol/2-propanol/1-propanol/13 mM aqueous ammonium acetate (16:3:3:2 by volume) for infusion and analysed in negative ion mode. Samples were infused directly into a linear ion trap mass spectrometer (LTQ-Orbitrap Discovery; Thermo Fisher Scientific) using a nanoelectrospray source at a syringe flow rate of 0.40 to 0.60 μ L/min and a capillary temperature set to 210°C. Automated acquisition of MS/MS fragmentation (at 35-50% collision energy) was obtained using the total ion mapping (TIM) functionality of the Xcalibur instrument control software (version 2.0, Thermo Scientific). As described by Aoki and Tiemeyer [15], in TIM analysis the m/z range from 200 to 2000 was scanned in successive 2.8 mass unit windows with a window-to-window overlap of 0.8 mass units. For subsequent manual MS/MS and MSⁿ analyses by collision-induced dissociation (CID), normalized collision energy of 35% was applied. As internal calibration, 10 pmol of deuteride-labelled permethylated maltooligosaccharides (Dp3 and Dp4) were spiked into the resuspended samples prior to analysis.

LC-MS of O-glycans

Residual glycopeptides bound to Dowex 50 after the PNGase A digestion (see above) were subject to gel filtration prior to reductive β -elimination and LC-MS. Released O-glycans were cleaned up as previously described [25] prior to analysis by LC-MS and LC-MSⁿ using a 10 cm \times 250 μ m I.D. column, prepared in-house, containing 5 μ m porous graphitized carbon (PGC) particles (Thermo Scientific). O-glycans were eluted using a linear gradient from 0-40% acetonitrile in 10 mM ammonium bicarbonate over 40 min at a flow rate of 10 μ L/min. The eluted O-glycans were detected using a LTQ XL ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of -33.0 V and capillary temperature of 300°C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analysed. Specified ions were isolated for MSⁿ fragmentation by CID with the collision energy set to 30%. The data were processed using Xcalibur software (version 2.2, Thermo Scientific).

RESULTS

The neutral N-glycomes of mosquito larvae

We examined the N-glycomes of *Aedes aegypti* and *Anopheles gambiae* by two different approaches (off-line MALDI-TOF MS of pyridylamino-labelled glycans or NSI-MS of permethylated glycans); in both cases we used PNGase F and then either released the residual N-glycans with PNGase A, which can also release core α 1,3-fucosylated glycans, or performed a parallel digest with PNGase A alone. For the off-line MS approach, the glycans were fractionated by solid phase extraction into pools of neutral and anionic PNGase F-released and neutral PNGase A-released glycans; these three pools were separately pyridylaminated and analysed by MALDI-TOF MS in conjunction with HPLC. In the case of the permethylated PNGase F- or PNGase A-released N-glycans, the pools of organic and aqueous phases (i.e., neutral and sulphated pools) were analysed by NSI-MS. Using either

approach, the basic repertoire of neutral N-glycans released with PNGase F, Hex₃₋₉HexNAc₂ as well as Hex₂₋₃HexNAc₂₋₅Fuc₀₋₁, was very similar in both mosquito species and is familiar from other glycomic studies on insects [9].

The most dominant glycan was Hex₃HexNAc₂Fuc₁ followed by Hex₃HexNAc₂ and Hex₉HexNAc₂ (see **Figure 1A** for the MALDI-TOF MS of *Aedes* N-glycans and **Supplementary Figure 1** for the NSI-MS spectra). It is to be noted that the higher oligomannosidic glycans tend to be underrepresented in the MALDI-TOF spectra as compared to the NSI-MS and HPLC profiles; this is probably due to less efficient ionisation of these structures, as we have no evidence for in-source fragmentation when analysing oligomannosidic glycans in individual HPLC fractions. Therefore, the qualitative estimation of N-glycan occurrence (**Figure 2**) is based primarily on HPLC data. Difucosylated N-glycans with compositions of Hex₂₋₃HexNAc₂₋₃Fuc₂ were found only in the PNGase A-released pools (**Figure 1B**). Such core difucosylated N-glycans are known to be the source of cross-reactivity towards anti-horseradish peroxidase [26]; MS/MS was performed to verify their composition, with an *m/z* 592 ion (HexNAc₁Fuc₂-PA) being a diagnostic fragment for such structures [27], but no data is presented here as they are well characterised from a number of insect sources [28]. An RP-amide HPLC column, calibrated in terms of glucose units, was used to fractionate the pyridylaminated neutral PNGase F- and PNGase A-released N-glycan pools from both species and MALDI-TOF MS (with MS/MS) was performed on each fraction (**Figure 3**). This resulted in detection of a higher number of N-glycans and aided also definition of their isomeric status; basic trends in terms of isomers have been well described for standard RP-HPLC columns over the past thirty years [29, 30] and these also apply (although not in terms of exact glucose units) for the RP-amide column. Examples of isomeric separations include distinguishing respectively α 1,3- from α 1,6-fucosylation (Hex₃HexNAc₂Fuc₁ at 5.2 and 9.6 g.u.; *m/z* 1135) or the presence of a non-reducing β 1,2-linked GlcNAc on either the α 1,3- or α 1,6-antennae (Hex₃HexNAc₃ at 6.7 and 9.0 g.u. [*m/z* 1192], Hex₃HexNAc₃Fuc₁ at 9.0 and 12.5 g.u. [*m/z* 1338] and Hex₃HexNAc₃Fuc₂ at 6.5 and 7.9 g.u. [*m/z* 1484]; see **Figure 2**).

As appropriate diagnostic exoglycosidase digestions were performed: for instance, two different isomers of Man₂GlcNAc₂Fuc₁ (*m/z* 973) and three different isomers of Man₃GlcNAc₃Fuc₁ (*m/z* 1338) could be resolved on the basis of their fragmentation pattern and sensitivity towards α 1,2/3- and α 1,6-specific mannosidases [31] and the FDL β 1,2-specific hexosaminidase [21]. Thus, for Man₂GlcNAc₂Fuc₁, the isomers of 8.5 g.u. and 9.6 g.u. have respectively α 1,3- and α 1,6-mannose residues (**Supplementary Figure 2 A-E**), while for Man₃GlcNAc₃Fuc₁, the single non-reducing terminal N-acetylglucosamine is β 1,2-linked to either the α 1,3- or α 1,6-mannose or β 1,4-linked to the α 1,3-mannose (9.0, 11.4 and 12.5 g.u. isomers shown; **Supplementary Figure 2 G-O**). Core fucosylation of these glycans, demonstrated by the protonated fragment ions of *m/z* 446 (see, *e.g.*, **Supplementary Figure 2 P-R**), was concluded to be in α 1,6-linkage due to the sensitivity towards bovine α -fucosidase (**Supplementary Figure 2F**) and the late RP-HPLC elution. Another example of isomeric separation is for Hex₃HexNAc₄Fuc₁ eluting at either 10.8 or 14.8 g.u. (*m/z* 1541); the former is predicted to be the standard biantennary Man₃GlcNAc₄Fuc₁ structure with two non-reducing GlcNAc residues (**Supplementary**

Figure 2W), while the latter is concluded to contain a LacdiNAc on the α 1,6-antenna as judged by the m/z 406 fragment and the α 1,3-mannosidase sensitivity of this structure (**Supplementary Figure 2 S-V**).

In general, the HPLC chromatograms of the two mosquito species are very similar; however, the *A. gambiae* chromatograms appeared somewhat simpler as compared to those for *A. aegypti*. Indeed, a close examination of the initial overall MALDI-TOF spectra indicated variations in the presence of glycans with more than four HexNAc residues. Especially an N-glycan with Hex₃HexNAc₅Fuc₁ (m/z 1744; [M+H]⁺) was found in *A. aegypti* (13.5 g.u.); the corresponding peak in *A. gambiae* was barely detected. The structure contains a core α 1,6-fucose as indicated by the m/z 446 fragment, which was missing after fucosidase digestion (**Figure 4 F and G**); the major question was whether this glycan was triantennary and, if so, which arm carried two non-reducing terminal *N*-acetylglucosamine residues. Therefore, the defucosylated glycan was reinjected onto the RP-amide column (**Figure 4A**); the digestion product co-eluted with an asialoagalacto-triantennary N-glycan isolated after PNGase F digestion of foetal calf serum proteins (primarily fetuin, whose triantennary glycan structures are well known [32]). Furthermore, no m/z 407 fragment (LacdiNAc) was present, while the m/z 1176 and 1379 fragments (m/z 1030 and 1233 for the defucosylated form) are compatible with two GlcNAc residues on the α 1,3-antenna (**Figure 4 F and G**). Thus, we conclude that the m/z 1744 glycan from *A. aegypti* contains two β 1,2- and one β 1,4-linked terminal *N*-acetylglucosamine residues, indicating that it is a product of the combined action of GlcNAc-TI, II and IV [33]; this is compatible by the presence of the relevant glycosyltransferase homologues in the *A. aegypti* and *A. gambiae* genomes (NCBI entries: XM_001661326 and XM_315359, XM_001651678 and XM_003436754 and XM_001664051 and XM_313212). A full list of neutral glycans is given in **Figure 2**.

Sulphated N-glycans of insect larvae

Due to the detection of sialylated N-glycans as minor components in *D. melanogaster* [11], our glycomic strategy reflected that we intended to also determine whether sialylated or other anionic N-glycans were present in mosquito larvae. Therefore, we used non-porous graphitised carbon for solid phase extraction of the underivatised oligosaccharides to enrich the anionic N-glycans (**Figure 1 C and D**), as we previously have done for *Dictyostelium* [14]. Unlike the 'neutral' fraction, the overall profiles for the anionic pools in the negative ion mode were different between the two species; the *A. gambiae* pool was simpler and dominated by a glycan of m/z 1416, whereas the major anionic glycan in *A. aegypti* was of m/z 1822. Indeed, the PNGase F-released glycans eluted from the carbon column under acidic conditions were enriched, as compared to the neutral pool, in structures predicted to carry modifications of either 80 or 176 Da. In the positive ion mode, the glycans with the 176 Da modification were also apparent, whereas those with the 80 Da modification were detected either as [M+H-80]⁺ or [M+Na]⁺ species. Thus, we hypothesised (based on our experience with *Dictyostelium* N-glycans) that the 80 Da modification is sulphate rather than phosphate, whereas the presence of the 176 Da modification primarily in the anionic fraction suggested that it was glucuronic acid, rather than non-anionic methylhexose. Similarly as with the neutral pools, the primary separation technique employed for anionic glycans was

HPLC using the RP-amide column, but the *A. gambiae* N-glycans were also fractionated on the AS11 HPLC column (**Figure 5**).

Digestion of the whole anionic pool of *A. aegypti* with bovine fucosidase and jack bean hexosaminidase resulted in a single major glycan with m/z 1067 ($\text{Hex}_3\text{HexNAc}_2\text{S}$; $[\text{M-H}]^-$). A subsequent specific α 1,2/3-mannosidase digestion indicated that the putative sulphate residue on the major anionic glycans was on the α 1,6-mannose, a result compatible with the negative ion mode MS/MS fragments of m/z 241, 403 and 565 ($\text{Hex}_{1-3}\text{S}_1$) (**Supplementary Figure 3**); we have also obtained data indicating sulphation of the α 1,6-mannose of N-glycans from hymenopteran and lepidopteran larvae (**manuscript in preparation**). There were also sulphated glycans in the fish and cat food controls which were the food sources for the larvae (*e.g.*, m/z 1416 in fish food); however, the amounts were low and the RP-HPLC elution positions were different, particularly when considering the dominance of m/z 1619 and 1822 in the overall mosquito anionic glycan spectra as compared to the food source; thus, these sulphated glycans were concluded to be indeed of insect origin, in contrast to traces of sialylated glycans which co-elute with those in the fish food control (**Supplementary Figure 4**).

As mentioned above, the most abundant anionic N-glycan in *A. aegypti* is $\text{Hex}_3\text{HexNAc}_5\text{Fuc}_1\text{S}$ (10.4 g.u.; m/z 1822 as $[\text{M-H}]^-$ in negative ion mode or 1846 as a minor $[\text{M}+\text{Na}]^+$ species in positive ion mode; **Figure 4 B and C**). Specific hexosaminidase treatments (*Apis* β 1,2-specific FDL and *Streptomyces* β 1,3/4-chitinase) suggested the presence of non-reducing terminal β 1,2- and β 1,4-*N*-acetylglucosamine residues on the α 1,3-arm (**Figure 4 D and E**). This glycan was also subject to solvolysis with methanolic HCl and, after subsequent fucosidase treatment, reinjected onto RP-HPLC. About 30% of the glycan had shifted its retention time to around 9.7 g.u. and, as for the defucosylated form of $\text{Hex}_3\text{HexNAc}_5\text{Fuc}_1$, co-eluted with the asialoagalacto-glycan derived from foetal calf serum (**Figure 4A**). The remaining 70% of the glycan retaining its sulphate after solvolysis eluted at 7.8 g.u., which suggests that sulphation reduces retention time of this structure on the RP-amide column by about 2 g.u.. The MS/MS fragmentation patterns were compatible with sulphation of mannose (**Figure 4 H and I**).

The second most dominant glycan in the anionic pool of *A. aegypti* had the composition $\text{Hex}_3\text{HexNAc}_4\text{Fuc}_1\text{S}$ (m/z 1619; $[\text{M-H}]^-$); analysis of the RP-HPLC fractions indicated that there were two isomers eluting at 7.9 and 8.7 g.u. which had very similar MS/MS fragmentation patterns (**Figure 6 A and B**); however, these could be distinguished due to their different sensitivity to FDL hexosaminidase. Thus, these glycans differ in the linkage (β 1,2- or β 1,4) of the *N*-acetylglucosamine residue on the α 1,3-arm (**Figure 6C-G**). The 7.9 g.u. glycan was also incubated with hydrofluoric acid, but was unaltered by this treatment, which is a further indication that it is sulphated rather than phosphorylated (**Supplementary Figure 5A-D**). The third most abundant anionic glycan in *A. aegypti* had the composition $\text{Hex}_3\text{HexNAc}_3\text{Fuc}_1\text{S}$ (m/z 1416; $[\text{M-H}]^-$); this glycan had the same retention properties (7.7 g.u.) as the major anionic one from *A. gambiae* and was susceptible to α 1,2/3-mannosidase digestion, which verifies that both the sulphate and non-reducing terminal *N*-acetylglucosamine are associated with the α 1,6-arm (**Figure 6 H and I**). In addition, we also detected disulphated N-glycans ($\text{Hex}_3\text{HexNAc}_{3-5}\text{S}_2$) in either early RP-amide and middle

HIAX fractions. Based on their hexosaminidase sensitivity, mannosidase resistance and fragmentation in negative ion mode (**Figure 6 M-O**; especially the fragment of m/z 667, which corresponds to sodiated Hex_3S_2), one sulphate residue is predicted to be on each of the core α -mannose residues. Although $\text{Man}_3\text{GlcNAc}_2\text{Fuc}_{0-1}$ are dominant structures in the neutral pool, a sulphated version of $\text{Man}_3\text{GlcNAc}_2$ was not observed, whereas $\text{Man}_3\text{GlcNAc}_2\text{Fuc}_1\text{S}_1$ was barely detected. In addition to these sulphated glycans with non-reducing terminal *N*-acetylglucosamine and one core fucose, there were other minor forms with additional hexose, *N*-acetylhexosamine and fucose residues. For instance, a glycan present in a HIAX fraction with the composition $\text{Hex}_4\text{HexNAc}_3\text{Fuc}_1\text{S}_1$ (m/z 1578; $[\text{M-H}]^-$) was susceptible to both β 1,3-galactosidase and α 1,2/3-mannosidase digestion (**Supplementary Figure 5E-G**), whereas $\text{Hex}_3\text{HexNAc}_4\text{Fuc}_2\text{S}_1$ (m/z 1765; $[\text{M-H}]^-$) detected in an another HIAX fraction possessed MS/MS fragments of m/z 1320 ($\text{Hex}_3\text{HexNAc}_3\text{Fuc}_1\text{S}_1$), 1117 ($\text{Hex}_3\text{HexNAc}_2\text{Fuc}_1\text{S}_1$) and 971 ($\text{Hex}_3\text{HexNAc}_2\text{S}_1$). This latter structure was susceptible to almond α 1,3-fucosidase and subsequent α 1,2/3-mannosidase (**Figure 6J-L**); thus, it was concluded to have a fucosylated LacdiNAc motif in addition to a sulphated mannose. The presence of LacdiNAc on the upper arm was also indicated for $\text{Hex}_3\text{HexNAc}_4\text{Fuc}_1\text{S}_1$ (m/z 1619; $[\text{M-H}]^-$) as judged by its α 1,2/3-mannosidase sensitivity (**Supplementary Figure 5G**). A full list of sulphated glycans is given in **Figure 7**.

We also searched for sulphated N-glycans by NSI-MS. As sulphated glycans are not found in the organic phase upon permethylation, we applied a new clean-up procedure to examine the glycans present in the aqueous phase [16]. The MS^2 data verified that sulphated N-glycans can be found not only in mosquito larvae, but also in *D. melanogaster* (**Figure 8 A-D**). As a further proof that the 80 Da modification on the permethylated glycans was sulphate, the same desulphation procedure with methanolic HCl as used on the aforementioned pyridylaminated $\text{Hex}_3\text{HexNAc}_5\text{Fuc}_1\text{S}$ was employed. However, following solvolysis, the desulphated glycan pool was repermethylated using deuterated iodomethane thereby placing an isotopically heavy methyl group in place of the sulphate. Thus new deuterated glycans were detected in the organic phase which are 2.2 a.m.u. different from corresponding neutral N-glycans in the original organic phase (**Figure 8E**). NSI- MS^5 data demonstrated that the deuteromethyl group was primarily associated with hexose (**Figure 8G**), a result compatible with the MALDI-TOF MS/MS data discussed above. A minor degree of sulphation on fucose was also detected as shown by MS/MS fragmentation and by the fucosidase resistance of glycans which on the RP column elute slightly before the neutral form of the same glycan, but somewhat later than the isobaric form with sulphated hexose (**Supplementary Figure 6 A-F**). This conclusion was also confirmed by NSI-MS, after desulphation and perdeuteromethylation, which showed association of the deuteromethyl group with fucose (fragment of m/z 274; **Figure 8H**).

Glucuronylated N-glycans of mosquito larvae

As mentioned above, some glycans in the anionic pools were hypothesised to contain glucuronic acid as the component conferring negative charge. Of these glycans, the most obvious from the overall negative ion mode spectra of the anionic pools were species of m/z 1877 and 2080 ($[\text{M-H}]^-$; **Figure 1 C and D**), which were also detected in positive ion mode

as m/z 1879 and 2082 ($[M+H]^+$; Hex₄HexNAc₄₋₅Fuc₁HexA₁). There appeared to be at least two isomers of Hex₄HexNAc₄Fuc₁HexA₁ (7.2 and 8.3 g.u.). These differed in their fragmentation patterns (**Supplementary Figure 7 A-B**), particularly in terms of the presence of an m/z 745 fragment (m/z 743 in negative ion mode; Hex₁HexNAc₂HexA₁); on the other hand, the m/z 542 fragment (m/z 540 in negative ion mode; Hex₁HexNAc₁HexA₁) was a feature of all putatively glucuronylated glycans. The presence of terminal glucuronic acid on the glycan in the 8.3 g.u. fraction was demonstrated by its sensitivity to β -glucuronidase (the m/z 745 fragment was lost and replaced by one of m/z 569; the product was in turn sensitive to a specific β 1,3-galactosidase (resulting in an m/z 407 fragment upon MS/MS), but resistant to β 1,4-galactosidase (**Figure 9 A-D**). Thus, we conclude that the antennal modification GlcA β 1,3Gal β 1,3GalNAc β 1,4GlcNAc may be present on the 8.3 g.u. form of the m/z 1879 glycan. This antenna can also be presumed for at least one isomer of Hex₄HexNAc₅Fuc₁HexA₁ (m/z 2082; **Supplementary Figure 7C**). Furthermore, there were a number of diglucuronylated N-glycans with compositions of Hex₅HexNAc₄₋₆Fuc₁GlcA₂ (m/z 2218, 2421 and 2624; $[M+H]^+$; **Supplementary Figure 7 D-F**). Indeed, the presence of two glucuronic acid moieties on such glycans could be demonstrated by exoglycosidase digestion with compatible changes in the MS/MS fragmentation (**Figure 9 E-J**). Thereby, two types of glucuronylated antennae (Hex₁HexNAc₁HexA₁ and Hex₁HexNAc₂HexA₁) were apparent for these diglucuronylated glycans; also NSI-MSⁿ data support these structural propositions (**Supplementary Figure 7 G-I**).

N-glycans modified with both sulphate and glucuronate residues

MALDI-TOF MS analysis of both RP-amide and HIAX fractions indicated that a range of N-glycans contained not only sulphate or glucuronic acid, but indeed possessed both modifications and sulphated forms of the various glucuronylated glycans described above were detected. Exoglycosidase digestions and MS/MS of simple examples of such structures were consistent with the presence of an extended upper arm with 'peripheral' glucuronic acid and 'internal' sulphate (Hex₄HexNAc₃₋₄Fuc₁HexA₁S₁; **Figure 10**); whereas one mannose was removed with α 1,2/3-specific mannosidase, the fragment of m/z 768 in negative ion mode is compatible with the linkage of sulphate to the trimannosyl core region.

In general, glycans predicted to have both anionic modifications were detected as $[M+H-80]^+$ in the positive ion mode and the MS/MS were similar to those for the glycans modified with glucuronic acid alone; both 'long' and 'short' antennal modifications (Hex₁HexNAc₁₋₂HexA₁) can be predicted (**Supplementary Figure 8**). In many cases, fragmentation of the $[M-H]^-$ showed the loss of 80 a.m.u., but was not generally more informative, leaving ambiguity regarding the position of the sulphate residue; indeed, in some cases, a fragment of m/z 458 in the negative ion mode would be compatible with sulphation of the terminal HexNAc₁HexA₁ motif on some glycans from *Anopheles* (**Supplementary Figure 8F and J**). In addition, the presence of a possibly triantennary structure with glucuronate and sulphate modifications could be predicted from MS of early-eluting RP-amide fractions or late HIAX fractions; it appeared that these glycans were, under the settings enabling their detection, rather labile and the major evidence for their

structure is from in-source fragmentation (**Supplementary Figure 9**). **Figure 7** summarises all glycans proposed to contain glucuronic acid with and without additional sulphate.

The O-glycomes of mosquito larvae

In order to examine the O-glycome of the larvae of the two mosquito species, β -elimination was performed on either the trypsinised acetone protein powder or residual glycopeptides of the pepsinised samples, followed respectively by NSI-MS of permethylated or LC-MS of reduced O-glycans (the former in comparison to O-glycans from *D. melanogaster* larvae). The permethylated organic phase showed the presence of a set of O-glycans ranging from mono- to pentasaccharides, including many core 1 variants (**Figure 11 A-C**; summarised in **Figure 12**). The glucuronylated O-glycans clearly detected in *D. melanogaster* larvae and previously found in *D. melanogaster* embryos [12], are seemingly less pronounced in the mosquitoes; nevertheless, the compositions of glycans such as HexA₁₋₂Hex₁HexNAc₁-ol could be verified by MS², but in part may also be of a different isomeric status as compared to those from *D. melanogaster* (**Figure 11 D-G**). On the other hand, the mosquito samples also contained a series of O-glycans with one or two extra hexoses and one extra *N*-acetylhexosamine; one of the hexoses attached to the core 1 structure could be α 1,4-linked galactose, if the mosquito larvae have share some of the same O-glycomic potential as lepidopteran cells [34]. In addition, low levels of permethylated Pnt₁Hex₁-ol (*m/z* 450) and HexNAc₁HexA₁dHex₁-ol (*m/z* 722) were detected which would be consistent with insect-type modifications of EGF domains on proteins such as Notch (**Supplementary Figure 10**). In the aqueous phase, three sulphated O-glycans were detected to varying extents in the mosquito and fruitfly samples; the MS² data does not resolve the position of some of the sulphate residues, but indicated that sulphate can be linked to *N*-acetylhexosamine (**Figure 13**). Both glucuronylated and sulphated O-glycans were detected from two *Drosophila* cell lines (S2 and BG2c6; **Supplementary Figure 11**).

LC-MS of native β -elimination products yielded comparable data, but also revealed even more variants of glycans with a terminal reduced *N*-acetylhexosamine (**Figure 12**). Again glucuronylated and sulphated structures could be detected (**Figure 11 H and I**). Striking were also the extended chains containing HexNAc₁Hex₁HexA₁ repeats, with or without terminal sulphate or internal phosphoethanolamine (**Figure 11 H and I**; for MS/MS data refer to **Supplementary Figure 12**), which are similar to structures previously found in insect cell lines [13]. The absence of some of these glycans from the NSI-MS data may be due to loss during the extraction after permethylation.

DISCUSSION

The first studies of mosquito glycans were performed some thirty years ago using cell lines [35-37] and demonstrated the predominance of oligomannosidic N-glycans. Also, a number of lectin-based studies of mosquitoes have been published [38-40], whereas actual mass spectrometric data on glycoproteins from mosquitoes or mosquito cell lines is limited [41, 42]. Certainly, a comprehensive analysis of the total glycomic complexity using unbiased high-sensitivity mass spectrometric methods has not been previously reported. Here we have examined the N- and O-glycomes of mosquito larvae using state-of-the-art mass

spectrometric techniques; these data are supplemented by NSI-MS glycomic analyses of larvae of another dipteran, *D. melanogaster*. Consistent with previous reports, oligomannosidic N-glycans are well represented in our chromatograms and spectra, as are mono- and di-fucosylated paucimannosidic N-glycans, which are familiar from other insect species [9]. The matter of anionic glycans in insects, however, has proven more controversial: only relatively recent studies proved the presence of sialic acid and glucuronic acid on N- and O-glycans of *D. melanogaster* embryos [11, 12, 15]. In mosquito larvae we find both sulphated and glucuronylated N-glycans, but have no convincing evidence for the presence of sialylated structures in the larvae of the two species studied. This is not a problem of detection, but of the absence of evidence that the traces of sialylated structures found in larvae are not derived from the food sources: indeed, after removal of sialic acid, such glycans in either larvae or food have a β 1,4-galactosylated ‘backbone’ structures and not β 1,3-galactosylated ones as found for the glucuronylated glycans (**Supplementary Figure 4** as compared to **Figure 9**). Nevertheless, a CMP-sialic synthase cDNA from *Aedes* may be able to complement a sialylation defective mammalian cell line as judged by lectin staining [43].

Sulphated N-glycans are well known in mammalian species, e.g., sulphation of galactose of thyroglobulin, of GalNAc in the case of pituitary glycoprotein hormones and of GlcNAc on bovine lung glycoproteins [44-46]; however, sulphation is less often described in lower organisms. We recently reported N-glycan sulphation of galactose residues of the oyster [19]. Sulphation of mannose was also reported on *Dictyostelium* N-glycans [14, 47] and the trypanosomatid protein cruzipain [48], as well as specifically on the α 1,6-mannose in the case of a lobster haemocyanin [49]. Thereby, the sulphation of mannose residues discovered in the current study is not without precedent. N-glycan sulphation has previously not been found in insects, but a single report suggests the presence of sulphate at an undefined position on the N-glycans of one peptide from another arthropod, the woodlouse [50]. Certainly, there is a wealth of sulphotransferases in the genomes of insects as judged by database searching (e.g., approximately 60 in *A. gambiae*), including a number of putative carbohydrate sulphotransferases; it remains to be discovered which of these may modify N-glycans.

Glucuronic acid is a component of chondroitin and heparan sulphates as well as part of the HNK-1 epitope (SO₃-GlcA-Gal) found on N- and O-glycans in mammalian neural tissue [51, 52]. In insects, glucuronic acid is present in glycolipids [53], where the GlcA-Gal-GalNAc motif constitutes the CAF-1 epitope which was recognised by a monoclonal antibody raised against blowfly glycolipids [54]. CAF-1 reactivity was also detected in the *Drosophila* brain [54]. In mosquito larvae we have found both GlcAGal-GlcNAc and GlcA-Gal-GalNAc-GlcNAc motifs as potential protein-bound CAF-1 epitopes; unfortunately the antibody is apparently no longer available for testing. Furthermore, the reported cross-reactivity of insect glycoproteins to L2/HNK-1 antibodies may not require the presence of sulphate [55] and thus may merely indicate terminal glucuronylation. Low levels of N-glycan glucuronylation have also been observed in *Drosophila* embryos by NSI-MS [15] and this result is now extended to larvae of this species as well as to mosquitoes. The minor abundance of glucuronic acid in the N-glycomes examined here is compatible with a high

level of protein and tissue specificity in expression of this modification. It is to be presumed that the two ‘broad-specificity’ glucuronyltransferases previously characterised from *Drosophila*, which have orthologues in *A. aegypti* and *A. gambiae*, are responsible for the expression of glucuronylated N-glycans [56].

The glucuronylated and sulphated N-glycans of mosquito larvae are often complex structures with up to three antennae and also containing putative LacdiNAc or LacNAc motifs. The Gal β 1,3GalNAc β 1,4GlcNAc motif, which appears to constitute the antennae for some of the glucuronylated glycans described here, has been previously found in unglucuronylated form as a ‘T-antigen’ on biantennary N-glycans from honeybee royal jelly [57], whereas GalNAc β 1,4GlcNAc is found in fucosylated form on honeybee venom glycoproteins [58]. Relevant β 1,3-galactosyltransferase and β 1,4-*N*-acetylgalactosaminyltransferase enzymes have been described from insect sources [59-62]. The occurrence of terminal galactose on N-glycans and its assumed presence on O-glycans (*e.g.*, the core 1 type) is also of interest as *Aedes* larvae are sensitive to a number of fungal lectins which bind galactose, *N*-acetylgalactosamine or fucose residues [63]. However, although CGL2 (*Coprinopsis cinerea* galectin 2) is also toxic to mosquito larvae, its preferred ligand (β 1,4-galactose linked to core fucose) was not detected in our study; on the other hand, *Aleuria aurantia*, *Clitocybe nebularis* and *Xerocomus chrysenteron* lectins respectively bind fucose, GalNAc and Gal β 1,3GalNAc which are components of N- and O-glycans found in larvae of both mosquito species studied.

The O-glycomes of mosquito larvae also show interesting features: while glucuronylation of *D. melanogaster* embryo and cell line O-glycans [12, 64] and phosphoethanolamine modifications of wasp nest O-glycans have been previously shown [65], structures containing both these modifications were previously unknown. Furthermore, we detected a set of extended O-glycan structures with HexNAc₁Hex₁HexA₁ trisaccharide repeats as found previously on a recombinant protein expressed in insect cells [13]; these structures are partially reminiscent of glycosaminoglycan repeats, albeit with an additional hexose, and also share some of their general composition with insect glycolipids [53, 66]. Glycans with a similar repeat motif are unknown from vertebrates. Thus, although simple core 1 (Gal β 1,3GalNAc) may be the dominant O-glycan in insect species [10], the diversity of O-glycans is far higher than was considered possible until a few years ago.

The significance of these various N- and O-glycan modifications (sulphation and glucuronylation) for the biology of the mosquito is as yet unclear; the structural similarities they share with glycosaminoglycans suggest that, like some proteoglycans, they might facilitate binding of malaria parasites to mosquito tissues [6]. Mosquitoes are also vectors for a number of viruses and mosquito-derived Dengue virus enters mammalian host cells in a lectin-dependent manner [67]. Interestingly, a lectin specific for certain sulphated glycans has recently been described from *Anopheles* [68], but has not yet been tested with sulphated paucimannosidic glycans as potential ligands. Although not mosquito-borne, many viruses utilise host glycans for binding, including Lassa fever virus and herpes virus which respectively require O-mannose-type glycans and 3-sulphation for entry into mammalian cells [69, 70]. Thus N-glycans of the types described here may, if also expressed in the gut or salivary gland of adult mosquitoes, be relevant to host/vector/pathogen interactions either

by mediating binding to vector tissues or by conferring these modifications to viruses which pass through these species. Furthermore, the glycomic complexities being now discovered in insects may well have repercussions for the use of insect cells as systems for expressing recombinant glycoproteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported in part by FWF (Austrian Science Fund) grants P21946 and P25058 to KP; SK was supported by the FWF-funded doctoral programme “Biomolecular Technology of Proteins” [W1224], NGK was supported by Swedish Research Council (2013-5895 and 2010-5322) and CJ was supported by the Knut and Alice Wallenberg Foundation. The LTQ mass spectrometer was obtained by a grant from the Swedish Research Council (342-2004-4434). The authors acknowledge the support and access to instrumentation provided through a grant from the National Institute of General Medical Sciences (P41GM103490, MT co-PI) of the National Institutes of Health. The authors also thank Pie Müller (Swiss Tropical and Public Health Institute) and Markus Künzler (ETH Zürich) for *Aedes* material, Rhoel Dinglasan (Johns Hopkins Bloomberg School of Public Health, Baltimore) for both *Aedes* and *Anopheles* larvae and Ján Mucha (Slovak Academy of Sciences, Bratislava) for access to the UltrafleXtreme MALDI-TOF MS used to record some of the presented spectra.

REFERENCES

- White NJ, Pukrittayakamee S, Hien TT, Faiz MA, Mokuolu OA, Dondorp AM. Malaria. *Lancet*. 2014; 383:723–735. [PubMed: 23953767]
- Karunamoorthi, K. Yellow Fever Encephalitis: An Emerging and Resurging Global Public Health Threat in a Changing Environment.. In: Tkachev, S., editor. Encephalitis. InTech; Rijeka, Croatia: 2013. p. 207-230.
- Guzman MG, Harris E. Dengue. *Lancet*. 2015; 385:453–465. [PubMed: 25230594]
- Tsetsarkin KA, Chen R, Sherman MB, Weaver SC. Chikungunya virus: evolution and genetic determinants of emergence. *Current opinion in virology*. 2011; 1:310–317. [PubMed: 21966353]
- Suthar MS, Diamond MS, Gale M Jr. West Nile virus infection and immunity. *Nature reviews. Microbiology*. 2013; 11:115–128. [PubMed: 23321534]
- Dinglasan RR, Alaganan A, Ghosh AK, Saito A, van Kuppevelt TH, Jacobs-Lorena M. Plasmodium falciparum ookinetes require mosquito midgut chondroitin sulfate proteoglycans for cell invasion. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:15882–15887. [PubMed: 17873063]
- Armistead JS, Wilson IBH, van Kuppevelt TH, Dinglasan RR. A role for heparan sulfate proteoglycans in Plasmodium falciparum sporozoite invasion of anopheline mosquito salivary glands. *The Biochemical journal*. 2011; 438:475–483. [PubMed: 21663594]
- Pinzon-Ortiz C, Friedman J, Esko J, Sinnis P. The binding of the circumsporozoite protein to cell surface heparan sulfate proteoglycans is required for plasmodium sporozoite attachment to target cells. *J Biol Chem*. 2001; 276:26784–26791. [PubMed: 11352923]
- Rendi D, Wilson IBH, Paschinger K. The glycosylation capacity of insect cells. *Croatica Chimica Acta*. 2008; 8:7–21.
- ten Hagen KG, Zhang L, Tian E, Zhang Y. Glycobiology on the fly: developmental and mechanistic insights from Drosophila. *Glycobiology*. 2009; 19:102–111. [PubMed: 18824561]
- 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]
- Aoki K, Porterfield M, Lee SS, Dong B, Nguyen K, McGlamry KH, Tiemeyer M. The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and

- tissue-specific requirements for cell signaling. *J Biol Chem.* 2008; 283:30385–30400. [PubMed: 18725413]
13. 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]
 14. Hykollari A, Balog CI, Rendic D, Bräulke 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. *Journal of proteome research.* 2013; 12:1173–1187. [PubMed: 23320427]
 15. Aoki K, Tiemeyer M. The glycomics of glycan glucuronylation in *Drosophila melanogaster*. *Methods in enzymology.* 2010; 480:297–321. [PubMed: 20816215]
 16. Kumagai T, Katoh T, Nix DB, Tiemeyer M, Aoki K. In-gel beta-elimination and aqueous-organic partition for improved O- and sulfoglycomics. *Anal Chem.* 2013; 85:8692–8699. [PubMed: 23937624]
 17. 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]
 18. Hase S, Ibuki T, Ikenaka T. Reexamination of the pyridylamination used for fluorescence labelling of oligosaccharides and its application to glycoproteins. *J Biochem (Tokyo).* 1984; 95:197–203. [PubMed: 6706908]
 19. 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]
 20. Dragosits M, Pflugl S, Kurz S, Razzazi-Fazeli E, Wilson IBH, Rendi D. Recombinant *Aspergillus* β -galactosidases as a robust glycomic and biotechnological tool. *Applied microbiology and biotechnology.* 2014; 98:3553–3567. [PubMed: 24037406]
 21. Dragosits M, Yan S, Razzazi-Fazeli E, Wilson IBH, Rendi 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]
 22. Anumula KR, Taylor PB. A comprehensive procedure for preparation of partially methylated alditol acetates from glycoprotein carbohydrates. *Analytical biochemistry.* 1992; 203:101–108. [PubMed: 1524204]
 23. Stoffyn P, Stoffyn A. Direct conversion of sulfatides into cerebroside. *Biochimica et biophysica acta.* 1963; 70:107–108. [PubMed: 13984258]
 24. Sugita M, Dulaney JT, Moser HW. Structure and composition of sulfatides isolated from livers of patients with metachromatic leukodystrophy: galactosyl sulfatide and lactosyl sulfatide. *Journal of lipid research.* 1974; 15:227–233. [PubMed: 4363968]
 25. Schulz BL, Packer NH, Karlsson NG. Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal Chem.* 2002; 74:6088–6097. [PubMed: 12498206]
 26. Paschinger K, Rendi D, Wilson IBH. Revealing the anti-HRP epitope in *Drosophila* and *Caenorhabditis*. *Glycoconj J.* 2009; 26:385–395. [PubMed: 18726691]
 27. 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 in press.
 28. Fabini G, Freilinger A, Altmann F, Wilson IBH. Identification of core α 1,3-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]
 29. Tomiya N, Awaya J, Kurono M, Endo S, Arata Y, Takahashi N. Analyses of N-linked oligosaccharides using a two-dimensional mapping technique. *Analytical biochemistry.* 1988; 171:73–90. [PubMed: 3407923]
 30. 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 Man₉GlcNAc₂. *Analytical biochemistry*. 1991; 193:90–100. [PubMed: 2042746]
31. Wong-Madden ST, Landry D. Purification and characterization of novel glycosidases from the bacterial genus *Xanthomonas*. *Glycobiology*. 1995; 5:19–28. [PubMed: 7772864]
 32. 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 ¹H NMR spectroscopy. *J Biol Chem*. 1988; 263:18253–18268. [PubMed: 2461366]
 33. Oguri S, Minowa MT, Ihara Y, Taniguchi N, Ikenaga H, Takeuchi M. Purification and characterization of UDP-N-acetylglucosamine: α1,3-D-mannoside β1,4-N-acetylglucosaminyltransferase (N-acetylglucosaminyltransferase-IV) from bovine small intestine. *Journal of Biological Chemistry*. 1997; 272:22721–22727. [PubMed: 9278430]
 34. Lopez M, Gazon M, Juliant S, Plancke Y, Leroy Y, Strecker G, Cartron JP, Bailly P, Cerutti M, Verbert A, Delannoy P. Characterization of a UDPGal:Galβ1-3GalNAc α1,4-galactosyltransferase activity in a *Mamestra brassicae* cell line. *J Biol Chem*. 1998; 273:33644–33651. [PubMed: 9837949]
 35. Butters TD, Hughes RC. Isolation and characterisation of mosquito cell membrane glycoproteins. *Biochimica et biophysica acta*. 1981; 640:655–671. [PubMed: 7213698]
 36. Butters TD, Hughes RC, Vischer P. Steps in the biosynthesis of mosquito cell membrane glycoproteins and the effects of tunicamycin. *Biochim. Biophys. Acta*. 1981; 640:672–686. [PubMed: 7213699]
 37. 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]
 38. Wilkins S, Billingsley PF. Partial characterization of oligosaccharides expressed on midgut microvillar glycoproteins of the mosquito, *Anopheles stephensi* Liston. *Insect biochemistry and molecular biology*. 2001; 31:937–948. [PubMed: 11483430]
 39. Rudin W, Hecker H. Lectin-binding sites in the midgut of the mosquitoes *Anopheles stephensi* Liston and *Aedes aegypti* L. (Diptera: Culicidae). *Parasitology research*. 1989; 75:268–279. [PubMed: 2649879]
 40. Andrews L, Laughinghouse A, Sina BJ. Lectin binding characteristics of male and female salivary gland proteins of *Anopheles gambiae*: Identification and characterization of female specific glycoproteins. *Insect biochemistry and molecular biology*. 1997; 27:159–166.
 41. Li JS, Li J. Characterization of N-linked oligosaccharides in chorion peroxidase of *Aedes aegypti* mosquito. *Protein science : a publication of the Protein Society*. 2005; 14:2370–2386. [PubMed: 16131661]
 42. Crispin M, Harvey DJ, Bitto D, Bonomelli C, Edgeworth M, Scrivens JH, Huiskonen JT, Bowden TA. Structural plasticity of the Semliki Forest virus glycome upon interspecies transmission. *Journal of proteome research*. 2014; 13:1702–1712. [PubMed: 24467287]
 43. Cime-Castillo J, Delannoy P, Mendoza-Hernandez G, Monroy-Martinez V, Harduin-Lepers A, Lanz-Mendoza H, Hernandez-Hernandez Fde L, Zenteno E, Cabello-Gutierrez C, Ruiz-Ordaz BH. Sialic Acid Expression in the Mosquito *Aedes aegypti* and Its Possible Role in Dengue Virus-Vector Interactions. *Biomed Res Int*. 2015; 2015:504187. [PubMed: 25874215]
 44. Spiro RG, Bhooyroo VD. Occurrence of sulphate in the asparagine-linked complex carbohydrate units of thyroglobulin. Identification and localisation of galactose-3-sulphate and N-acetylglucosamine-6-sulphate residues in the human and calf proteins. *J Biol Chem*. 1988; 263:14351–14358. [PubMed: 3170547]
 45. Green ED, Baenziger JU. Asparagine-linked oligosaccharides on lutropin, follitropin, and thyrotropin. I. Structural elucidation of the sulfated and sialylated oligosaccharides on bovine, ovine, and human pituitary glycoprotein hormones. *J Biol Chem*. 1988; 263:25–35. [PubMed: 3121609]
 46. Murakami T, Natsuka S, Nakakita S, Hase S. Structure determination of a sulfated N-glycans, candidate for a precursor of the selectin ligand in bovine lung. *Glycoconj J*. 2007; 24:195–206. [PubMed: 17356912]

47. Freeze HH. Mannose 6-sulfate is present in the N-linked oligosaccharides of lysosomal enzymes of Dictyostelium. Arch Biochem Biophys. 1985; 243:690–693. [PubMed: 2935083]
48. Barboza M, Duschak VG, Fukuyama Y, Nonami H, Erra-Balsells R, Cazzulo JJ, Couto AS. Structural analysis of the N-glycans of the major cysteine proteinase of Trypanosoma cruzi. Identification of sulfated high-mannose type oligosaccharides. The FEBS journal. 2005; 272:3803–3815. [PubMed: 16045752]
49. van Kuik JA, Breg J, Kolsteeg CEM, Kamerling JP, Vliegthart JFG. Primary Structure of the Acidic Carbohydrate Chain of Hemocyanin from Panulirus interruptus. FEBS letters. 1987; 221:150–154.
50. Martin G, Sorokine O, Moniatte M, Bulet P, Hetru C, Van Dorsselaer A. The structure of a glycosylated protein hormone responsible for sex determination in the isopod, Armadillidium vulgare. European journal of biochemistry / FEBS. 1999; 262:727–736. [PubMed: 10411634]
51. Gallego RG, Blanco JL, Thijssen-van Zuylen CW, Gotfredsen CH, Voshol H, Duus JO, Schachner M, Vliegthart JFG. Epitope diversity of N-glycans from bovine peripheral myelin glycoprotein P0 revealed by mass spectrometry and nano probe magic angle spinning ¹H NMR spectroscopy. J Biol Chem. 2001; 276:30834–30844. [PubMed: 11410585]
52. Morise J, Kizuka Y, Yabuno K, Tonoyama Y, Hashii N, Kawasaki N, Manya H, Miyagoe-Suzuki Y, Takeda S, Endo T, Maeda N, Takematsu H, Oka S. Structural and biochemical characterization of O-mannose-linked human natural killer-1 glycan expressed on phosphacan in developing mouse brains. Glycobiology. 2014; 24:314–324. [PubMed: 24352591]
53. 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]
54. Keller M, Dennis RD, Weske B, Wiegandt H. Immunochemical analysis of a monoclonal antibody recognizing a terminal glucuronic acid-containing epitope of insect acidic glycolipids. Hybridoma. 1990; 9:295–307. [PubMed: 1698706]
55. Dennis RD, Antonicek H, Wiegandt H, Schachner M. Detection of the L2/HNK-1 carbohydrate epitope on glycoproteins and acidic glycolipids of the insect Calliphora vicina. Journal of neurochemistry. 1988; 51:1490–1496. [PubMed: 2459311]
56. 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]
57. Kimura Y, Ushijima T, Maeda M, Hama Y, Kimura M, Okihara K, Sugimoto H, Yamada H. Tumor antigen occurs in N-glycan of royal jelly glycoproteins: honeybee cells synthesize T-antigen unit in N-glycan moiety. Bioscience, biotechnology, and biochemistry. 2006; 70:2583–2587.
58. Kubelka V, Altmann F, Staudacher E, Tretter V, März L, Hård K, Kamerling JP, Vliegthart JFG. Primary structures of the N-linked carbohydrate chains from honeybee venom phospholipase A₂. European journal of biochemistry / FEBS. 1993; 213:1193–1204. [PubMed: 8504812]
59. Kimura Y, Sakamura S, Ushijima T, Hama Y, Kajiura H, Fujiyama K, Okihara K, Hashimoto K, Sugimoto H, Yamada H. Evidence for new β1-3 galactosyltransferase activity involved in biosynthesis of unusual N-glycan harboring T-antigen in Apis mellifera. Bioscience, biotechnology, and biochemistry. 2007; 71:1111–1114.
60. 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. Glycoconjugate Journal. 2015 in press, doi: 10.1007/s10719-10015-19585-10717.
61. Yamamoto-Hino M, Yoshida H, Ichimiya T, Sakamura S, Maeda M, Kimura Y, Sasaki N, Aoki-Kinoshita KF, Kinoshita-Toyoda A, Toyoda H, Ueda R, Nishihara S, Goto S. Phenotype-based clustering of glycosylation-related genes by RNAi-mediated gene silencing. Genes Cells. 2015 10.1111/gtc.12246.
62. 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]
63. Bleuler-Martinez S, Butschi A, Garbani M, Walti MA, Wohlschlager T, Potthoff E, Sabotic J, Pohleven J, Luthy P, Hengartner MO, Aebi M, Kunzler M. A lectin-mediated resistance of higher fungi against predators and parasites. *Molecular ecology.* 2011; 20:3056–3070. [PubMed: 21486374]
64. Breloy I, Schwientek T, Lehr S, Hanisch FG. Glucuronic acid can extend O-linked core. 1 glycans, but it contributes only weakly to the negative surface charge of *Drosophila melanogaster* Schneider-2 cells. *FEBS letters.* 2008; 582:1593–1598. [PubMed: 18417079]
65. Garenaux E, Maes E, Leveque S, Brassart C, Guerardel Y. Structural characterization of complex O-linked glycans from insect-derived material. *Carbohydrate research.* 2011; 346:1093–1104. [PubMed: 21536259]
66. Weske B, Dennis RD, Helling F, Keller M, Nores GA, Peter-Katalinic J, Egge H, Dabrowski U, Wiegandt H. Glycosphingolipids in insects. Chemical structures of two variants of a glucuronic-acid-containing ceramide hexasaccharide from a pupae of *Calliphora vicina* (Insecta: Diptera), distinguished by a N-acetylglucosamine-bound phosphoethanolamine sidechain. *European journal of biochemistry / FEBS.* 1990; 191:379–388. [PubMed: 2384087]
67. Navarro-Sanchez E, Altmeyer R, Amara A, Schwartz O, Fieschi F, Virelizier JL, Arenzana-Seisdedos F, Despres P. Dendritic-cell-specific ICAM3-grabbing nonintegrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. *EMBO reports.* 2003; 4:723–728. [PubMed: 12783086]
68. Francischetti IM, Ma D, Andersen JF, Ribeiro JM. Evidence for a lectin specific for sulfated glycans in the salivary gland of the malaria vector, *Anopheles gambiae*. *PloS one.* 2014; 9:e107295. [PubMed: 25207644]
69. Kunz S, Rojek JM, Kanagawa M, Spiropoulou CF, Barresi R, Campbell KP, Oldstone MB. Posttranslational modification of α -dystroglycan, the cellular receptor for arenaviruses, by the glycosyltransferase LARGE is critical for virus binding. *Journal of virology.* 2005; 79:14282–14296. [PubMed: 16254363]
70. Shukla D, Liu J, Blaiklock P, Shworak NW, Bai X, Esko JD, Cohen GH, Eisenberg RJ, Rosenberg RD, Spear PG. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus. 1 entry. *Cell.* 1999; 99:13–22. [PubMed: 10520990]

Highlights

- The N- and O-glycans of mosquito larvae were analysed by off-line MALDI-TOF MS and by ESI-MS or LC-MS
- Powerful combination of HPLC, exoglycosidase treatments and mass spectrometry for detailed structure elucidation
- Sulphated and glucuronylated N- and O-glycans are present also in *Drosophila*

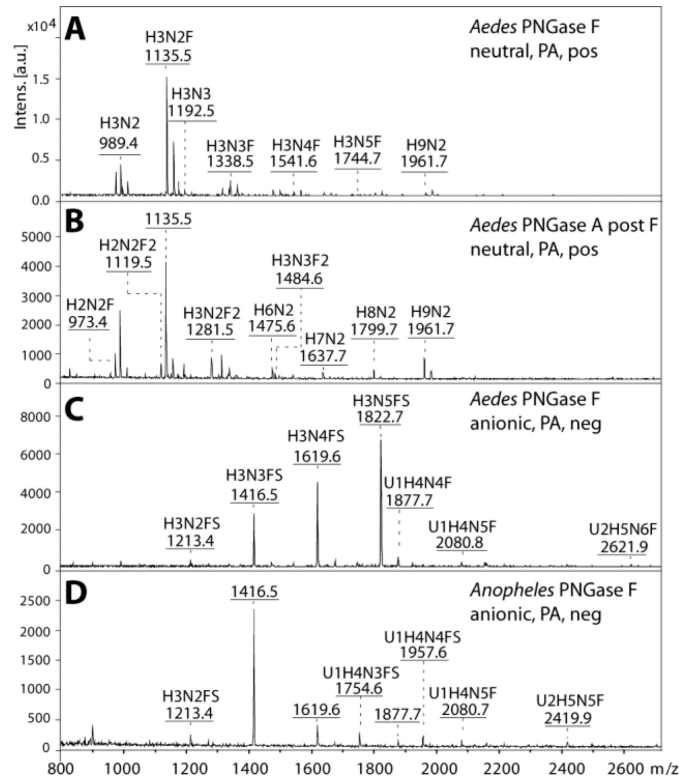


Figure 1. Mass spectrometric analysis of neutral and anionic mosquito N-glycan pools
 Released N-glycans were enriched in neutral and anionic pools prior to pyridylamino-
 labelling and MALDI-TOF/TOF MS. The spectra shown are those in positive (A, B) and
 negative ion mode (C, D) of the N-glycan pools released with PNGase F (A, C, D) or
 PNGase A post PNGase F (B) from *Aedes aegypti* (A-C) and *Anopheles gambiae* (D).
 Abbreviations for annotated $[M+H]^+$ and $[M-H]^-$ ions are: H Hexose, N N-
 acetylhexosamine, F Fucose, S Sulphate and U Glucuronic Acid. The positive ion mode
 spectra of the neutral *Anopheles* PNGase F and A pools are not shown due to their similarity
 to the *Aedes*; the differences between the two species, limited to the lower intensity glycans,
 were only resolved after HPLC fractionation.

Compos.	[M+H] ⁺	<i>Aedes</i>	<i>Anoph.</i>	RP g.u.	Structure
H1N2F (00F ³)	811.35	+	+	8.5 (F)	
H2N2 (M)	827.34	+	+	7.3 (F)	
H2N2F (0MF ³)	973.40	+	+	8.5 (F)	
H2N2F (M0F ³)	973.40	+++	+++	9.6 (A/F)	
H3N2 (MM)	989.39	++	++	7.3 (A/F)	
H2N2F2 (M0F ³ F ³)	1119.46	++	nd	6.9 (A)	
H3N2F (MMF ³)	1135.45	+	+	5.2 (A)	
H3N2F (MMF ³)	1135.45	+++	+++	9.6 (A/F)	
H4N2 (Man4)	1151.45	+	+	7.3 (F)	
H2N3F	1176.48	trace	trace	8.7 (F)	
H3N3 (MGn)	1192.47	++	++	6.7 (A/F)	
H3N3 (GnM)	1192.47	trace	trace	9.0 (A/F)	
H3N2F2 (MMF ³ F ³)	1281.51	++	+	6.9 (A)	
H5N2 (Man5)	1313.50	++	++	7.3 (A/F)	
H3N3F (MGnF ³)	1338.53	+	+	9.0 (A/F)	
H3N3F	1338.53	+	trace	11.4 (F)	
H3N3F (GnMF ³)	1338.53	+	+	12.5 (A/F)	
H4N3	1354.53	+	+	6.7 (F)	
H3N4 (GnGn)	1395.55	+	+	8.0 (F)	
H6N2 (Man6)	1475.55	+	+	6.3 (A/F)	
H6N2 (Man6)	1475.55	trace	trace	8.0 (F)	
H3N3F2 (MGnF ³ F ³)	1484.59	+	nd	6.5 (A)	
H3N3F2 (GnMF ³ F ³)	1484.59	+	+	7.9 (A)	
H4N3F	1500.58	trace	trace	9.0 (F)	
H5N3	1516.58	+	+	6.7 (F)	
H3N4F (GnGnF ³)	1541.61	+	+	10.8 (A/F)	
H3N4F	1541.61	trace	nd	14.8 (F)	
H7N2 (Man7)	1637.60	++	++	5.5 (A/F)	
H7N2 (Man7)	1637.60	+	+	6.1 (A/F)	
H5N3F	1662.64	trace	trace	8.7 (F)	
H3N4F2	1687.67	nd	trace	8.2 (F)	
H3N5F	1744.69	trace	trace	11.4 (F)	
H3N5F	1744.69	+	+	13.5 (F)	
H8N2 (Man8B)	1799.66	++	++	5.2 (A/F)	
H8N2 (Man8A)	1799.66	+	+	5.9 (A/F)	
H3N6F	1947.77	trace	nd	13.5 (F)	
H9N2 (Man9)	1961.71	++	++	5.5 (A/F)	
H10N2 (GlcMan9)	2123.76	+	nd	6.7 (A/F)	

Figure 2. Neutral N-glycans from *Aedes* and *Anopheles* larvae

The predicted and observed m/z values for pyridylaminated neutral N-glycans isolated by RP-amide HPLC and analysed by MALDI-TOF MS are listed along with elution times (in glucose units, g.u.) and proposed structures according to the pictorial nomenclature of the Consortium for Functional Glycomics, supplemented in certain cases with the Schachter-style nomenclature (MM, MGn, GnM, etc). A, F or A/F indicates a glycan detected only in the PNGase A (post PNGase F) pool, PNGase F pool or both; nd, not detected in fractions of either species. Estimation of glycan amounts is indicated (scale from ‘+++’ through to ‘trace’), whereby ‘+++’ is for the major components present in the HPLC fractions with the highest fluorescence intensity.

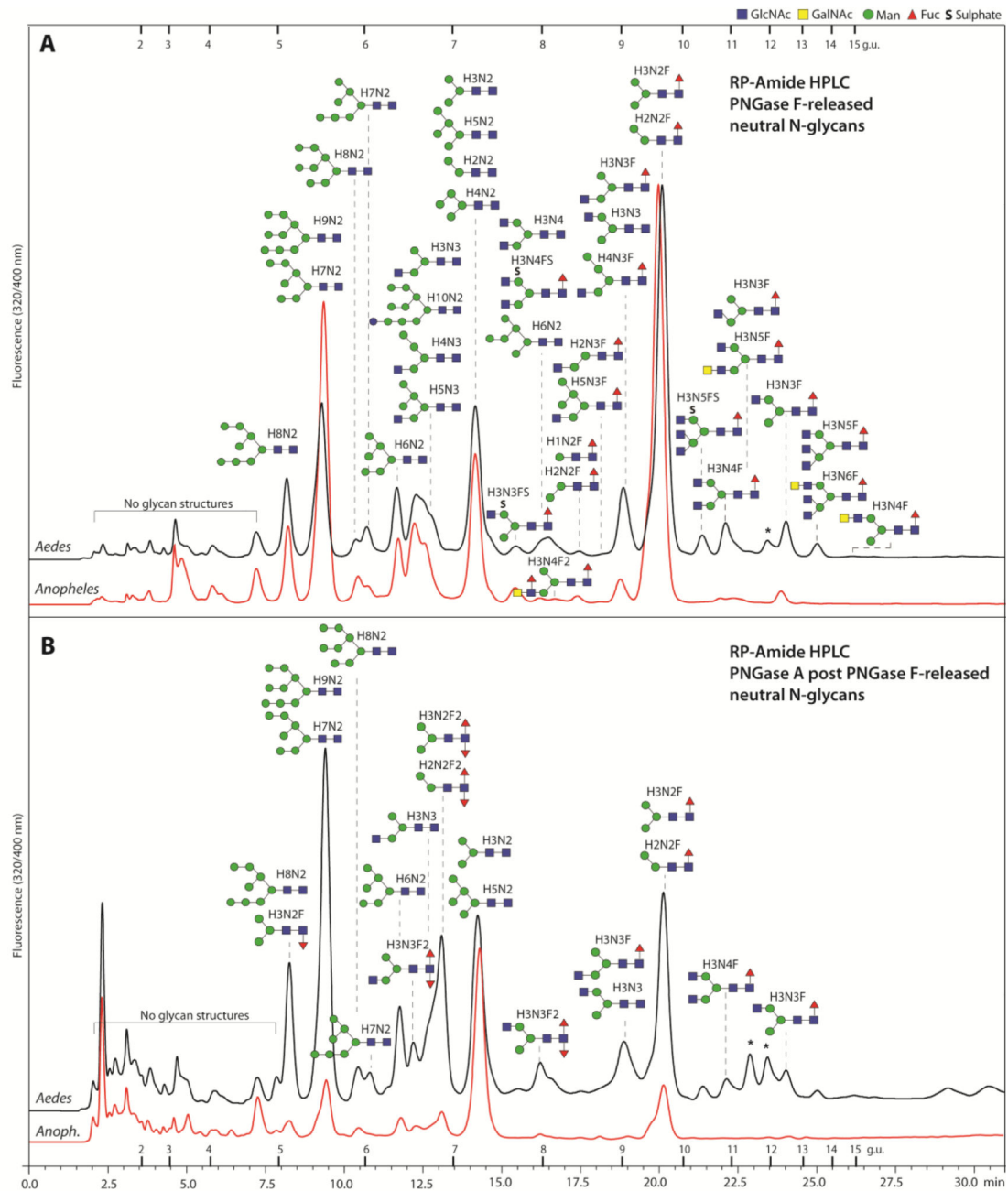


Figure 3. RP-Amide HPLC of mosquito neutral N-glycans

The pyridylaminated neutral N-glycan pools of both mosquito species (*Aedes* and *Anopheles*) released by PNGase F (**A**) and PNGase A post PNGase F (**B**) were fractionated using a RP-Amide column. Selected peaks are annotated with the proposed N-glycan structures using the symbolic nomenclature of the Consortium for Functional Glycomics (circles, hexose; squares, N-acetylhexosamine; triangles, deoxyhexose; S, sulphate) as well as abbreviated compositions of the form $H_xN_yF_z$; glycans eluting in one peak are shown according to relative intensity in the MALDI-TOF MS spectra (most abundant uppermost). The downward or upward depiction of the core fucose indicates respectively an α 1,3 or α 1,6-linkage; most of the monofucosylated species in the PNGase A post F chromatogram

are α 1,6-fucosylated as judged by retention time, whereas the occurrence of GalNAc is presumed from other studies on insect N-glycans. The elution positions of dextran standards are annotated with the relevant glucose units and peaks representing structures derived from the larval food source are marked with asterisks. The chromatograms are normalised with respect to each other; the original scales are: 4000 and 1850 mV (*A. gambiae* and *A. aegypti* F digest) and 125 and 85 (*A. gambiae* and *A. aegypti* A after F digest).

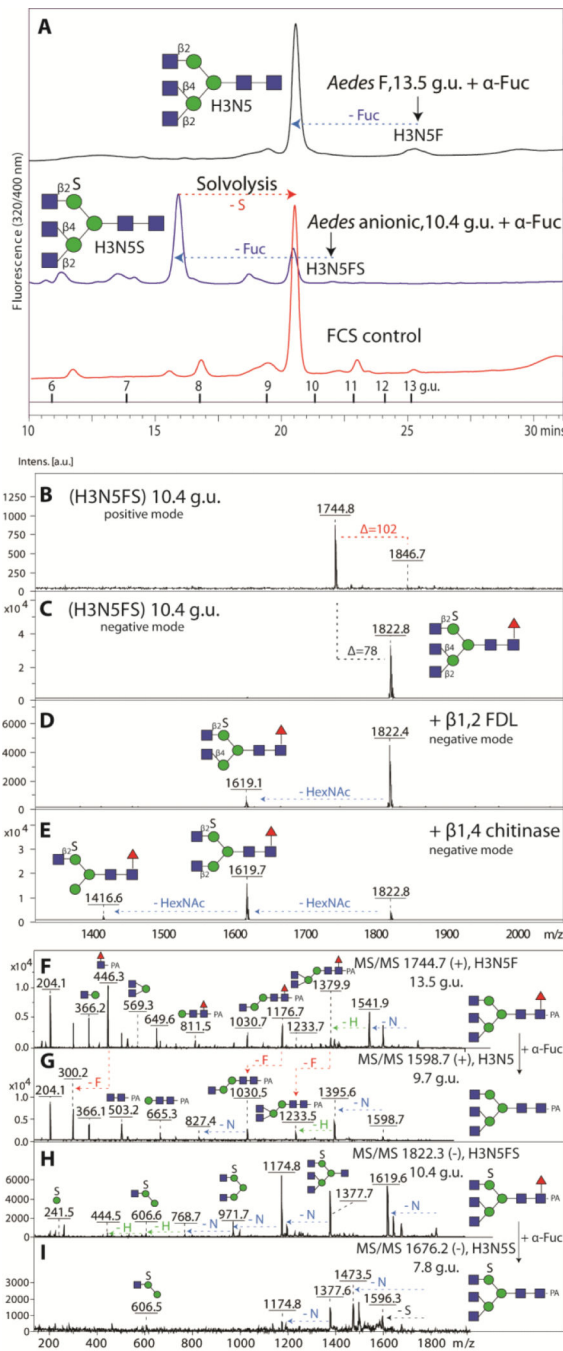


Figure 4. Elucidation of neutral and anionic triantennary N-glycan structures
(A) Coelution of *Aedes* derived triantennary glycans after either defucosylation of the neutral glycan species or solvolysis and defucosylation of the sulphated glycan species with an asialoagalacto-triantennary N-glycan isolated from FCS; note that solvolysis under these conditions is incomplete, but that further incubation resulted in unspecific reaction products. The arrows indicate the shifts in retention time. MALDI TOF/TOF MS spectra in positive **(B)** and negative ion mode **(C-E)** of the *Aedes* anionic triantennary structure eluting at 10.4 g.u. on RP-Amide HPLC show that the anionic triantennary structure is sensitive to β 1,2-

specific FDL (**D**) and β 1,4-specific 'chitinase-type' hexosaminidase (**E**). MS/MS spectra of parent and product glycan species upon bovine fucosidase digests in positive (**F**, **G**) and negative ion mode (**H**, **I**) are shown.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

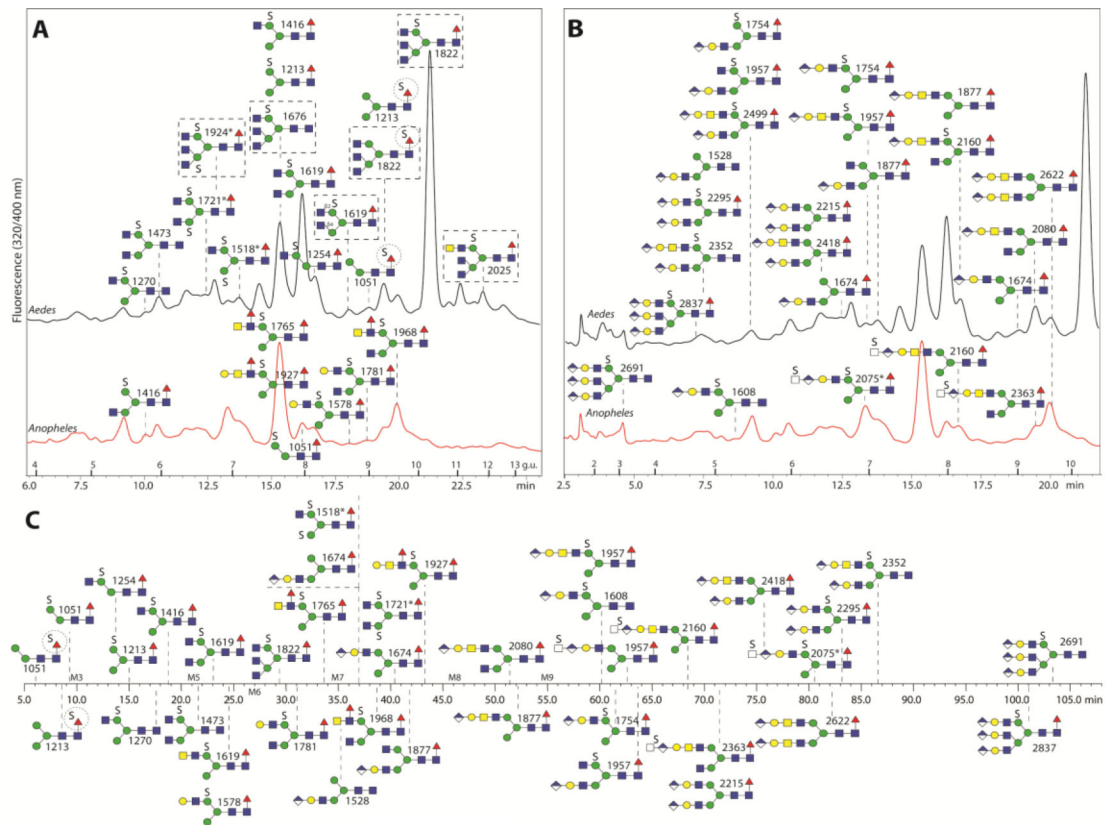


Figure 5. HPLC analysis of anionic N-glycans

Anopheles and *Aedes* anionic-enriched N-glycans were chromatographed either on RP-amide (both; respectively red and black chromatograms) or HIAX (*Anopheles* alone). The same RP-amide chromatograms are annotated with either the sulphated (**A**) or glucuronylated (**B**) structures. Glycans above the *Aedes* chromatogram are found in both species, other than those in boxes which are found in *Aedes* alone; glycans depicted below the *Aedes* chromatogram are found solely in *Anopheles*. The circles highlight putative sulphation of fucose residues. The HIAX schematic chromatogram (**C**) is shown as a time-scale annotated with selected glycan structures. The triantennary structures in **B** and **C** are putative and are based on in-source fragmentation rather than standard MS/MS (see Supplementary Figure 9). In general the higher the number of anionic groups, the earlier the elution is on RP-HPLC and the later on the HIAX column, which also separates according to size; on the other hand, α 1,6-fucosylation is associated with late RP-HPLC retention. Structures are annotated with m/z values from negative ion mode; the disulphated structures marked with an asterisk were detected as $[M-2H+Na]^-$ or $[M-2H+K]^-$.

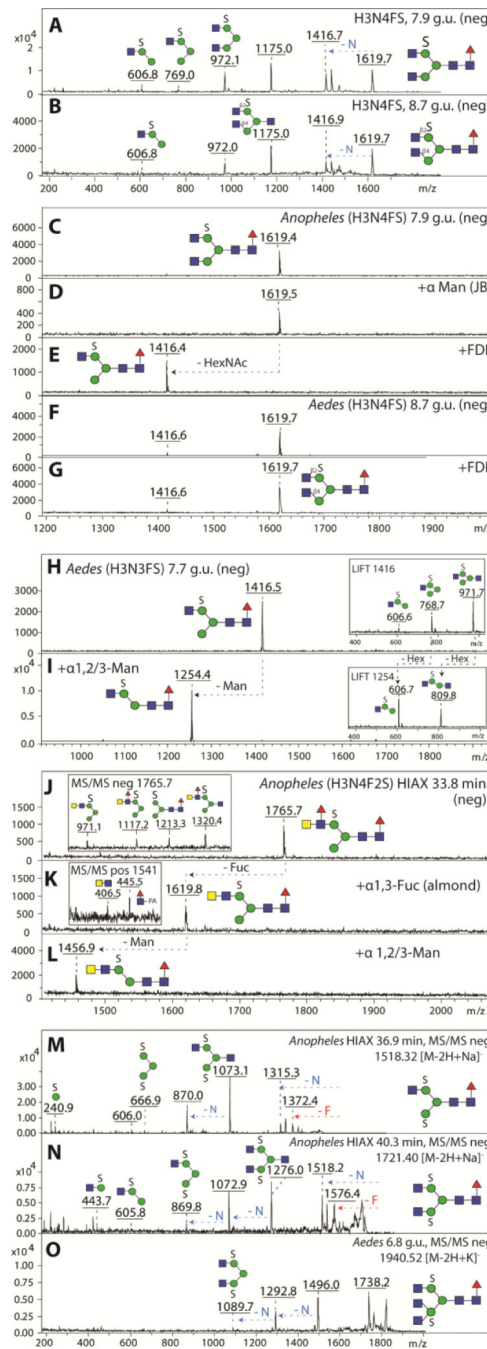


Figure 6. Elucidation of sulphated N-glycan structures in negative ion mode MS
 MS/MS of isomeric biantennary sulphated glycans (H3N4FS, m/z 1619) eluting at 7.9 (A) and 8.7 g.u. (B) are rather similar, but showed different susceptibility to β 1,2-specific FDL hexosaminidase. The earlier eluting isomer of *Anopheles* (C) is resistant to jack bean α -mannosidase (D), and sensitive to β 1,2-specific FDL (E), whereas the later eluting isomer predominantly found in *Aedes* (F) remains resistant to β 1,2-specific FDL treatment (G) suggesting a β 1,4-linked GlcNAc residue on the α 1,3-mannose arm. A sulphated pseudohybrid glycan species of *Aedes* (H3N3FS, m/z 1416, H) is sensitive towards α -1,2/3-

specific mannosidase treatment (**I**) thereby reducing the MS/MS key fragments by $m/z = 162$ (*insets*). A glycan (H3N4F2S, m/z 1765, **J**) in an *Anopheles* HIAX fraction displays MS/MS key fragments indicating the presence of a fucosylated LacdiNAc motif in addition to sulphated mannose (*inset* in **J**) and loses one fucose upon treatment with α 1,3-fucosidase (**K**) and one mannose upon subsequent α 1,2/3-mannosidase treatment (**L**). Positive ion mode MS/MS upon α 1,3-fucosidase indicated the presence of core fucose (*inset* in **K**). MS/MS of disulphated N-glycans of the compositions H3N3FS2 (**M**), H3N4FS2 (**N**) and H3N5FS2 (**O**) are compatible with sulphation of both α -mannose residues; a further indication for this position is that these glycans were sensitive to jack bean β -hexosaminidase, but resistant to jack bean α -mannosidase.

Compos.	predicted masses [M+H] ⁺ or [M+H-S] ⁺	[M-H] ⁻	Aedes RP [M-H] ⁻	Anoph. RP [M-H] ⁻	Anoph. HIAX [M-H] ⁻	Anionic structures	Compos.	predicted masses [M+H] ⁺ or [M+H-S] ⁺	[M-H] ⁻	Aedes RP [M-H] ⁻	Anoph. RP [M-H] ⁻	Anoph. HIAX [M-H] ⁻	Anionic structures
H2N2FS	973.40	1051.34	nd	1051.37 (8.0 g.u.)	1051.31 (9.5 min)		H3N5FS	1744.69	1822.63	1822.80 (9.2 g.u.)	nd	nd	
H2N2FS	973.40	1051.34	1051.43 (9.0 g.u.)	1051.4 (9.0 g.u.)	1051.06 (5.8 min)		H3N5FS	1744.69	1822.63	1822.82 (10.4 g.u.)	1822.76 (10.4 g.u.)	1822.78 (29.3 min)	
H3N2FS	1135.45	1213.39	1213.48 (7.7 g.u.)	1213.45 (7.7 g.u.)	1213.42 (15.0 min)		U1H4N4F	1879.69	1877.68	1877.84 (7.2 g.u.)	1877.81 (7.2 g.u.)	1877.76 (41.8 min)	
H3N2FS	1135.45	1213.39	1213.50 (9.2 g.u.)	1213.45 (9.2 g.u.)	1213.38 (8.8 min)		U1H4N4F	1879.69	1877.68	1877.85 (8.3 g.u.)	1877.74 (8.3 g.u.)	1877.66 (52.2 min)	
H2N3FS	1176.48	1254.42	1254.53 (8.3 g.u.)	1245.46 (8.3 g.u.)	1254.41 (13.8 min)		H3N5FS2		1924.57 [M2H+Na]	1924.73 (6.8 g.u.)	nd	nd	
H3N3S	1192.47	1270.41	1270.55 (5.8 g.u.)	1270.48 (5.8 g.u.)	1270.44 (17.6 min)		H4N4F2S	1849.72	1927.66	nd	1927.62 (7.7 g.u.)	1927.78 (43.3 min)	
H3N3FS	1338.53	1416.47	nd	1416.5 (5.8 g.u.)	nd		U1H4N4FS	1879.69	1957.64	1958.45 (5.5 g.u.)	1957.75 (5.5 g.u.)	1957.8 (63.4 min)	
H3N3FS	1338.53	1416.47	1416.57 (7.7 g.u.)	1416.5 (7.7 g.u.)	1416.49 (18.8 min)		U1H4N4FS	1879.69	1957.64	1957.74 (7.0 g.u.)	1957.76 (7.0 g.u.)	1957.78 (60.3 min)	
H3N4S	1395.55	1473.49	1473.64 (6.0 g.u.)	1473.56 (6.0 g.u.)	1473.57 (21.8 min)		U1H4N4FS	1879.69	1957.64	nd	nd	1957.81 (62.4 min)	
H3N3FS2		1518.41 [M2H+Na]	1518.52 (7.2 g.u.)	1518.43 (7.2 g.u.)	1518.42 (36.9 min)		H3N5FS2	1890.71	1968.69	nd	1967.79 (9.6 g.u.)	1968.8 (38.7 min)	
U1H4N3	1530.56	1528.54	1528.46 (4.8 g.u.)	1528.61 (4.8 g.u.)	1528.59 (35.2 min)		H3N6FS	1947.77	2025.71	2025.94 (11.9 g.u.)	nd	nd	
H4N3FS	1500.58	1578.53	nd	1578.6 (8.7 g.u.)	1578.57 (24.4 min)		U1H4N4FS2	1879.69	2075.55 [M2H+K]	nd	2074.90 (7.0 g.u.)	2075.44 (80.6 min)	
H3N4FS	1541.61	1619.55	1619.68 (7.9 g.u.)	1619.59 (7.9 g.u.)	1619.58 (23.0 min)		U1H4N5F	2082.77	2080.76	2080.97 (9.2 g.u.)	2080.87 (9.2 g.u.)	2080.70 (51.2 min)	
H3N4FS	1541.61	1619.55	1619.72 (8.7 g.u.)	nd	nd		U1H4N5FS	2082.77	2160.72	2160.89 (8.3 g.u.)	nd	nd	
H3N4FS	1541.61	1619.55	nd	nd	1619.59 (24.4 min)		U1H4N5FS	2082.77	2160.72	nd	2160.76 (8.3 g.u.)	2160.88 (68.5 min)	
U1H4N3S	1530.56	1608.50	nd	1608.58 (60.3 min)	1608.57 (60.3 min)		U2H5N4F	2217.78	2215.77	2215.92 (6.4 g.u.)	2215.88 (6.4 g.u.)	2215.91 (71.3 min)	
U1H4N3F	1676.62	1674.60	1674.74 (6.7 g.u.)	1674.65 (6.7 g.u.)	1674.51 (36.9 min)		U2H5N4FS	2217.78	2295.72	2295.62 (4.8 g.u.)	2295.60 (4.8 g.u.)	2295.86 (83.4 min)	
U1H4N3F	1676.62	1674.60	1674.74 (9.0 g.u.)	1674.71 (9.0 g.u.)	1674.71 (40.3 min)		U2H5N5S	2274.80	2352.74	2352.65 (4.8 g.u.)	2352.65 (4.8 g.u.)	2352.65 (86.5 min)	
H3N5S	1598.63	1676.57	1676.7 (7.7 g.u.)	nd	nd		U1H4N6FS	2285.85	2363.80	nd	2363.88 (9.2 g.u.)	2363.95 (71.3 min)	
H3N4FS2		1721.49 [M2H+Na]	1721.65 (6.7 g.u.)	1721.53 (6.7 g.u.)	1721.62 (40.3 min)		U2H5N5F	2420.86	2418.84	2419.02 (6.4 g.u.)	2418.97 (6.4 g.u.)	2419.94 (75.4 min)	
U1H4N3FS	1676.62	1754.56	1754.69 (5.5 g.u.)	1754.67 (5.5 g.u.)	nd		U2H5N5FS	2420.86	2498.80	2498.93 (5.5 g.u.)	2498.92 (5.5 g.u.)	nd	
U1H4N3FS	1676.62	1754.56	1754.69 (7.0 g.u.)	1754.67 (7.0 g.u.)	1754.68 (61.5 min)		U2H5N6F	2623.94	2621.92	2622.17 (9.5 g.u.)	2622.04 (9.5 g.u.)	2622.18 (82.1 min)	
H3N4F2S	1687.67	1765.61	nd	1765.57 (7.7 g.u.)	1765.71 (33.8 min)		U3H6N5S	2612.88	2690.83	nd	2691.79 (3.1 g.u.)	2692.62 (103.4 min)	
H4N4FS	1703.66	1781.61	nd	1781.7 (8.9 g.u.)	1781.71 (31.0 min)		U3H6N5FS	2758.94	2836.89	2835.88 (4.7 g.u.)	2836.86 (4.7 g.u.)	2836.85 (101.0 min)	

Figure 7. Sulphated and glucuronylated N-glycans from *Aedes* and *Anopheles* larvae
 The predicted and observed m/z values (negative ion mode) for pyridylaminated sulphated and glucuronylated N-glycans isolated by RP-amide or HIAX HPLC and analysed by MALDI-TOF MS are listed along with elution times (in glucose units, g.u.) and proposed structures according to the pictorial nomenclature of the Consortium for Functional Glycomics; nd, not detected.

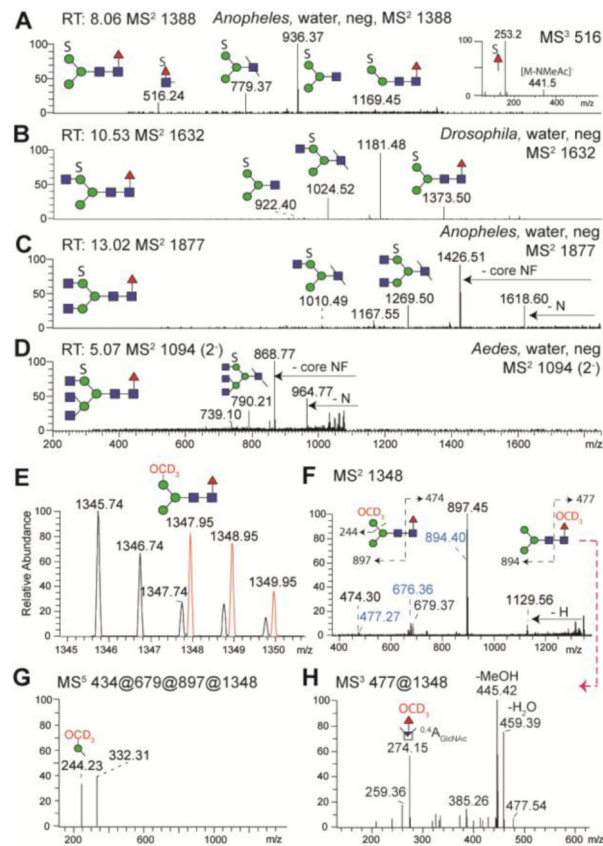


Figure 8. Detection of sulphated N-glycans by NSI-MS

PNGase F released sulphated N-glycans from *Anopheles*, *Aedes* and *Drosophila* were enriched upon permethylation in the water-phase and analysed by NSI-MS; automated TIM MS² negative ion mode spectra at certain scan run times (RT) are shown for the major sulphated N-glycans (A-D). After solvolysis and repermethylation with deuterated iodomethane CD₃I, a new deuterated glycan isotopic distribution (in red, E) was detected in positive ion mode of the organic phase which is 2.2 a.m.u. different from the neutral N-glycan (H₃N₂F, *m/z* 1345). The positive ion mode MS² spectrum (F) shows daughter ions indicating the loss of hexose (*m/z* 1129) and monofucosylated core N-acetylglucosamine (*m/z* 897). MS⁵ indicates that the deuteromethyl group is associated with hexose (*m/z* 244; G). The *m/z* 516 ion in panel A is consistent with the presence of sulphated fucose on a small proportion of the core fucosylated glycans (see MS³ inset and also **Supplementary Figure 6**). The MS² of the *m/z* 1348 repermethylated species further indicates the presence of low abundance fragments (F, in blue) originating from an originally sulphated core fucose; (H) MS³ of the *m/z* 477 ion results in an *m/z* 274 fragment indicative of a perdeuteromethylated fucose.

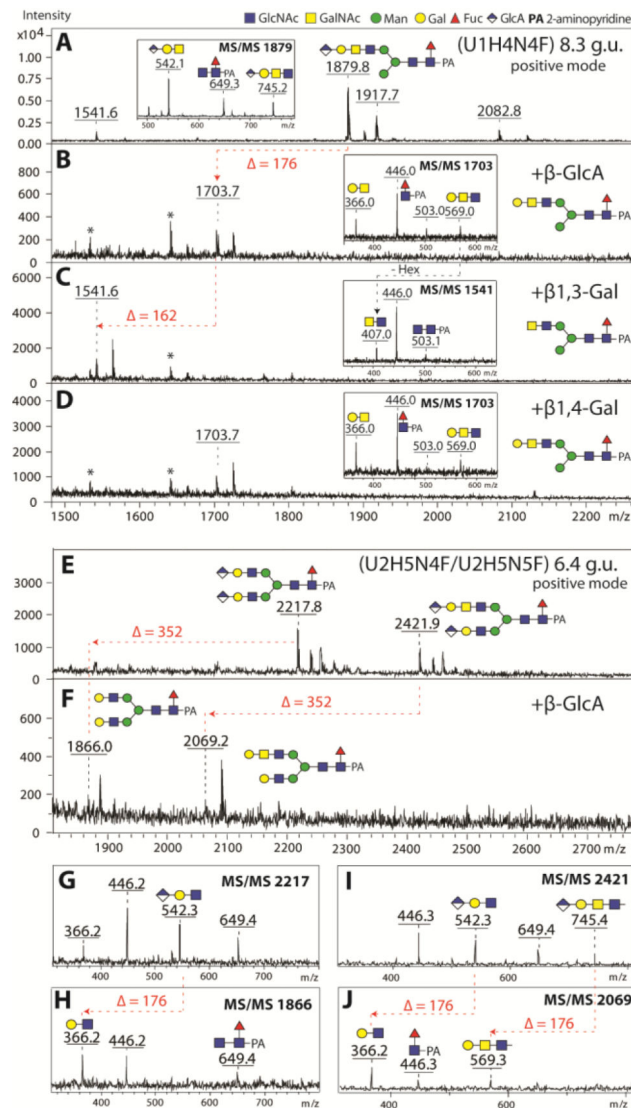


Figure 9. Structure elucidation of the glucuronylated N-glycans

Positive ion mode MALDI-TOF/TOF MS spectra of selected RP-Amide HPLC fractions (8.3 and 6.4 g.u.) of *Aedes* PNGase F released anionic N-glycans before (A, E) and after treatment with β -glucuronidase (β -GlcA; B, F) and subsequent β 1,3-specific (C) or β 1,4-specific (D) galactosidases. Successful removal of glucuronic acid and β 1,3-linked galactose moieties are indicated by the loss of 176 and 162 Da respectively. Removal of two GlcA residues from the biantennary glucuronylated N-glycan structures (m/z 2217 and 2422) is indicated by the loss of 352 Da (F). Insets in the spectra represent positive ion mode MS/MS spectra of untreated parent (m/z 1879) or exoglycosidase product glycan species (m/z 1703, 1541 and 1703). Panels G-J show the MS/MS for the parent and digested forms of the diglucuronylated glycans whose MS spectra are shown in E and F. Asterisks are indicating contaminating signals from the enzyme preparations. The proposed structure of the major glycan in each spectrum is also shown.

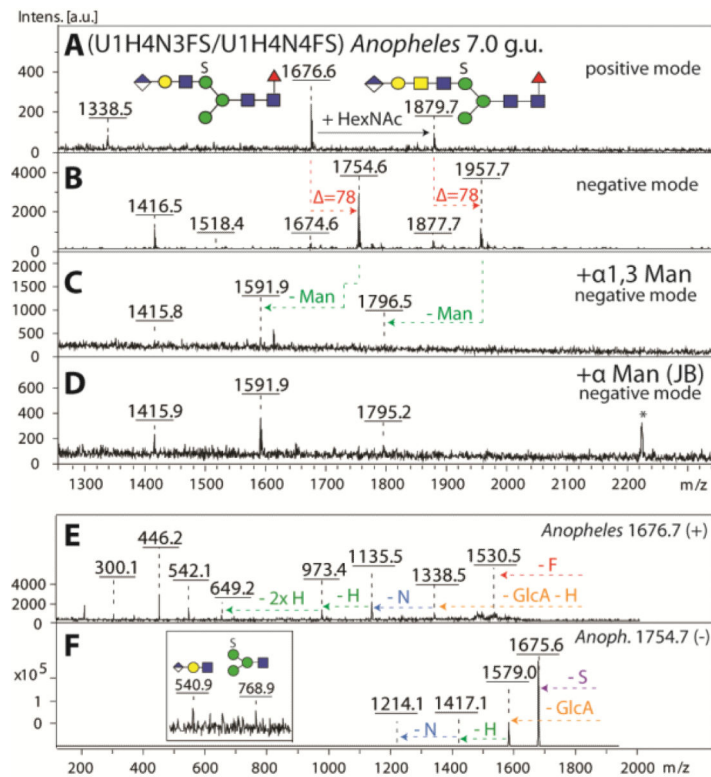


Figure 10. Definition of the antennal isomeric status of a N-glycan modified with both sulphate and glucuronic acid

A selected RP-Amide HPLC fraction (7.0 g.u.) of PNGase F released anionic N-glycans from *Anopheles* was analysed in positive (A) and negative (B) ion modes before (A, B) and after exoglycosidase treatments (C, D). As judged by negative ion mode MS, both the α 1,3-specific (C) and the jack bean α -mannosidases (D) were able to liberate one mannose residue from the parent glycan species (m/z 1754 and 1957). MS/MS of the most abundant parent glycan species in panel A was performed in positive (m/z 1676, E) and negative (m/z 1754, F) ion modes and offered further evidence for the structure of the backbone as well as for the position of the anionic modifications. The asterisk in (D) indicates a contaminating signal from the jack bean enzyme preparation.

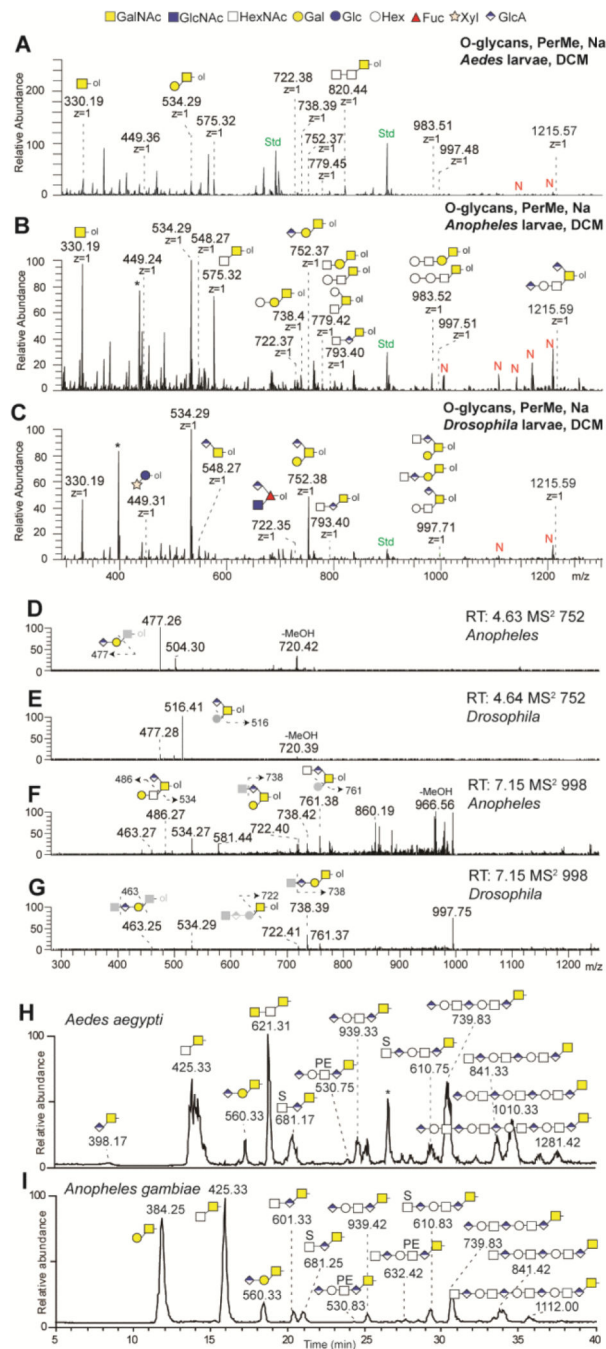


Figure 11. O-glycan analyses by NSI-MS and LC-MS

O-glycans from *Aedes* (A), *Anopheles* (B) and *Drosophila* (C) larvae were released by reductive β -elimination, permethylated, enriched in the organic (DCM) phase and analysed by NSI-MS in positive ion mode. Singly charged mainly core-1 modified O-glycans were detected. Glucuronylated O-glycans (m/z 752, 998, 1216) were predominantly found in *Drosophila* larvae. The automated TIM MS² spectra of the three glucuronylated structures found in *Anopheles* (D, F) and in *Drosophila* (E, G) larvae are presenting a mixture of similar daughter ions. Fragmentation schemes are shown in each MS² spectra. Native, α -

eliminated O-glycans from *Aedes* (**H**) and *Anopheles* (**I**) were also analysed by LC-MS in negative ion mode.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Composition	predicted m/z [M+Na] ⁺	observed m/z [M+Na] ⁺			predicted m/z [M-H] ⁻ ([M-2H] ²⁻)	observed m/z	
		Aedes	Ag	Dm		Aedes	Ag
Hex-ol	289.16	289.18	289.17	289.29	181.07	nd	nd
HexNac-ol	330.19	330.19	330.19	330.19	222.10	nd	nd
Pent ₁ Hex-ol	449.24	449.24	449.24	449.31	313.11	nd	nd
Hex ₁ HexNac-ol	534.29	534.29	534.29	534.29	384.15	384.17	384.25
HexA ₁ HexNac-ol	548.27	548.27	548.27	548.27	398.13	398.17	398.25
HexNac ₁ HexNac-ol	575.32	575.32	575.32	575.32	425.18	425.33	425.33
HexA ₁ HexNac ₁ Fuc-ol	722.36	722.38	722.37	722.35	544.19	nd	nd
Hex ₂ HexNac-ol	738.40	738.39	738.40	738.36	546.20	nd	nd
HexA ₁ Hex ₁ HexNac-ol	752.37	752.37	752.37	752.38	560.18	560.33	560.33
Hex ₁ HexNac ₁ HexNac-ol	779.42	779.45	779.42	779.58	587.23	587.33	587.33
HexA ₁ HexNac ₁ HexNac-ol	793.40	793.42	793.40	793.40	601.21	601.33	601.25
HexNac ₂ HexNac-ol	820.44	820.44	820.45	820.45	628.26	628.31	nd
S-HexA ₁ Hex ₁ HexNac-ol	840.29* [794.31]	nd	840.31 (794.32)	840.31 (794.32)	640.14	nd	nd
S-HexA ₁ HexNac ₁ HexNac-ol	881.32* [835.34]	881.47* [835.34]	881.32* [835.34]	881.33* [835.34]	681.17	681.17	681.25
Hex ₂ HexNac ₁ HexNac-ol	983.52	983.51	983.52	983.61	749.28	749.33	nd
HexA ₁ Hex ₁ HexNac ₁ HexNac-ol	997.49	997.48	997.51	997.71	763.26	763.33	763.33
S-Hex ₂ HexNac ₁ HexNac-ol	1071.44* [1025.46]	1071.60* [1024.42]	1071.45* [1025.46]	1071.40* [1025.41]	829.24	nd	nd
HexA ₂ Hex ₁ HexNac ₁ HexNac-ol	1215.57	1215.57	1215.58	1215.59	939.29	939.33	939.42
PE-HexA ₂ Hex ₁ HexNac ₁ HexNac-ol	1338.58 [1314.59]	nd	nd	nd	1062.3 (530.65)	530.75	530.83
S-HexA ₂ Hex ₁ HexNac ₂ HexNac-ol	1548.62* [1502.64]	nd	nd	nd	1222.33 (610.66)	610.75	610.83
PE-HexA ₂ Hex ₁ HexNac ₂ HexNac-ol	1583.71 [1559.71]	nd	nd	nd	1265.38 (632.19)	632.25	632.42
HexA ₃ Hex ₂ HexNac ₂ HexNac-ol	1882.88	nd	nd	nd	1480.46 (738.72)	739.83	739.83
HexA ₃ Hex ₂ HexNac ₃ HexNac-ol	2128.00	nd	nd	nd	1683.54 (841.27)	841.33	841.42
HexA ₄ Hex ₃ HexNac ₃ HexNac-ol	2550.18	nd	nd	nd	2021.62 (1010.31)	1010.42	1010.42
HexA ₄ Hex ₃ HexNac ₄ HexNac-ol	2795.31	nd	nd	nd	2224.7 (1111.85)	1111.75	1112.00
HexA ₅ Hex ₄ HexNac ₄ HexNac-ol	3217.49	nd	nd	nd	2562.79 (1280.89)	1280.92	1281.00

Figure 12. O-glycans identified in dipteran species by NSI or LC-MS

Permethylated O-glycans released by β -elimination of *Aedes*, *Anopheles* (Ag) or *Drosophila* (Dm) were analysed in positive mode by NSI-MS as sodiated species, except for sulphated species (S) which were detected either in the form [M-H+2Na]⁺ in positive ion mode (*) or [M-H]⁻ in negative mode as indicated in square brackets. Reduced, but not permethylated, O-glycans of *Aedes* and *Anopheles* were independently released by β -elimination and analysed by LC-MS in negative mode either in the form [M-H]⁻ or [M-2H]²⁻ as indicated in brackets. nd, not detected; S, sulphate; PE, phosphoethanolamine.

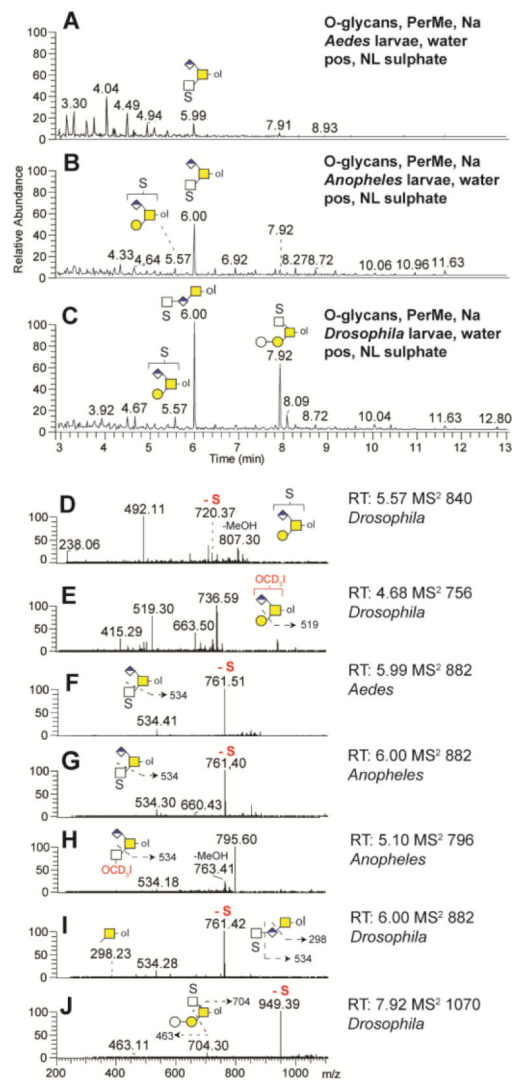


Figure 13. Sulphated O-glycan analyses by NSI-MS

Sulphated O-glycans were released by reductive β -elimination, permethylated, enriched in the water phase and analysed by NSI-MS in positive and negative ion mode. In automated positive ion mode TIM, spectra were filtered for neutral loss (NL) of sodiated sulphate moieties (120 Da) in *Aedes* (A), *Anopheles* (B) and *Drosophila* (C) larvae preparations. Three singly charged sulphated O-glycans were detected, two O-glycans carrying both sulphate and glucuronic acid modifications (m/z 840 and 882) and one sulphated O-glycan (m/z 1070). The automated TIM MS² spectra of selected O-glycan structures found in *Anopheles* (G, H), *Drosophila* (D, E, I, J) and *Aedes* (F) larvae are presented together with potential fragmentation schemes.