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

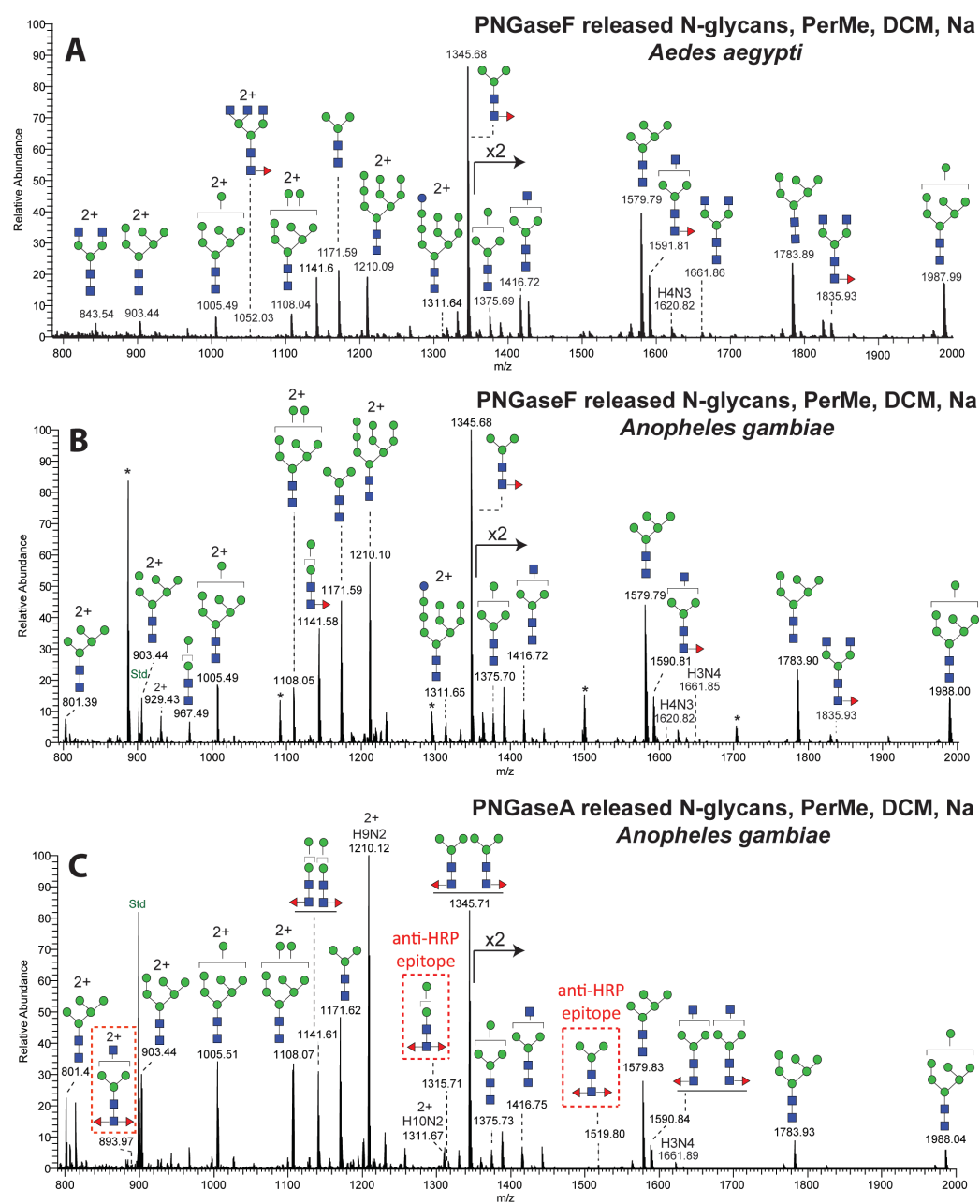Simone Kurz[1], Kazuhiro Aoki[2], Chunsheng Jin[3], Niclas G. Karlsson[3], Michael Tiemeyer[2], Iain B.H. Wilson[1] and Katharina Paschinger[1]

[1] Department für Chemie, Universität für Bodenkultur, 1190 Wien, Austria,
[2] Complex Carbohydrate Research Centre, University of Georgia, Athens, Georgia 30602, USA
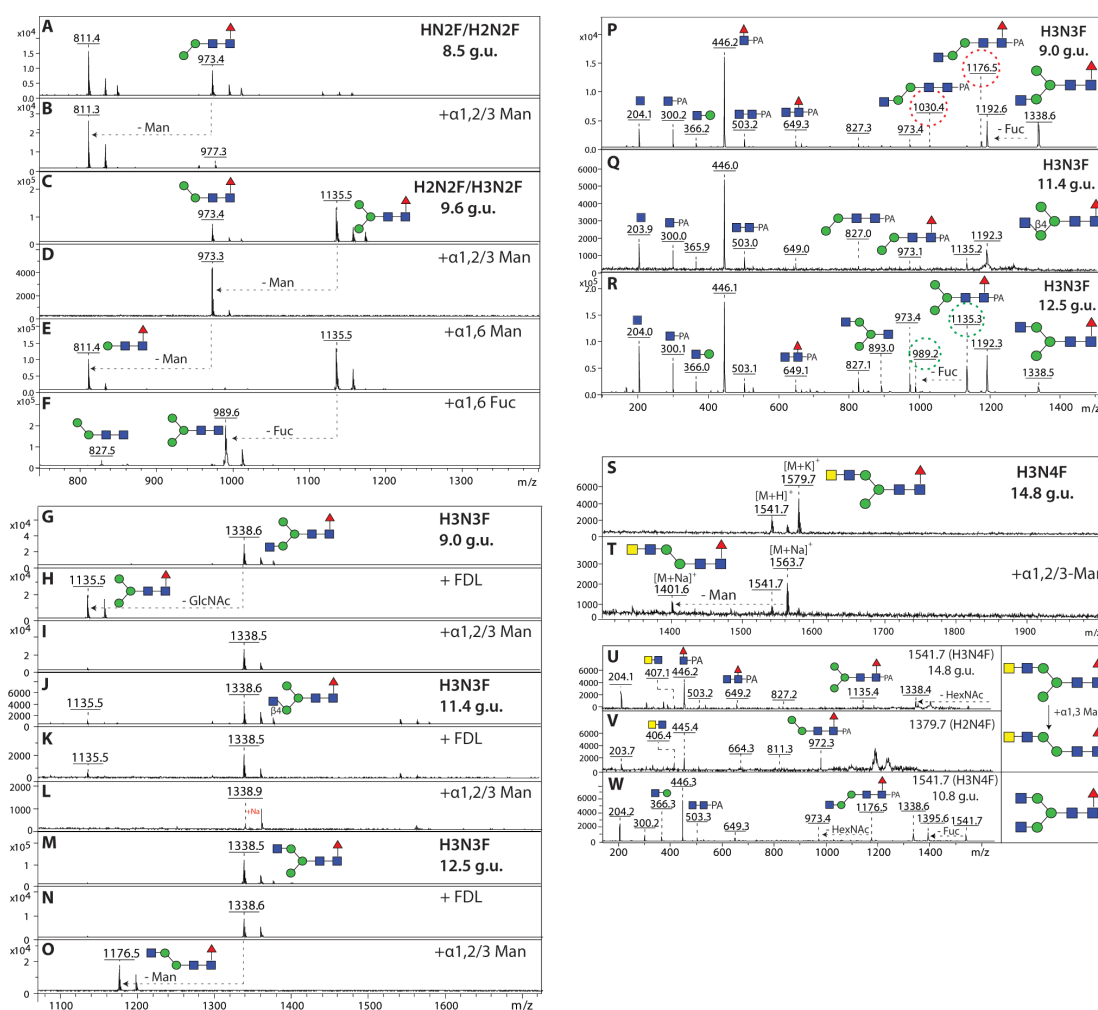[3] Department of Medical Biochemistry, University of Gothenburg, SE-405 30 Göteborg, Sweden
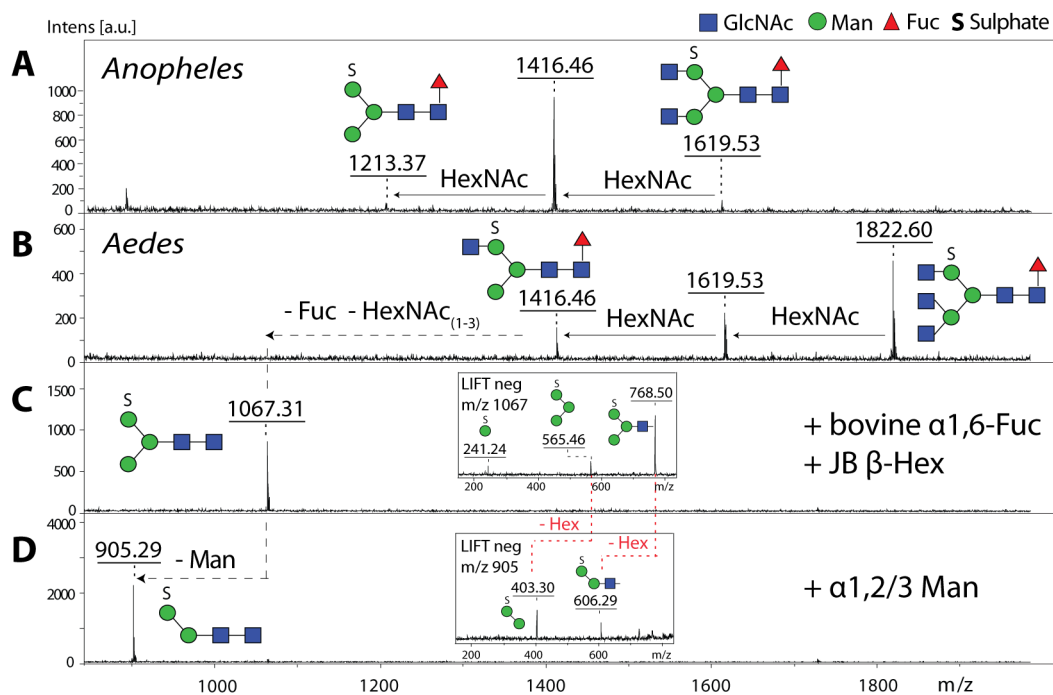
Supplementary Data

Supplementary Figure 1 - NSI-MS analysis of mosquito N-glycans: Non-sulphated N-glycans were enriched upon permethylation in the organic (DCM) phase and analysed by NSI-MS in the positive ion mode. PNGase F released N-glycans from *Aedes* (**A**) and *Anopheles* (**B**) larvae include a variety of oligomannosidic structures (doubly and singly charged) as well as paucimannosidic structures. Difucosylated N-glycans (dashed boxes) were detected upon PNGase A release from *Anopheles* (**C**) and are responsible for cross-reactivity towards anti-HRP. Peaks marked with asterisks correspond to a hexose series.
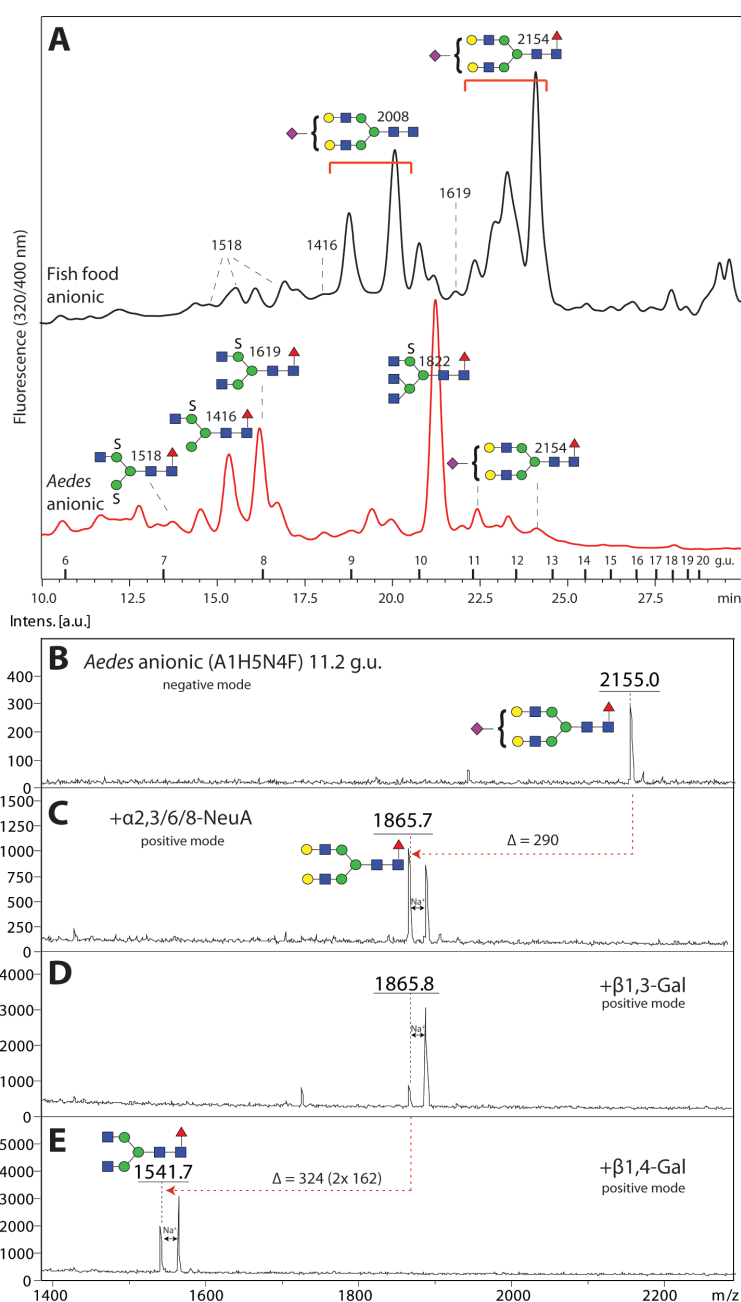
Supplementary Figure 2 - Selected digestion and MS/MS data for neutral *Aedes* N-glycans: (**A-F**) Exoglycosidase digestions of small paucimannosidic glycans indicating the isomeric status; MALDI-TOF MS of RP-amide fractions containing $Man_{1-3}GlcNAc_2Fuc_1$ (A, C) before and after specific $\alpha1,2/3-$ (B, D) or $\alpha1,6$-mannosidase (E) or bovine α-fucosidase (F) treatment. (**G-O**) Exoglycosidase digestions of glycans indicating the different positions of the single non-reducing *N*-acetylglucosamine on isomers of $Man_3GlcNAc_3Fuc_1$; the three isomers eluting 9, 11.4 and 12.5 g.u. (G, J, M) were incubated with either the specific FDL hexosaminidase (H, K, N) or $\alpha1,2/3$-mannosidase (I, L, O) prior to analysis by MALDI-TOF MS. (**P-R**) MS/MS spectra for the three $Man_3GlcNAc_3Fuc_1$ isomers with key diagnostic fragments highlighted (especially *m/z* 1030 and 1176 as compared to *m/z* 989 and 1135). (**S-V**) Specific mannosidase digestion of a LacdiNAc-containing $Hex_3HexNAc_4Fuc_1$ structure eluting at 14.8 g.u.; a partial sensitivity of the glycan to $\alpha1,2/3$-mannosidase (compare S and T) and the MS/MS spectra as well as the late elution time suggest that the LacdiNAc modification is on the $\alpha1,6$-arm. (**W**) The MS/MS fragmentation of the 10.8 g.u. $Hex_3HexNAc_4Fuc_1$ structure differs from that for the LacdiNAc-containing isobaric form. The glycans were predominantly detected in their protonated forms (unless otherwise indicated).
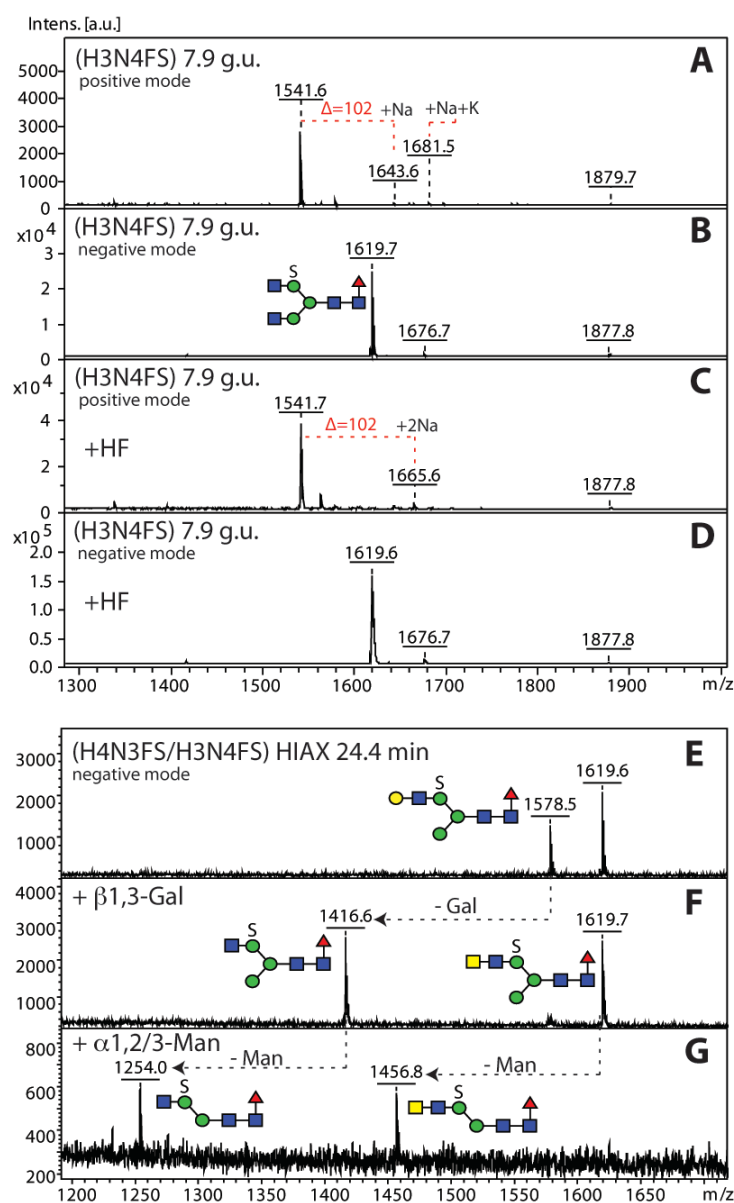

3

Supplementary Figure 3 - Structural determination of the major anionic mosquito N-glycans: Negative ion mode MALDI-TOF/TOF MS spectra of the anionic N-glycan pools of *Anopheles* (**A**) and *Aedes* (**B-D**) before (**A, B**) and after (**C-D**) exoglycosidase treatments to determine the position of the anionic modification (sulphate). Bovine α-fucosidase and jack bean β-hexosaminidase were able to liberate the core fucose and up to three N-acetylglucosamine residues from the parent glycan species $Hex_3HexNAc_{2-5}Fuc_1[SO_3]_1$ to the product glycan species of $Hex_3HexNAc_2[SO_3]_1$ (*m/z* 1067 as [M-H]$^-$; **C**). The resulting product was then incubated with α1,3-specific mannosidase (**D**) which was able to liberate one mannose residue indicating the sulphate group to be associated on the α1,6 mannose; jack bean mannosidase also removed only one mannose (**not shown**). MS/MS fragmentation of the $Man_3GlcNAc_2[SO_3]_1$ substrate and the α1,3-mannosidase digestion product are shown in insets and are compatible with sulphation of mannose.
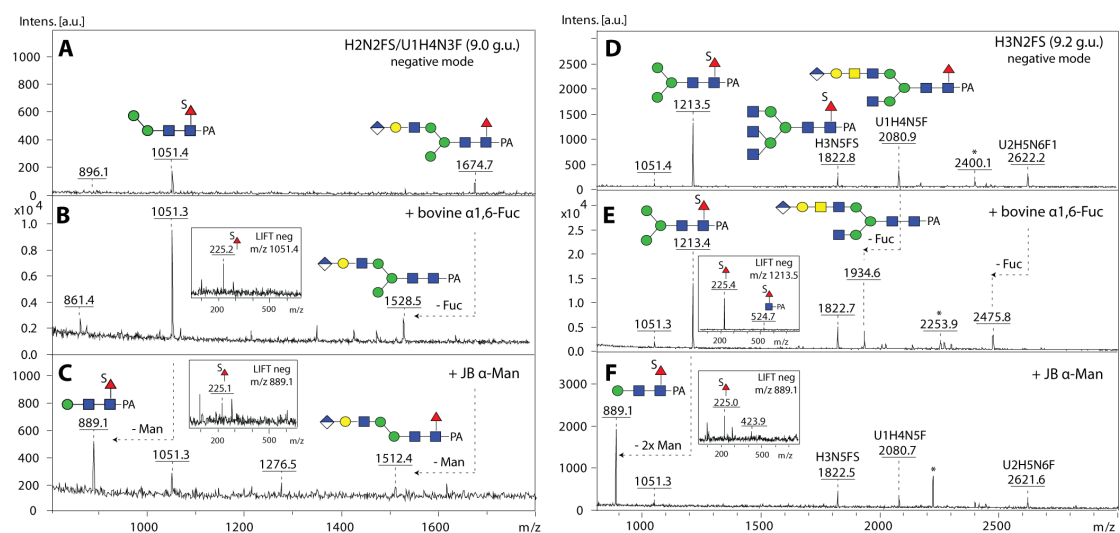
Supplementary Figure 4 – Comparison of anionic N-glycans present in *Aedes* larvae and the corresponding food control. (**A**) RP-amide HPLC of the anionic pools of PNGase F-released N-glycans from *Aedes* larvae (see also Figure 5A of the main text) and the corresponding food control (Vitakraft Premium VITA 'Flockenfutter für alle Zierfische'). Whereas the sulphated glycans differ in terms of elution (structures only shown for those from the mosquito; otherwise the *m/z* as [M-H]⁻ or [M-2H+Na]⁻ is indicated), sialylated glycans are present in a broad region (11-13 g.u.) as dominant components of the food control; these co-elute with trace amounts of sialylated glycans in the *Aedes* sample (11.2 and 12.5 g.u.). (**B-E**) Exoglycosidase treatments of the 11.2 g.u. sialylated N-glycans present in the *Aedes* sample show that the underlying structure contains β1,4-galactose. Due to the co-elution and the type of galactosylation, it is unproven as to whether the sialylated glycans in the *Aedes* sample originate from the larvae or the food source.
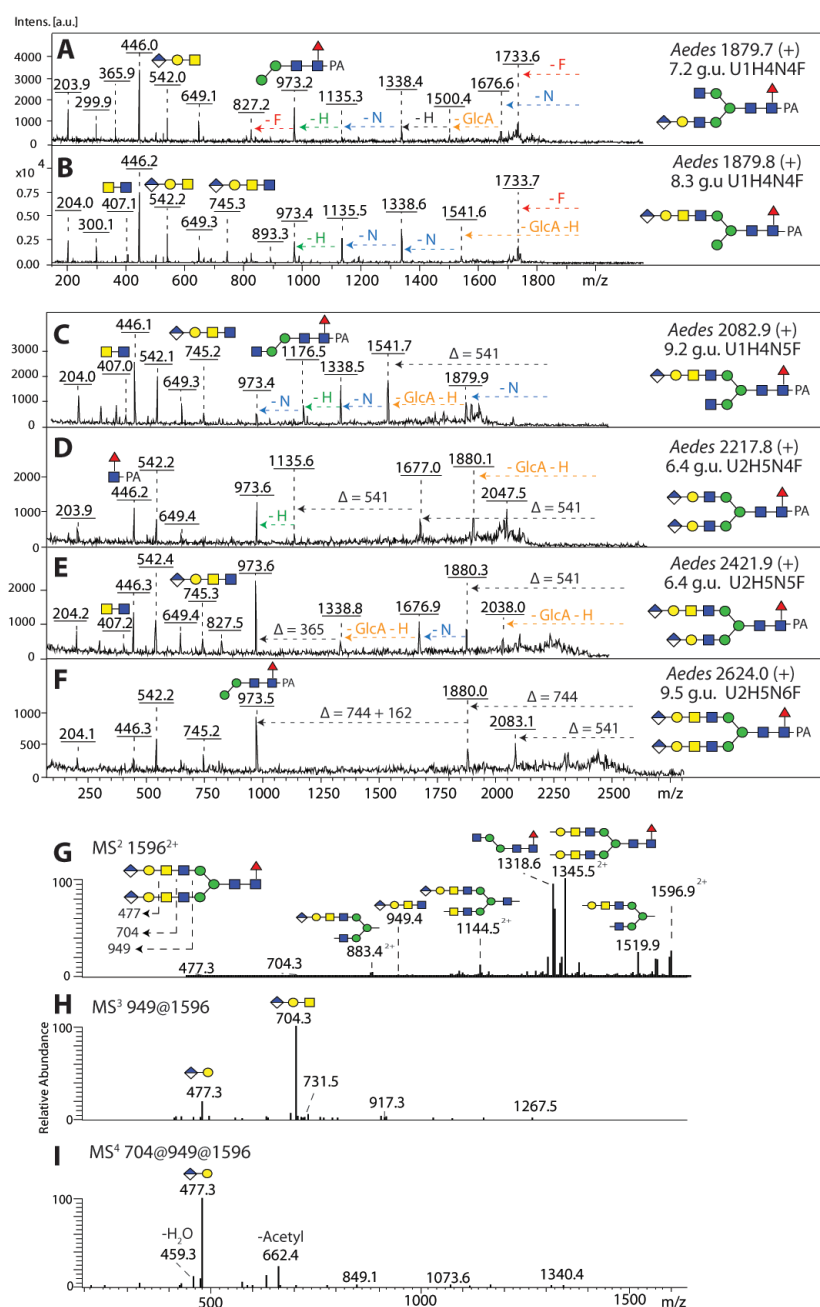
Supplementary Figure 5 - Further incubations of mosquito sulphated N-glycans with chemical or enzymatic reagents. (**A-D**) Incubation with hydrofluoric acid of the *Aedes* 7.9 g.u. anionic fraction containing $Hex_3HexNAc_4Fuc_1[SO_3]_1$ and traces of $HexA_1Hex_4HexNAc_{3-4}Fuc_1$ resulted in no alteration of the negative and positive ion mode spectra, which is a further indication for sulphation rather than phosphorylation as is the low level of the $[M+Na]^+$ ions for the intact glycan in positive mode. (**E-G**) Incubation with *Xanthomonas* β1,3-specific galactosidase followed by *Xanthomonas* α1,2/3-mannosidase of an *Anopheles* anionic fraction (HIAX, 24.4 mins, see Figure 5C of the main text) containing $Hex_{4/3}HexNAc_{3/4}Fuc_1[SO_3]_1$; the data indicate that the sulphated *m/z* 1578 glycan contains a β1,3-linked galactose and further show that both the *m/z* 1578 and 1619 glycans have the sulphate and galactose/N-acetylgalactosamine modifications on the upper α1,6-arm.
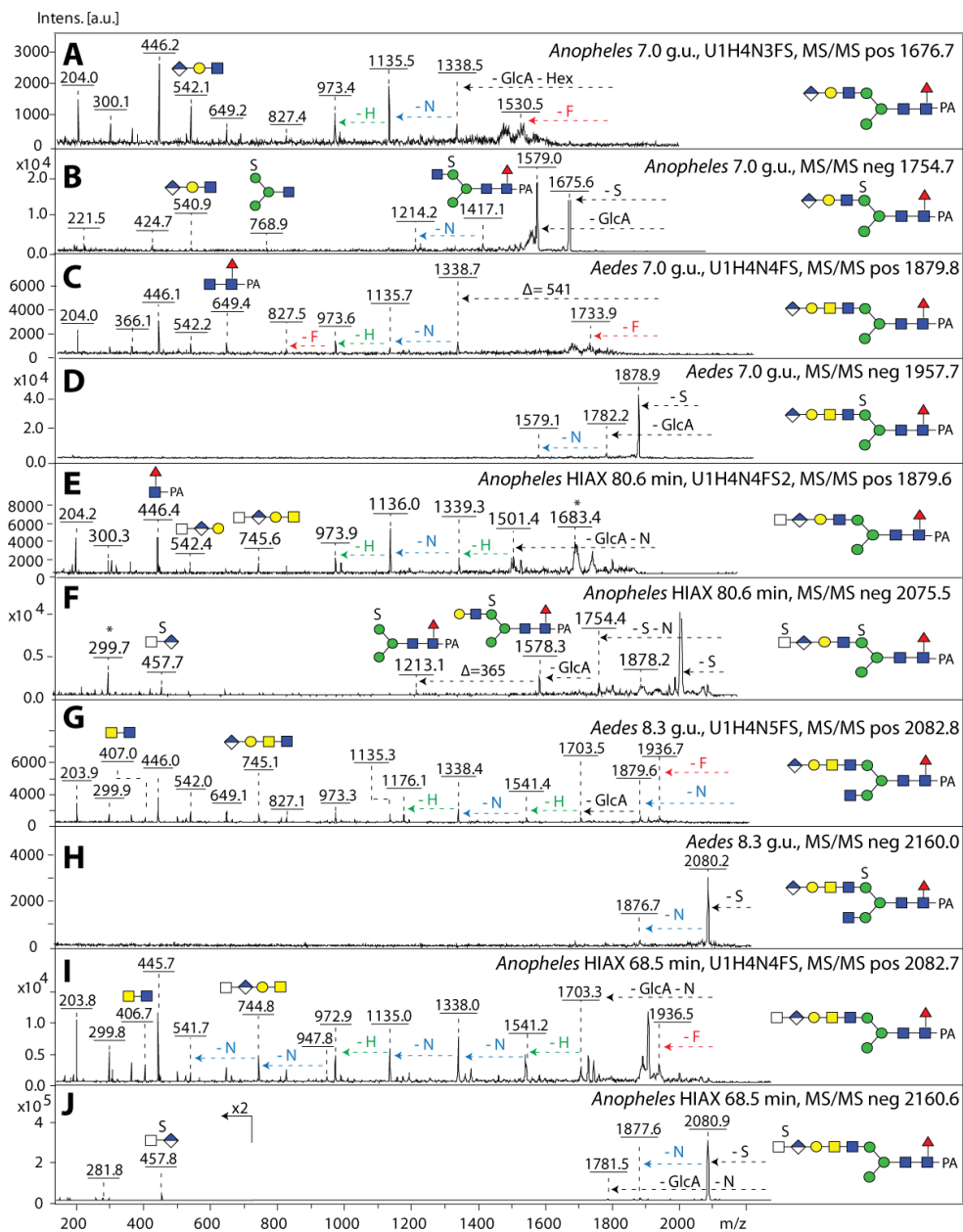
Supplementary Figure 6 – Evidence for sulphation of core α1,6-fucose. The *Aedes* anionic 9.0 and 9.2 g.u. fractions (eluting before the major triantennary sulphated N-glycan; see Figure 5 of the main text) were incubated with either bovine kidney α-fucosidase or jack bean α-mannosidase prior to MALDI-TOF MS (compare **A** with **B** and **C** or **D** with **E** and **F**). Whereas fucosidase treatment of the major triantennary glycan resulted in loss of fucose (see Figure 4 in the main text) as well as loss of the fucose from glucuronylated glycans in the 9.0 and 9.2 g.u. fractions, the sulphated N-glycans in these fractions (*m/z* 1051, 1213 and 1822) were resistant and possessed *m/z* 225 fragments (see insets); however, the *m/z* 1051 and 1213 glycans were fully sensitive to mannosidase treatment unlike the complex glycans pre-treated with β-hexosaminidase (see Supplementary Figure 3). In combination with the NSI-MS[n] data showing an *m/z* 516 fragment (compatible with the permethylated composition $GlcNAc_1Fuc_1[SO_3]_1$; Figure 8 of the main text), we conclude that a low level of sulphation of fucose is present.
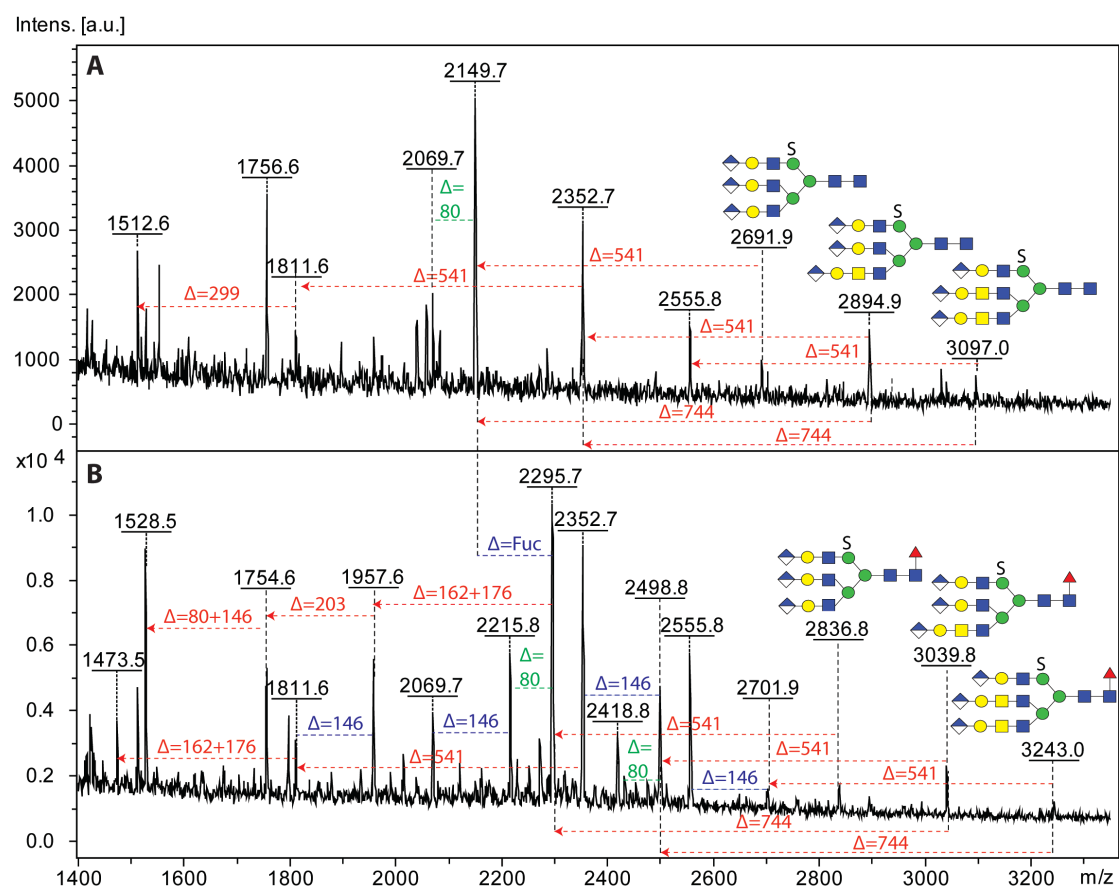
Supplementary Figure 7 - MS/MS of selected mono- and diglucuronylated N-glycans: Example positive ion mode MALDI-TOF MS/MS of glucuronylated N-glycans are shown for (**A-B**) two isomers of $Hex_4HexNAc_4Fuc_1GlcA_1$ present in two different *Aedes* RP-amide fractions as well as for (**C-F**) four different mono- or diglucuronylated glycans in three *Aedes* RP-amide fractions. (**G-I**) Example positive ion mode NSI-MS^n of a permethylated, glucuronylated, doubly charged N-glycan species enriched in the DCM phase. In the MS^2 fragmentation of m/z 1596^{2+} (**G**), a minor singly charged fragment ion of m/z 949 ($HexA_1Hex_1HexNAc_2$) was detected and selected for MS^3 fragmentation. A major fragment ion of m/z 704 ($HexA_1Hex_1HexNAc_1$) was detected upon MS^3 fragmentation (**H**) and further selected for MS^4 fragmentation. MS^4 of m/z 704 yield a major fragment ion of m/z 477 ($HexA_1Hex_1$) and minor fragment ions indicative for the loss of water (m/z 459) and the acetyl group (m/z 662) from the parent ion selected for MS^4 (**I**).
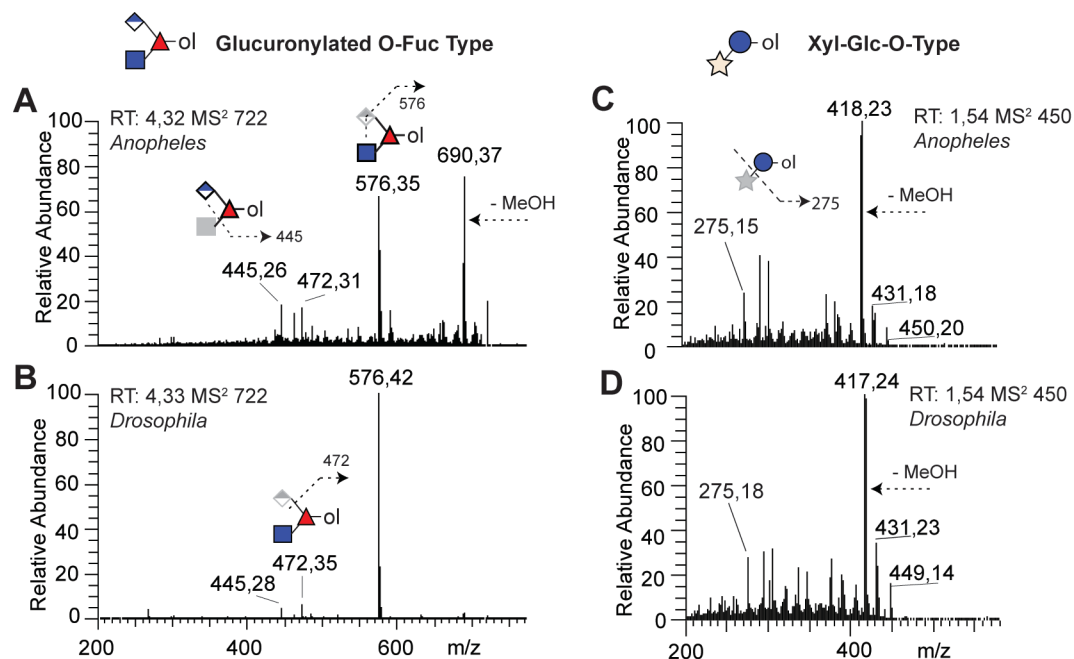
Supplementary Figure 8 – MS/MS analysis of glycans containing both glucuronic acid and sulphate residues: Example MALDI-TOF MS/MS of N-glycans carrying both glucuronic acid and sulphate are shown (alternating positive and negative ion modes) for glycans with compositions of $Hex_4HexNAc_{3-5}Fuc_1HexA_1[SO_3]_1$ present in different *Anopheles* (**A/B**) and *Aedes* (**C/D** and **G/H**) RP-amide fractions or in *Anopheles* HIAX fractions (**E/F** and **I/J**). The glycan structures shown are proposed, but alternative interpretations or the presence of mixed isomers are possible; two potential isomeric variations of $Hex_4HexNAc_5Fuc_1HexA_1[SO_3]_1$ are shown with the position of the sulphate in the two *Anopheles* HIAX fractions being ambiguous, but the *m/z* 458 fragment (apparent in some other spectra of other structures not shown) in negative mode suggestive that the sulphate is associated with the $HexNAc_1HexA_1$ motif rather than mannose and is also similar to a motif found on some O-glycans (Supplementary Figure 12). Asterisks indicate fragments of unknown origin.

Supplementary Figure 9 - MALDI-TOF MS analysis of early-eluting *Anopheles* anionic glycan RP-HPLC fractions. (**A**) Negative ion mode MALDI-TOF MS of the 3.1 g.u. fraction appeared to result in in-source fragmentation of triantennary tetra-anionic N-glycans. A potential interpretation of the mass spectrum is shown (arrows) in terms of the loss of glucuronylated antennae (*Δm/z* 542 and 745) as well as of the core GlcNAc-PA (*Δm/z* 299). (**B**) Negative ion mode MALDI-TOF MS of the 4.7 g.u. fraction yields a range of presumed in-source fragments of core fucosylated triantennary tetra-anionic N-glycans; losses of fucose, sulphate and glucuronylated antennae as well as serial losses of glucuronic acid, *N*-acetylhexosamine and hexose are indicated. The apparent instability of the triantennary structures is not obvious with hybrid and biantennary glycans with one or two glucuronic modifications (see, e.g., Figure 9 in the main text). The spectra were obtained using an UltrafleXtreme instrument; the analysis on the Autoflex Speed resulted only in detection of the *m/z* 2149, 2215, 2295 and 2352 ions. Possible structures are only shown for the presumed triantennary parent ions.

Supplementary Figure 10 – O-fucosyl- and O-glucosyl O-glycans: After reductive β-elimination, permethylation and phase partition, O-Fuc and O-Hex glycans were enriched in the DCM phase and subjected to NSI-MS and MS^2 analysis. Positive ion mode MS^2 fragmentation of glucuronylated O-Fuc glycans (HexNAc$_1$HexA$_1$dHex$_1$-ol, m/z 722) demonstrated Z-ions at m/z 445 and 472 corresponding to the loss of terminal HexNAc and terminal HexA from a branched O-glycan species in both *Anopheles* (**A**) and *Drosophila* (**B**) larvae. In fragmentation of the O-Hex glycan species (Pent$_1$Hex$_1$-ol, m/z 450), Y-ions at m/z 275 indicated the presence of single substituted Hex$_1$-ol in both dipteran species (**C, D**).

Supplementary Figure 11 – NSI-MS analysis of O-glycans from *Drosophila* cell lines: 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, indicated by -S) in *Drosophila* Schneider S2 (**A**) and *Drosophila* BG2-c6 (**B**) cell line preparations. Two low abundant singly charged sulphated O-glycans were detected carrying both sulphate and glucuronic acid modifications (*m/z* 840 and 882). The automated TIM MS^2 spectra of selected O-glycan structures found in *Drosophila* Schneider S2 (**C**) and BG2-c6 (**D**) cells are presented together with potential fragmentation schemes.

Supplementary Figure 12 – LC-MS/MS analysis of extended O-glycans: Examples of MS/MS spectra of O-glycans predicted to contain two glucuronic acid residues are shown; the glycan in panel B contains sulphate (S) and those in panels D and E contain phosphoethanolamine (PE). The terminal $HexNAc_1HexA_1Hex_1$ unit with and without sulphate may also feature on some N-glycans (see Supplementary Figure 8).