

Glycan Complexity and Heterogeneity of Glycoproteins in Somatic Extracts and Secretome of the Infective Stage of the Helminth *Fasciola hepatica*

Authors

Carolina De Marco Verissimo, Krystyna Cwiklinski, Jonas Nilsson, Ekaterina Mirgorodskaya, Chunsheng Jin, Niclas G. Karlsson, and John P. Dalton

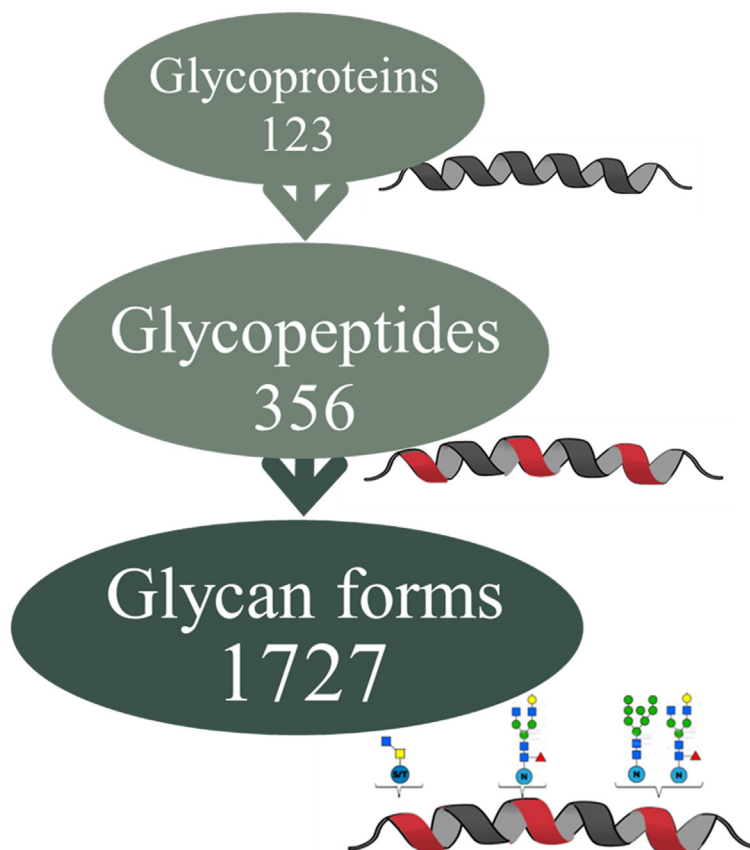
Correspondence

carolina.verissimo@universityofgalway.ie

In Brief

Fasciola hepatica, a worm parasite of humans and livestock, employs surface and secreted molecules to escape the host's attack and establish infection. While the role of proteins during infection is known, we lack understanding about the sugars (glycans) associated with them. Our analyses unveiled that a diverse range of glycans can modify the same parasite protein, generating variability and a complex relationship with the host. These new insights into *F. hepatica*'s glycan biology will improve diagnostics and parasite-targeted vaccines.

Graphical Abstract



Highlights

- 123 glycoproteins of the invasive stage of *Fasciola hepatica* were characterized.
- 71 glycoproteins are secreted-excreted into the host space.
- 356 individual glycosylated sites were mapped to their respective glycoproteins.
- 56 *N*-glycan and 16 *O*-glycan structures modify *F. hepatica* NEJs proteins.
- There is major heterogeneity within the glycosylation of NEJs' glycoproteins.



Glycan Complexity and Heterogeneity of Glycoproteins in Somatic Extracts and Secretome of the Infective Stage of the Helminth *Fasciola hepatica*

Carolina De Marco Verissimo^{1,*}, Krystyna Cwiklinski^{1,2}, Jonas Nilsson³, Ekaterina Mirgorodskaya³, Chunsheng Jin³, Niclas G. Karlsson⁴, and John P. Dalton¹

Fasciola hepatica is a global helminth parasite of humans and their livestock. The invasive stage of the parasite, the newly excysted juvenile (NEJs), relies on glycosylated excreted-secreted (ES) products and surface/somatic molecules to interact with host cells and tissues and to evade the host's immune responses, such as disarming complement and shedding bound antibody. While -omics technologies have generated extensive databases of NEJs' proteins and their expression, detailed knowledge of the glycosylation of proteins is still lacking. Here, we employed glycan, glycopeptide, and proteomic analyses to determine the glycan profile of proteins within the NEJs' somatic (Som) and ES extracts. These analyses characterized 123 NEJ glycoproteins, 71 of which are secreted proteins, and allowed us to map 356 glycopeptides and their associated 1690 *N*-glycan and 37 *O*-glycan forms to their respective proteins. We discovered abundant micro-heterogeneity in the glycosylation of individual glycosites and between different sites of multi-glycosylated proteins. The global heterogeneity across NEJs' glycoproteome was refined to 53 *N*-glycan and 16 *O*-glycan structures, ranging from highly truncated paucimannosidic structures to complex glycans carrying multiple phosphorylcholine (PC) residues, and included various unassigned structures due to unique linkages, particularly in pentosylated *O*-glycans. Such exclusive glycans decorate some well-known secreted molecules involved in host invasion, including cathepsin B and L peptidases, and a variety of membrane-bound glycoproteins, suggesting that they participate in host interactions. Our findings show that *F. hepatica* NEJs generate exceptional protein variability via glycosylation, suggesting that their molecular portfolio that communicates with the host is far more complex than previously anticipated by transcriptomic and proteomic analyses. This study opens many avenues to understand

the glycan biology of *F. hepatica* throughout its life-stages, as well as other helminth parasites, and allows us to probe the glycosylation of individual NEJs proteins in the search for innovative diagnostics and vaccines against fascioliasis.

Proteins undergo post-translational glycosylation, which involves the addition of *N*- and *O*-linked glycans in the endoplasmic reticulum (ER) and Golgi apparatus. This process enhances the biological significance and complexity of these molecules by changing their biochemical and biological properties (1). Glycans exhibit enormous variations in composition, structure, arrangement, and modifications. Such heterogeneity has the potential to modify and diversify protein functions, offering numerous opportunities for interactions with other molecules (2, 3). Glycosylation can also impact the presentation and properties of proteins on cell surfaces or organelles, as well as alter their antigenicity/immunogenicity, redesigning how they are recognized and influencing other cells, particularly those of the immune system. Glycans can elaborate the functionality of glycoproteins within the same species or redefine their activity when encountered in different species (4).

Parasites use glycans to decorate mainly proteins on their surface and those associated with their secretions (5–7), and these are critical to their ongoing communication with the host. Since helminth parasites go through a number of developmental stages during their migration, growth, and maturation within different tissues of the mammalian host, glycans become essential in the continual conversation between them and the neighboring tissues, and in inducing immunopathogenesis, progression of disease, and development of immunity and immune tolerance (5, 8). Every parasite

From the ¹Molecular Parasitology Lab (MPL) - Centre for One Health and Ryan Institute, School of Natural Science, National University of Ireland Galway, Galway, Republic of Ireland; ²Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom; ³Proteomics Core Facility, Sahlgrenska Academy of Science, University of Gothenburg, Gothenburg, Sweden.; ⁴Department of Life Science and Health, Faculty of Health Science, Oslo Metropolitan University, Oslo, Norway

*For correspondence: Carolina De Marco Verissimo, carolina.verissimo@universityofgalway.ie.

appears to generate a distinct array of glycoconjugates, encompassing *N*-/*O*-glycoproteins and glycolipids modified by both unique and conserved glycan motifs (9). These glycoconjugates are crucial for the parasite's specific adaptation and survival strategies and, therefore, unique glycan moieties expressed by *Schistosoma* spp. (10), *Trichinella spiralis* (6, 11), *Acanthocheilonema viteae* (12), for example, have been largely studied in the context of diagnosis and vaccine development.

The zoonotic helminth *Fasciola hepatica* is endemic in more than 70 countries where it infects over 2.6 million people and their livestock, leading to annual losses of ~€2.5 billion to the agriculture sector (13–16). Infection by *F. hepatica* is initiated following ingestion of vegetation contaminated with an encysted metacercarial stage from which newly excysted juveniles (NEJs) emerge in the intestine (17). The microscopic parasites rapidly traverse the intestinal wall before migrating through the liver tissue and entering the bile ducts, where adult parasites reside and produce progeny for many years (18). Advances in -omics technologies have provided large genome, transcriptome and proteome datasets and have been pivotal for the extensive investigation of proteins in the liver fluke biology, most particularly of those involved in NEJs tissue invasion (19). The data collected show that NEJs express >8000 genes during the first few days migrating to the liver and secreting a complex mixture of over 100 different proteins with various biochemical and immunomodulatory properties (20, 21). However, there is a dearth of information regarding the glycans expressed by liver fluke molecules, how these are used to modify the parasite proteins, and the impact they have on the host.

Reports from over 5 decades show that *F. hepatica* express glycans in their tissues, particularly within the thick glycocalyx layer that covers their entire surface (22–27). Ravidá *et al.* (28) and Garcia-Campos *et al.* (25) provided the first description of glycan structures for glycoproteins in extracts from the tegument of adults and immature parasite stages and revealed the presence of truncated and high-mannose sugars. Lectin array studies also indicated that a range of sugars are exposed on the adult parasite tegument and are part of the surface cargo of released extracellular vesicles (EVs), which could be involved in modulating innate immune cell phenotypes, eliciting immune tolerance, and/or inducing anergy in infected mice (29). Taking a different approach, McVeigh *et al.* (30) used *in silico* analysis to identify 149 orthologs of human genes involved in protein *N*- or *O*-glycosylation and predicted that the *F. hepatica* glycome contains principally small glycans of mannosidic nature.

In the present study, we exploited up-dated mass spectrometry and bioinformatics methods for glycan and glycoprotein analyses, together with extensive proteome sequence databases, to probe and characterize the glycome of the infective *F. hepatica* NEJs stage. Furthermore, we delved deeper into the glycome and proteome of the NEJs somatic extract and secretome to map the glycosylated sites of

individual glycoproteins and their associated *N*- and *O*-glycan structures. These studies resulted in a detailed and comprehensive picture of each glycoprotein and uncovered abundant glycan heterogeneity within the glycosites that generate vast complexity within these extracts. This in-depth analysis of glycosylation in *F. hepatica* complements and extends the vast publicly-available liver fluke -omics datasets and will support future advances in our knowledge of liver fluke biology, particularly in the area of parasite invasion, immune evasion, and pathogenicity, which is critical for the advancement for the development of diagnostics and vaccines.

EXPERIMENTAL PROCEDURES

Experimental Design and Statistical Rationale

Glycomics and glycopeptide analyses were carried out on two commercial *F. hepatica* isolates (Italian and Aberystwyth; both susceptible to triclabendazole). *N*- and *O*-glycan analyses were carried out using 200 µg for both the somatic (Som) and ES extracts in triplicate derived from replicates of 2000 newly excysted juveniles (NEJs). Glycan abundance was calculated as the intensity derived from the area of under the curve (AUC) normalized to the total AUC (calculated for all glycans in a sample), represented as a percentage.

N-glycopeptide and *O*-glycopeptide analyses were carried out in triplicate derived from replicates NEJs and using 500 µg for the Som extracts and 80 µg for the ES extracts. Post-glycopeptide enrichment, three and two injections were acquired for in-depth *N*- and *O*-glycoproteomics, respectively. We used the glycopeptide data to complement and curate the glycomics data obtained, and to ensure that only glycans associated to *F. hepatica* proteins would be included in the comparison. The number of the glycopeptides and glycan forms associated with each glycoprotein was determined by manual curation. Lectin labeling experiments were carried out on 20 NEJs per treatment and the respective results were reflective of at least three individual NEJs analyzed for each lectin.

Parasites and Extracts

F. hepatica NEJs were obtained by excysting metacercariae (Italian and Aberystwyth isolates, Ridgeway Research Ltd, UK) as previously described by Robinson *et al.* (21). To obtain the excretory/secretory products (ES), *F. hepatica* NEJs were cultured in RPMI-1640 media (Thermo Fisher Scientific) supplemented with 30 mM HEPES (Thermo Fisher Scientific), 0.1% glucose and 50 µg/ml gentamycin, at 37 °C with 5% CO₂, for 24 h. The culture media was collected, centrifuged, and concentrated using Amicon Ultra 3 kDa columns (Merck Millipore). The protein concentration was determined by Bradford Protein Assay (Bio-Rad), and the samples were stored at –70 °C until use (27).

NEJs were recovered after 24 h of culture, washed five times in Dulbecco's phosphate-buffered saline (DPBS) (Thermo Fisher Scientific) and used to produce the somatic extract (Som) and for lectin labeling experiments. For the somatic extract, the NEJs (replicates of 2000) were freeze-dried and extracted in lysis buffer (7 M urea, 2 M thiourea, 40 mM Tris, and 1% CHAPS) by homogenization using a glass homogenizer on ice, for 20 min, followed by overnight incubation at 4 °C on a rotary shaker (10 rpm) and centrifugation at 10,000g for 15 min. The supernatant was collected and protein concentration was determined using the BCA assay. The samples were subsequently reduced with 5 mM dithiothreitol (DTT) at 56 °C for 35 min and alkylated with 11 mM iodoacetamide (IAM) at room temperature (RT) for 20 min, in the dark.

Release and Analysis of the N- and O-Glycans From *F. hepatica* NEJs' Glycoproteins

F. hepatica NEJs Som and ES extracts (200 µg, in triplicate) were dot blotted onto polyvinylidene difluoride (PVDF) membrane. The membrane was blocked with 200 µl 1% PVP (polyvinyl pyrrolidone 40,000, w/w) in 50% methanol with shaking for 10 min and washed with 500 µl water three times for 5 min each. The N-glycans were released from the glycoproteins by the addition of PNGase F (CarboClip, Asparia Glycomics) in 50 mM ammonium acetate (pH 8.4) at 37 °C for 24 h. The released N-glycans were reduced by incubation with 0.5 M NaBH₄ in 20 mM NaOH at 50 °C for 16 h, the reaction was quenched by acetic acid and the released N-glycans desalted with AG50WX8 cation exchange beads (Bio-Rad). Borate complexes were removed by repeated addition/evaporation with methanol. After washing of the membrane, the O-glycans were released from the proteins by reductive β-elimination by incubating with 0.5 M NaBH₄ and 50 mM NaOH at 50 °C for 16 h, and the O-glycans were desalted as described for the N-glycans (31).

The samples containing the released N- and O-glycans were re-suspended in water and analyzed by liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS). The oligosaccharides were separated on a column (10 cm × 250 µm) packed in-house with 5 µm porous graphite particles (Hypercarb, Thermo-Hypersil). The oligosaccharides were injected into the column and eluted with an acetonitrile (ACN) gradient (Buffer A, 10 mM ammonium bicarbonate; Buffer B, 10 mM ammonium bicarbonate in 80% ACN). The gradient (0–45%, Buffer B) was eluted for 46 min, followed by a wash step with 100% Buffer B, and equilibrated with Buffer A for the next 24 min. A 40 cm × 50 µm i.d. fused silica capillary was used as a transfer line to the ion source.

The samples were analyzed in negative ion mode on an LTQ linear ion trap mass spectrometer (Thermo Electron), with an IonMax standard ESI source equipped with a stainless-steel needle kept at –3.5 kV. Compressed air was used as nebulizer gas. The heated capillary was kept at 270 °C, and the capillary voltage was –50 kV. The mass ranges were defined depending on the specific structure to be analyzed. Data acquisition and processing were conducted with Xcalibur software (Version 2.0.7, Thermo Fisher Scientific). Glycans were identified from their MS/MS spectra by manual annotation. Glycan structural characterization was based on the diagnostic fragment ions (32).

Preparation, Enrichment, and Analysis of N- and O-Glycopeptides From ES and Somatic Extracts of *F. hepatica* NEJs

Sample Preparation—For the proteomics and glycoproteomics analyses, NEJs ES and Som samples were processed using the modified filter-aided sample preparation (FASP) method (33). Freeze-dried NEJs were homogenized using the lysis matrix B on the FastPrep-24 instrument (MP Biomedicals) in lysis buffer (50 mM triethylammonium bicarbonate (TEAB), 2% sodium dodecyl sulfate (SDS)), and protein concentration determined using the BCA assay. NEJs Som (500 µg) and ES (80 µg) extracts were reduced with 100 mM DTT at 60 °C for 30 min, and transferred to 30 kDa MWCO Pall Nanosep centrifugation filters (Sigma-Aldrich). Two individual filters were used for Som extract, resulting in a total amount of 250 µg per filter. All filters were washed several times with 8 M urea and once with digestion buffer (DB, 50 mM TEAB, 0.5% sodium deoxycholate (SDC)) prior to alkylation with 18 mM IAM for 30 min at room temperature in the dark. Samples were digested with trypsin (Pierce MS grade Trypsin, Thermo Fisher Scientific, ratio 1:100) at 37 °C overnight, followed by an extra trypsin addition (ratio 1:100) and 2 h reaction at 37 °C. The resulting proteolytic peptides were collected by centrifugation and filters were additionally washed with 50 mM TEAB to

ensure all material eluted from the filters. SDC was removed by acidification with 10% trifluoroacetic acid (TFA) and subsequent centrifugation for 20 min at 21,000g. The supernatant was further purified using Pierce peptide desalting spin columns (Thermo Fisher Scientific) according to the manufacturer's instructions. The bound peptides were eluted with 80% ACN in 1% TFA and directly loaded onto in-house prepared hydrophilic interaction liquid chromatography (HILIC) spin columns (20 mg ZIC-HILIC resins per column, SeQuant # 2942) for glycopeptide enrichment. Both HILIC flow-through and eluates were collected and evaporated prior to further analysis. For glycoproteomic analysis, the HILIC eluate was reconstituted in 2% ACN and 0.2% formic acid (FA). Once the global N-glycoproteomic analysis was finalized, the remaining material was lyophilized and subjected to PNGase F digestion prior to O-glycoproteomic analysis.

NanoLC-MS/MS Analysis—All preparations were analyzed on a QExactive HF mass spectrometer interfaced with the Easy-nLC1200 liquid chromatography system (Thermo Fisher Scientific). Peptides were trapped on an Acclaim Pepmap 100 C18 trap column (100 µm × 2 cm, particle size 5 µm, Thermo Fischer Scientific) and separated on an in-house packed analytical column (75 µm × 300 mm, particle size 3 µm, Reprosil-Pur C18, Dr Maisch). Solvent A was 0.2% FA and solvent B was 80% ACN in 0.2% FA.

The HILIC eluate preparations, enriched for glycopeptides, were analyzed using a 90-min gradient: from 3% to 35% solvent B over 75 min, followed by an increase to 100% B for 5 min at a flow of 300 nl/min. To facilitate glycosylated peptide characterization, triplicate injections were acquired, all with the precursor detection in the 600 to 2000 m/z range at a resolution of 120,000 with different settings for the normalized higher-energy collision dissociation (HCD) energies. MS/MS analysis was performed in a data-dependent mode, where the ten most intense ions with charge states 2 to 6 were selected for fragmentation using HCD at collision energy settings of 20, 28, or 38. The isolation window was set to 3 m/z and dynamic exclusion set to 10 ppm for 20 s.

For O-glycosylated peptides analysis, the PNGase F-treated preparations were analyzed using the same LC gradient. Double injections were acquired, both with the precursor detection in the 375 to 1500 m/z range at a resolution of 120,000. MS/MS analysis was performed in a data-dependent mode. The ten most intense ions with charge states 2 to 5 were selected for fragmentation using HCD at collision energy settings of either 20 or 28. The isolation window was set to 1.2 m/z and dynamic exclusion was set to 10 ppm for 20 s.

Glycoproteomic Data Analysis

The Byonic software (v2.10.21, Protein Metrics) was used as a node within Proteome Discoverer (v2.4, Thermo Fisher Scientific) for the identification of N- and O-glycopeptides. The N-glycan database contained 200 compositions including oligomannose, sialylated complex biantennary as well as phosphatidylethanolamine (PE)- and phosphorylcholine (PC)-containing structures (Supplementary Material 1). The O-glycan database contained 53 compositions ranging from HexNAc (1) to HexNAc (8)Hex (8), including a number of pentose, PE-, and PC- containing glycan structures. The glycan databases (Supplementary Material 2) were based on the initial glycomic analysis and also included PC-containing glycoforms that were identified by manual analysis of the glycopeptide MS/MS data, as the approach used for glycomic analyses in this study did not result in the identification of PC-modified structures. In addition, the corresponding PE glycoforms were included as PC and PE are closely related structures and may be present simultaneously (Supplementary Material 2). The LC-MS/MS datasets were searched against *F. hepatica* entries in two libraries, (a) the UniProt database (<https://www.uniprot.org/proteomes/UP000230066>) and (b) an internal database (1629 proteins) generated from previous *F. hepatica* NEJs

proteomic analyses (34) corresponding to gene models identified in the *F. hepatica* genome available at WormBase ParaSite and NCBI/ENA: PRJEB6687 and PRJEB25283. Trypsin was specified as the cleavage enzyme allowing up to two missed cleavages. The mass tolerance for precursor ions was set at 20 ppm in the first search and 5 ppm in the main search. The mass tolerance for fragment ions was set to 30 ppm. Carbamidomethyl on the cysteine residue was specified as a fixed modification, and oxidation of methionine was defined as a variable modification. Variable modification of Asn to Asp was included in the O-glycopeptide analysis due to the PNGase F treatment. The false discovery rate (FDR) target was set to 1% at the protein level; 1% (strict) and 5% (relaxed) at the glycopeptide/peptide identity levels. Furthermore, the following acceptance criteria for the glycopeptide identities were used for the MS/MS spectra: (1) the correct peptide ion and/or peptide+HexNAc ion with respect to the assigned glycan-peptide combination; (2) the presence of the expected profile of saccharide oxonium ions, for instance, *m/z* 163.06 and *m/z* 145.05 for oligomannose; and *m/z* 184.07 and *m/z* 534.20 for PC-glycans (see Supplementary Material 3); and (3) inclusion of at least two b- and/or y-ions for at least one of the identified glycoforms.

Bioinformatics Analyses

Identification of the signal peptide sequence and prediction of N- and O-glycosylation sites within the *F. hepatica* NEJs glycoproteins identified was carried out using SignalP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalP-4.1/>), NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) and NetOGlyc 4.0 (<https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0>), respectively.

Occupation of the predicted N- and O-glycosylation sites as well as the number of glycans and glycoforms per glycoprotein and glycopeptide were assessed after manual curation of the peptides identified in the glycoproteomic analysis. A search for *F. hepatica*-specific N- and O-glycosylation amino acid sequence motifs was performed using MoMo suite (V 5.4.1, <http://meme-suite.org/tools/momo/>) (35). Amino acid residues comprising 21 residues within the N-glycosylation sites (10 residues before and 10 after the Asn (N) residue glycosylated) or 31 residues within the O-glycosylation sites (15 residues before and 15 after the Thr (T) or Ser (S) residue glycosylated) were

included in the analyses. The accessibility of glycosylated sites was assessed by secondary structure analysis using the NetSurfP (<https://services.healthtech.dtu.dk/service.php?NetSurfP-2.0>) tool with settings as described previously (36).

The Gene Ontology Annotation (GOA) was based on the UniProt-GOA database (<http://www.ebi.ac.uk/GOA/>). Several putative or uncharacterized proteins were further interrogated using BLAST and InterPro database (<https://www.ebi.ac.uk/interpro/>) for the detection of signature protein motifs. The PSORTb v3.0 software package was used for predicting the subcellular location of proteins (<http://www.psорт.org/>).

Lectin Labeling of *F. hepatica* NEJs

F. hepatica 24 h cultured NEJs as described above were fixed in 4% paraformaldehyde (PFA) in PBS (Sigma-Aldrich), pH 7.4, for 4 h at RT, followed by three washes in antibody diluent (AbD buffer: PBS containing 0.1% (v/v) Triton X-100, 0.1% (w/v) bovine serum albumin and 0.1% (w/v) sodium azide). The NEJs were then incubated in AbD containing a biotinylated lectin: GSL-I (1:200), PNA (1:500), ECL (1:200), WGA (1:200); Table 1). Alternatively, the NEJs were labeled with the monoclonal human anti-Man3/Man5 (Mannotriose/Mannopentaose; clone M3-7, Creative Biolabs), which identifies the most common observed glycan motifs in NEJs, Man3/Man5 (1:500) and anti-phosphorylcholine (PC) (IgA, Kappa from murine myeloma - clone TEPC 15, 1:500; Sigma). As a technical control, NEJs in the same experiment were immunostained with polyclonal rabbit anti-Fh cathepsin L3 (FhCL3) and pre-immune rabbit serum, at a 1:500 dilution (27). All samples were incubated ON at 4 °C. After washing steps with AbD buffer, the lectin-labeled samples were incubated with fluorescein isothiocyanate (FITC)-labeled streptavidin (1:200; Invitrogen) ON at 4 °C, in the dark. Specific secondary antibodies, namely FITC-labelled goat anti-human IgG Fc (1:200; Invitrogen), FITC-labelled goat anti-Mouse IgA (1:200; Sigma-Aldrich); Goat anti-human IgG Fc (1:200; Invitrogen) and goat anti-rabbit IgG (1:200; Sigma-Aldrich) were added to the anti-Man3/Man5 and anti-FhCL3-labeled samples, respectively. A separated group of NEJs was incubated with FITC-labeled Concanavalin A (ConA, 1:200; Sigma-Aldrich). After washes, all samples were counterstained with AbD

TABLE 1
Lectin- and Immuno-labelling reagents and conditions for *F. hepatica* NEJs

Lectin/Primary antibody	Specificity	Secondary labelling	Counter-staining
<i>Griffonia simplicifolia</i> – GSL I (1:200)	Terminal α-Gal Terminal α-GalNAc (Tn-antigen).	FITC-Streptavidin (1:200)	Phalloidin-(TRITC) (200 µg/ml)
<i>Arachis hypogaea</i> –PNA (1:500)	Galβ1,3GalNAc (T-antigen)		
<i>Erythrina cristagalli</i> – ECL (1:200)	Galβ1,4GlcNAc (LacNAc) βGal Lactose		
<i>Triticum vulgare</i> – WGA (1:200)	Terminal βGlcNAc Neu5Ac		
Concanavalin A – FITC-ConA (1:200)	α-Man Glc	N/A	
Monoclonal anti-Man3/Man5 (1:500)	Man3 and Man5	FITC-goat anti-human IgG (1:200)	
Anti- PC (TEPC)(1:500)	PC	FITC-goat anti-mouse IgA (1:200)	
Polyclonal rabbit anti-FhCL3 (1:500)	FhCL3	FITC-Goat anti-rabbit IgG (1:200)	
Rabbit pre-immune sera (1:500)	N/A		

Gal, Galactose; GalNAc, N-acetylgalctosamine; GlcNAc, N-acetylglucosamine; Man, Mannose; FhCL3, *F. hepatica* cathepsin L3.

containing phalloidin-tetramethyl-rhodamine isothiocyanate (TRITC) (200 µg/ml; Sigma-Aldrich) ON at 4 °C, in the dark. The specimens were mounted on slides using a 10% glycerol solution containing 0.1 M propyl gallate. The slides were examined using an Olympus Fluoview 3000 Laser Scanning Confocal Microscope under the PL APO CS 60× oil objective lens. Olympus type F immersion oil was used in viewing and all images were taken at room temperature.

RESULTS AND DISCUSSION

NEJs of the worm *F. hepatica* initiate infection by penetrating the mucosal layer of the intestine, which stimulates parasite growth and development. They are microscopic in size (~40 × 100 µm) and, therefore, only since the development of improved methods of molecular biological techniques, have they become amenable to multi-omics analysis. In the present study, we integrated glycan profiling, glycopeptide structure analysis and proteomic analyses with a series of bioinformatics tools to elucidate the glycosylation status of individual *F. hepatica* NEJs proteins. Until now there has been a dearth of information regarding the glycans of *F. hepatica* but using our approach, we characterized 114 glycoproteins in Som extract and 71 in ES products of NEJs, and we identified and mapped 356 glycosylation sites that altogether exhibited a total of 1727 glycan forms (1690 *N*-glycans and 37 *O*-glycans forms). In doing so, we have revealed the high complexity and heterogeneity of the glycoproteins of *F. hepatica* NEJs (Fig. 1). These molecules are important in virulence, pathogenesis, and immunomodulation of host

immune responses and represent important targets for the development of new diagnostics and vaccines for the control of fasciolosis.

Glycosylation is not encoded by the genome and thus it is impossible to predict which glycan structure will be used to modify a determined glycosite within a protein backbone. In addition, distinct glycan structures can be covalently attached to single or multiple Asn (*N*-glycosylation) or Ser/Thr (*O*-glycosylation) residues on a single polypeptide (1, 37), making the characterization of a glycoprotein in terms of its glycan profile extremely challenging, especially when they are part of complex extracts and carry previously uncharacterized structures in low abundance. In addition, bioinformatics tools and databases for glycan analysis are primarily mammalian-centered and do not favor the study of parasites-derived structures, which have often been demonstrated to be unique (38–40). Consequently, to date, only a few individual helminth parasite molecules have their glycosylation profile fully resolved (12, 41–46).

Curation of the Glycome Data

We assembled an extensive list of *N*- and *O*-glycans in our initial analyses of the total glycans released from the total proteins within *F. hepatica* NEJs Som and ES extracts. Figure 2 and Supplementary Material 1 highlight a curated version of the glycome data obtained by complementary *N*- and *O*-glycopeptide analyses. In relation to the complete list of *N*- and *O*-glycans (non-curated) presented in Supplementary Material 1, it is worth noting that, because NEJs are obtained by excystment from metacercariae in medium that contain ox bile proteins (within the sodium tauroglycocholate preparation), it is possible that some glycans in our analyses are of bovine origin. The structures identified in this study that were not mapped to NEJ proteins may reflect the methodology we have used, but also the limited sensitivity of *O*-glycopeptide analysis to detect *O*-glycans with higher mass (*i.e.*, mucin-types). Moreover, glycomic analysis after beta-elimination release allows more *O*-glycans from all proteins (both in level and type) to be recovered. All glycan structures that were successfully mapped to specific *F. hepatica* NEJs glycoproteins are graphically represented in Supplementary Material 4.

F. hepatica NEJs Express Generic Mannosylated and Short *N*-Glycans

Using *N*-glycan and *N*-glycoproteomic analyses we revealed 53 different *N*-glycan structures that modify the NEJs Som and ES proteins (Fig. 2A), which are comprised of paucimannose ($n = 7$), oligomannose ($n = 6$), and hybrid/complex-like ($n = 38$) structures. Based on the frequency at which these structures were observed we can conclude that the NEJs proteins are predominantly decorated by oligomannose type *N*-glycans, with HexNAc₂Hex₅ (*i.e.*, Man₅GlcNAc₂) being the most frequent *N*-glycan form identified. Other oligomannose

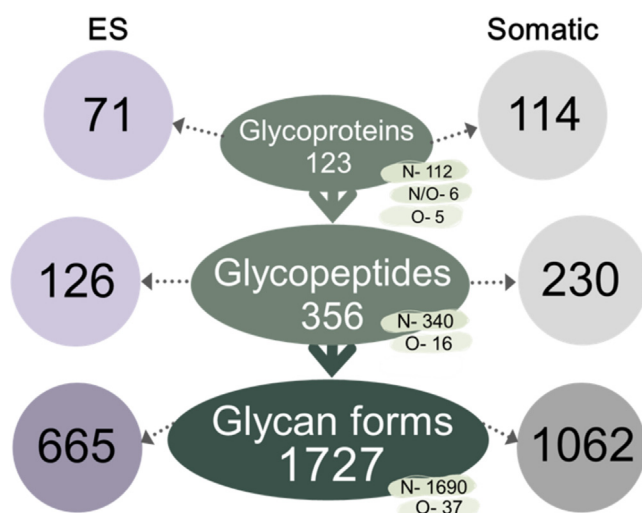


FIG. 1. Summary of the results obtained from the glycomic and glycoproteomic analyses of *F. hepatica* NEJs ES and Somatic extracts. The schematic highlights the diversity generated by protein glycosylation. The number of glycan forms presented represents the total number that modify the 356 glycopeptides mapped to *F. hepatica* proteins by complementary glycopeptide analysis. The breakdown in terms of *N*- or *O*- structures is shown in blue and yellow. From top to bottom: *N*-: *N*-glycoproteins, *N*-glycopeptide, or *N*-glycan form, respectively. *N/O*-: proteins *N*- and *O*-glycosylated; *O*-: *O*-glycoproteins; *O*-glycopeptide; *O*-glycan form, respectively. ES, excretory-secretory extract.

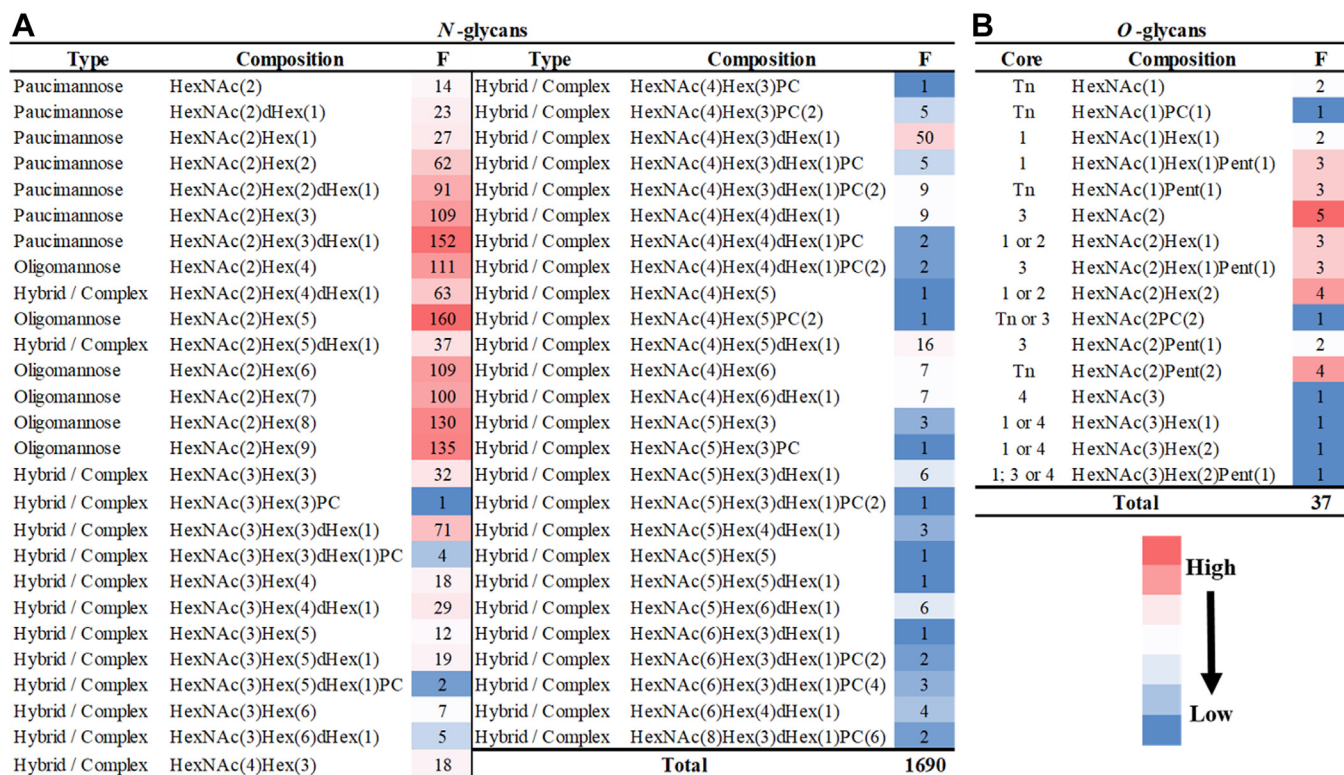


FIG. 2. *N*- and *O*-glycans associated with glycoproteins in *F. hepatica* NEJs Somatic extracts and ES products, classified according to type or core. A, *N*-glycans; (B) *O*-glycans composition and the frequency (F) with which each glycan structure occurs in NEJs glycoproteins. Tn: *O*-glycans without a classical core, known as Tn-antigen. Total: represents the number of *N*- or *O*-glycan forms found in our analyses.

structures, such as Man₆₋₉GlcNAc₂, previously identified in NEJ extracts by Garcia-Campos *et al.* (25), were also common across the samples analyzed in this study. Since Man₅₋₉GlcNAc₂ are also present within the glycome of mammalian cells and tissues it supports the suggestion that *F. hepatica* may use host-like glycans to evade the host immune response (8) or to hide antigenic epitopes of their proteins (47–49), a practice that could be conserved across various helminths (8, 50, 51). Furthermore, mammalian host recognition of these conserved structures *via* lectin receptors may promote pathogen sensing or, conversely, aid pathogen transmission and dissemination by activation of transcription factors that stimulate an immune response that favors the parasite (52–54).

Paucimannose/truncated structures make up approximately a quarter of all the *N*-glycan forms associated with NEJs' glycoproteins (Fig. 2A). Our analysis also revealed that a number of NEJs glycoproteins carry small *N*-glycans that consist only of the diacetylchitobiose core, with or without the α 1,6-linked fucose (dHex, HexNAc₂Hex₀₋₁dHex₀₋₁, *i.e.*, Man₀₋₁GlcNAc₂Fuc₀₋₁). It is worth noting that core fucose was identified in 35% (446/1268) of the *N*-glycopeptide identities, which is typically α -6 linked in vertebrates, but can alternatively be α -3 linked and/or core-difucosylated in helminths (9). However, digestion of the glycopeptides samples with PNGase F resulted in the cleavage of ~90% *N*-glycans. As the

number of *N*-glycopeptide spectra that matched structures containing core fucose (nine *N*-glycopeptide identities) did not increase following this digestion, it follows that α -3 linkages were not present and, therefore, NEJs only produce core α -6-fucose (Supplementary Material 5).

While, as mentioned earlier, many NEJ glycans mimic those of mammalian cells, the short paucimannose structures resemble plant glycans (55). The biosynthesis of *N*-glycans begins in the ER and progresses in the Golgi apparatus, where the maturation (formation of complex structures by core fucosylation and addition of terminal GlcNAc residues) occurs. Whereas this process is highly conserved in different eukaryotic organisms, plants can further process *N*-glycans inside vacuoles post-Golgi maturation (during the transit of glycoproteins to their final cellular destination), which results in characteristic truncated structures like those we observed in NEJ glycoproteins. This truncation that occurs in plants is associated with the activity of exoglycosidases present in vacuoles and in the extracellular environment that hydrolyze glycosidic bonds of complex glycans (56–58), and with endoglycosidases that cleave *N*-glycans near the protein backbone and reveal the highly truncated chitobiose cores (55). While the functional impact of this truncation of *N*-glycans in NEJs is still unknown, lectins capable of binding to oligomannose and paucimannose glycans on the NEJs

tegument have been shown to significantly reduce their ability to invade the host's gut wall (26), suggesting that the interaction of these glycans with host cells play an important role in liver fluke infection.

There is a paucity of data available regarding the *F. hepatica* glycosidases that may be involved in both *N*- and *O*-glycan truncation pathways and their compartmentalization within the parasite cells/tissues. However, intense glycosidase activity has been verified in the secreted proteins of *F. hepatica* adult parasites, which has been linked to several enzymes, including β -*N*-acetylhexosaminidases, α -fucosidases, α -mannosidases, and α -galactosidases (59); α -mannosidases, and α -galactosidases were also identified proteomically in the NEJs Som extract and ES products in this study (Supplementary Material 6). The presence of truncated *N*-glycans associated with the ES secreted proteins suggests an unusual compartmentalization of enzymes involved in *N*-glycan biogenesis by *F. hepatica*, which is supported by our previous data showing that glycosyltransferases, namely GALNT10, responsible for *O*-glycan formation, and B4GALT2 that adds Gal to terminal GlcNAc to form the LacNAc moiety, are found amongst the cargo of *F. hepatica* EVs from adult parasites (60). Indeed, glycosidases present in the NEJs' ES or on the NEJs surface may play a role in parasite-host interactions by modifying the structure of surface glycans (61, 62), and/or by breaking down and assimilating host-derived oligosaccharides and monosaccharides. The recent *in silico* characterization of *F. hepatica* protein glycosylation genes by McVeigh *et al.* (30) revealed a temporally regulated dynamic expression of glycosylating genes that is likely important to the growing and developing parasite that needs to change and match their surface according to the host environment.

F. hepatica NEJs Express Unique Multi-Phosphorylcholine (PC) N-Glycans

We discovered that glycan forms of the hybrid/complex-type make up ~25% ($n = 467$) of the structures on NEJs glycoproteins (Fig. 2A). In line with previous glycomic data, NEJs glycoproteins often carry short complex biantennary glycans such as HexNAc3-4He x 3dHex0-1 (*i.e.*, Man3GlcNAc3-4Fuc0-1, $n = 171$) (25, 28). However, our MS data also revealed various bi- and tri-antennary complex-type glycans elongated with Hex and HexNAc (HexNAc4-5He x 5-6dHex0-1 and HexNAc4-6He x 3dHex0-1), which are consistent with the structure of LacNAc (*i.e.*, Gal β 1,4GlcNAc-) and LacdiNAc (*i.e.*, GalNAc β 1,4GlcNAc-) motifs, respectively. In addition, we observed some common monosaccharide structures presenting unusual linkages; for example, we could not assign the putative structure of the *N*-glycan m/z 1219 to 1 (HexNAc2He x 4dHex1, Supplementary Material 1) because the spectrum indicates an unusual position/linkage for the fourth Man residue, other than the expected α -3, α -6 or α -2-linked sugars and, therefore, further work will be necessary for the detailed structural assignment of these complex glycans.

Vertebrates typically extend their *N*-glycans with LacNAc, and, therefore, LacdiNAc is considered an unusual aspect of glycosylation found in helminth glycomes besides *F. hepatica*, including *Dictyocaulus viviparus* (63) and *T. spiralis* (64). Parasite LacdiNAc motifs belong to a group of host-like glycans that nevertheless induce antibody responses; this is due to the expression of the LacdiNAc in unusual linkage configurations (65, 66) and possibly because, unlike mammals, they do not add terminal sialic acid, but in some instances modify them with zwitterionic phosphorylcholine (PC) and/or phosphoethanolamine (PE) terminals (5, 8). Expression of LacdiNAc-containing glycans by *F. hepatica* was first suggested by specific lectins and anti-glycan antibodies binding to glycoproteins in adult worm extracts (40). Of note, the structure of Gal β 1,4GlcNAc- motifs was verified by our glycomic analysis (Supplementary Material 1) but the HexNAc-HexNAc motifs modified by PC could not be assigned in terms of composition and linkages due to the presence of terminal PC in these branches. As enzymatic release of these PC terminals is not possible, future analysis following chemical removal will allow for the proper assignment of these interesting parasite glycans.

The presence of non-sugar moieties such as PC on the terminal GalNAc/GlcNAc residues of the NEJs glycans adds additional complexity, heterogeneity, and most likely antigenicity to these glycoproteins. The identification of PC on *F. hepatica* glycoproteins is consistent with early studies by Sloan *et al.* (67), which demonstrated that more than 60% of the antibody reactivity of serum from cattle and rats infected with liver fluke was directed against PC epitopes present in parasite somatic extracts. Our analyses revealed that ~2.5% of the glycan forms associated with NEJs glycoproteins contain terminal PC (Fig. 2, A and B). Unexpectedly, we found that *F. hepatica* NEJs produce unique glycan motifs by adding two PC residues to the same terminal HexNAc monosaccharide, which we demonstrated by in-depth MS/MS analyses of an *N*-glycopeptide from the NEJs cysteine peptidase cathepsin B11 (Fig. 3). The annotated b- and y-ions, glycosidic fragmentation and oxonium ions of the spectrum in Figure 3A is presented in Supplementary Material 7. All annotated spectra of the identified *N*- and *O*-glycopeptides are available *via* the Byonic viewer files at the Proteome Xchange repository. Further inspection of glycopeptides from the same protein revealed that even more complex *N*-glycans, HexNAc6He x 3dHex1PC4 and HexNAc8He x 3dHex1PC6, can modify the same glycosite (Supplementary Material 8).

Taken together, many B-ions that were diagnostic of the terminal (PC)₂HexNAc on the *N*-glycopeptides were observed (Fig. 3E), which can be used to screen LC-MS/MS data for their occurrence. The high mass resolution of orbitrap-measured fragment ions enables high mass accuracy of fragment ions, which adds confidence to the correct assignment of the presented compositions. For instance, the 14

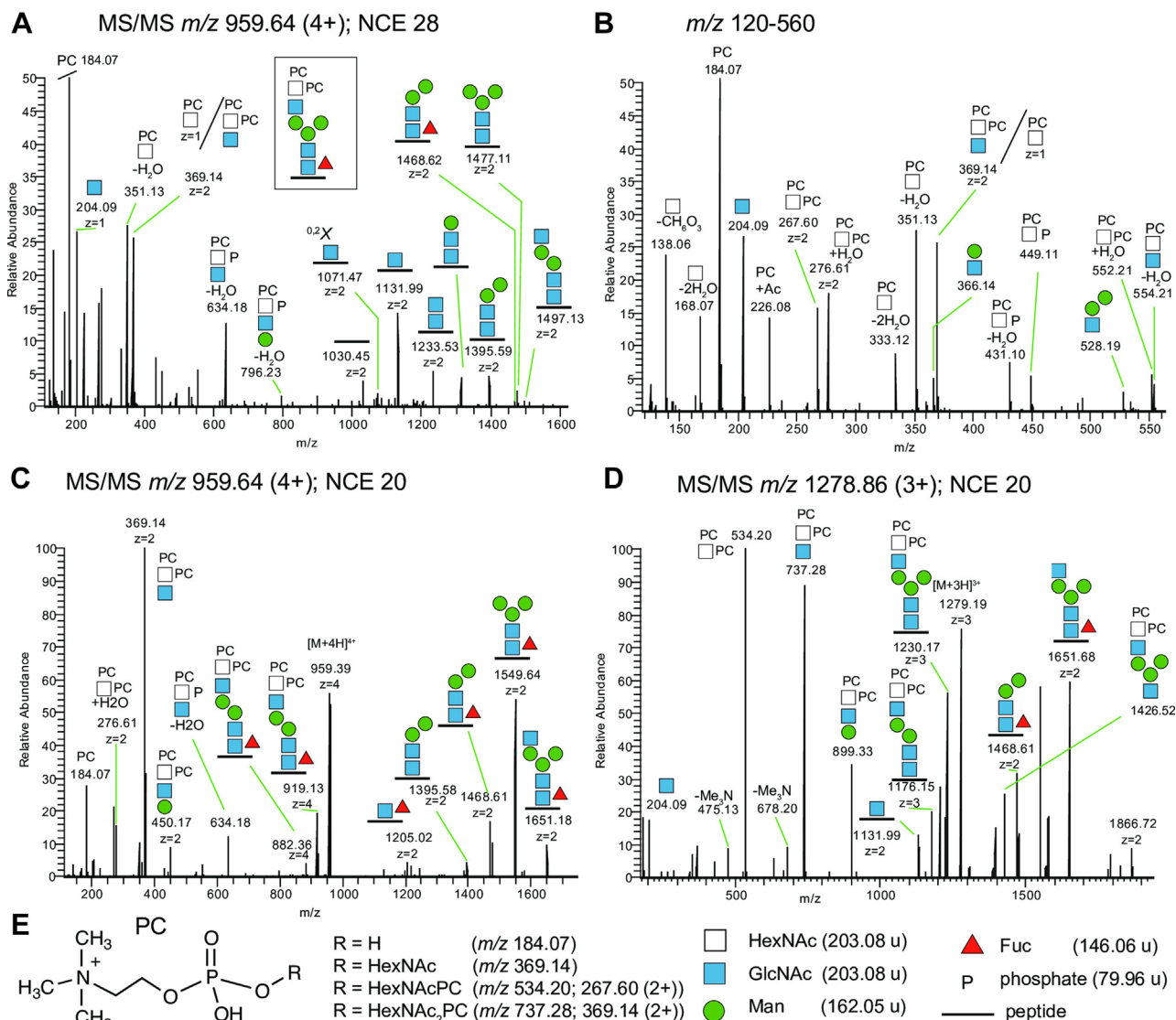


FIG. 3. Glycoproteomic analysis of *N*-glycopeptides carrying multiple PC residues in *F. hepatica* NEJs glycoproteins. A, MS/MS of a 4+ glycopeptide precursor ion (m/z 959.64) with the peptide sequence [K].TYDEDKFYGNSSYNVEK.[S] from *F. hepatica* cathepsin B11 (maker-scaffold10x_889_pilon-snap-gene-0.33) and HexNAc₄Hex₃dHex₁PC₂ glycan composition. m/z 184.07 corresponding to phosphorylcholine (PC) was the base peak (relative abundance scale starts at 50%). The Y1 peptide+GlcNAc (m/z 1131.99) to Y5 peptide+GlcNAc₂Man₃ (m/z 1497.13) ions showed the characteristic *N*-glycan chitobiose core structure, and diagnostic ^{0,2}X ring cleavage into m/z 1071.47 verified that the glycan was *N*-glycosidically attached to Asn. The ion m/z 1497.13 suggests that both PC residues are attached to a terminal HexNAc. Direct evidence for HexNAc+(PC)₂ composition came from the B-ions at m/z 267.60 (2+) and m/z 276.61 (2+), and a significant ion corresponding to (HexNAc)₂+(PC)₂ at m/z 369.14 (2+). Further, B-ions corresponding to loss of PC (and nH_2O) included m/z 554.21, m/z 352.13, and most likely m/z 369.14 (1+) coinciding with m/z 369.14 (2+); and loss of PC, less the mass of phosphate, included ions at m/z 634.18 and m/z 449.11. B, The expansion of the m/z 120 to 560 region revealed many of the PC associated B-ions. C, MS/MS of the same glycopeptide structure at a lower NCE. The deduced structure (3A, boxed) was inspected at a lower normalized collision energy (NCE) of 20% confirming that the core Fuc was correctly assigned due to intense presence of fucosylated Y-ions at m/z 1205.02 to m/z 1651.18. D, the MS/MS of the corresponding 3+ precursor ion also gave a complementary fragmentation profile, showing intense B-ions at m/z 899.33; m/z 737.28; and m/z 534.20. Also, loss of (CH₃)₃N (-59.07 u), was observed, which further verifies the PC identifications. Charge states of fragment ions are indicated when $z \geq 2$. E, structure of PC and the masses of critical PC-containing fragment ions. Glycan symbols are according to the SNFG format (ref: PMID: 26543186).

annotated PC-associated B-ions shown in Figure 3 had an average mass accuracy of 2.4 ± 1.2 ppm, whereas 12 identified HexNAc and Hex oxonium ions had an average mass accuracy of 1.3 ± 0.7 ppm (Supplementary Material 3).

Some of the PC-modified *N*-glycans discovered in NEJs resemble those previously observed in the nematodes *A. viteae* (12) and *Trichuris suis* (41), for example, where PC was determined to modify residues of GlcNAc and GalNAc in the

antenna of the glycans. The addition of PC in parasite glycans depends on the acquisition of choline from the lipid membranes of their hosts and the subsequent transfer of PC into their glycoconjugates by PC-transferases (12, 38). The identification of a collection of PC-modified glycans in NEJs glycoproteins, which includes multi-PC residues in the same monosaccharide residue of both *N*- and *O*-glycans, suggests that *F. hepatica* expresses a range of PC-transferases capable of adding PC residues to different monosaccharides and linkages positions. Given the key role PC plays in parasite-host interactions (68, 69), these enzymes should represent new targets for exploration in the development of novel treatments against liver fluke.

PC moieties in parasite *N*-glycans can prevent complement activation *via* different pathways including binding to host C-reactive protein (CRP) (67, 70) or mannose-binding lectin (MBL) (68). We have recently shown that *F. hepatica* can circumvent complement activation *via* the lectin pathway by preventing binding of MBL to the parasite surface raising the possibility that PC is involved in this immune evasion strategy (27). PC also has been shown to affect the proliferative signaling pathways in B and T-cells, the maturation of dendritic cells (DC), and the degranulation of mast cells (69, 71–74). Furthermore, it has been shown that the PC-modified ES-62 glycoprotein from *A. viteae*, performs a number of immunomodulatory functions including skewing of the immune response towards Th2 (72, 75, 76), enhancing DCs maturation and IL-4 levels, reducing IFN- γ production in T-cells, desensitizing B-cell activation and proliferation, and increasing IgG1 whilst decreasing IgG2 production (77). Similar modulation of immune responses is also observed in animals naturally infected with *F. hepatica* (78–81), indicating a putative key role for *F. hepatica* PC moieties in manipulating the host immune response during infection.

F. hepatica NEJs Glycoproteins are Decorated With Novel Canonical O-Glycans Containing GalNAc-pentosylated Cores

There is a severe lack of information regarding the structure and diversity of parasite-derived *O*-glycans, somewhat due to their small size and low-level expression that makes them difficult to analyze (82). The limited studies available suggest that *O*-glycans also play key roles in recognition, binding, and signaling events and those few parasite *O*-glycans characterized to-date display unique immunogenic motifs that have been exploited for immunodiagnosics development (83, 84). In relation to *F. hepatica*, only the antigenic Tn-antigen motif (GalNAc with a glycosidic α linkage to Ser/Thr) has been identified in adult worm extracts and this was detected using specific monoclonal antibodies (85).

O-linked glycans include any single monosaccharide or oligosaccharide that is glycosidically linked to the side chain hydroxyl of Ser/Thr residues of proteins. In the case of the canonical *O*-glycosylation, elongation of *O*-GalNAc results in

the generation of mucin-type *O*-glycan cores, which can be further modified in a specific manner depending on the glycosyltransferases expressed by the organism (86, 87). The collection of *O*-glycans linked to NEJs glycoproteins include structures containing 1, 2, 3, and 4-types cores, with those that contain core 1 (Gal β 1-3GalNAc-) being more frequent in the Som and ES preparations (Fig. 2B). We found that the NEJs core 1 *O*-glycans may be extended with HexNAc (commonly GlcNAc) and Gal residues (forming the LacNAc motif) and, in some cases, terminal PC residue(s) is added. Also, the immunogenic Tn-antigen motif present within the NEJs *O*-glycopeptides could be modified by PC or, more uniquely, by a pentose. Glycans containing 3-type core (GlcNAc β 1,3GalNAc-), make up about a quarter of the *O*-glycan forms identified in the present study and were observed to be elongated with repeating units of LacNAc (poly-LacNAc; Supplementary Material 1). We also observed that all the *F. hepatica* NEJs *O*-glycans, irrespective of the core-type, could be modified with a pentose (Pent), which is added to the *O*-GalNAc residue (Fig. 2B; Supplementary Material 1).

Glycosyltransferases capable of forming the observed Tn-antigen and cores 1-4-type *O*-glycans were anticipated based on our previous *in silico* identification of orthologous sequences in *F. hepatica* (30). In line with this, UDP-glucuronate decarboxylase, which converts UDP-GlcA to UDP-xylose, was listed among our glycoproteomics data (Supplementary Material 6). However, modification of the canonical *O*-glycan cores by a α -pentose, which is present in ~43% of the NEJs *O*-glycopeptides, is a novel finding, and further research to unveil this biosynthetic pathway could provide new insights for *F. hepatica* *O*-glycobiology. Indeed, due to their unusual architecture, we interrogated one of these motifs containing *O*-GalNAc-Pent by MS fragmentation following specific exoglycosidase treatment to better characterize the glycan structure. The MS/MS fragmentation of parent ions all showed loss of pentose ($M - 132 - \text{water}$ (18)), indicating that a pentose was located at a non-reducing end of the glycan. *Salanum tuberosum* β 1,2 xylosidase (Europa, EU224) was not active on these structures (reaction was carried out at 37 °C overnight) and, as C2 position of GalNAc alditol is N-acetylated, the pentose is more likely linked to C4 or C6 position of GalNAc alditol (Fig. 4). Altogether, our results confirmed that this glycan represents a novel pentosylated structure, which was structurally defined both by the glycoproteomic and glycomic analyses and to date has only been observed in *F. hepatica* NEJs glycoproteins.

In addition to the PC and Pent modifications, another feature of the *O*-linked glycans decorating proteins present in the NEJ extracts is that they are relatively short structures. This may be due to challenges associated with mucin domain analysis using standard proteomic workflows, as mucin domains commonly lack tryptic cleavage sites. This limits the sensitivity of the glycoproteomic techniques to detect large

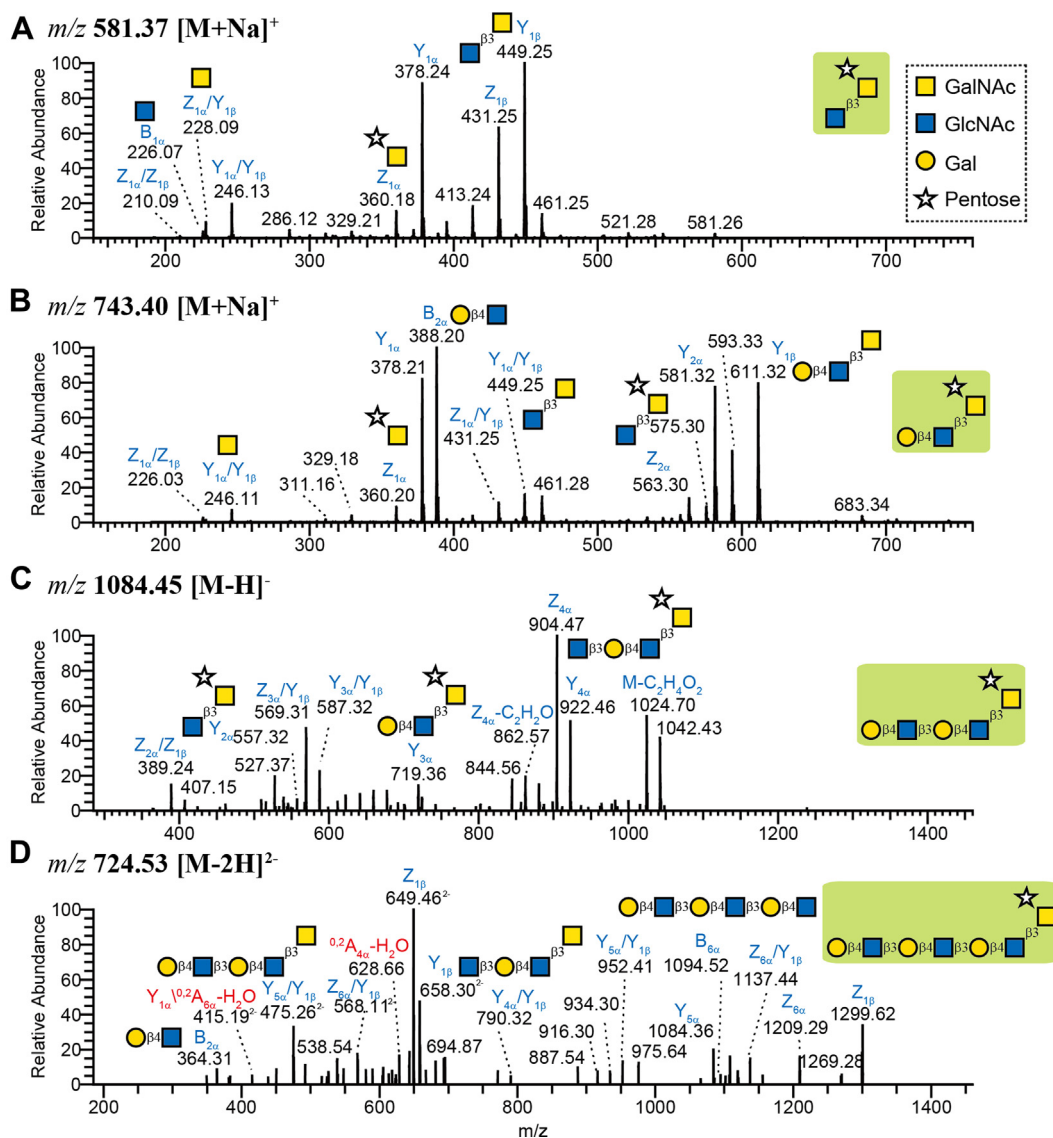
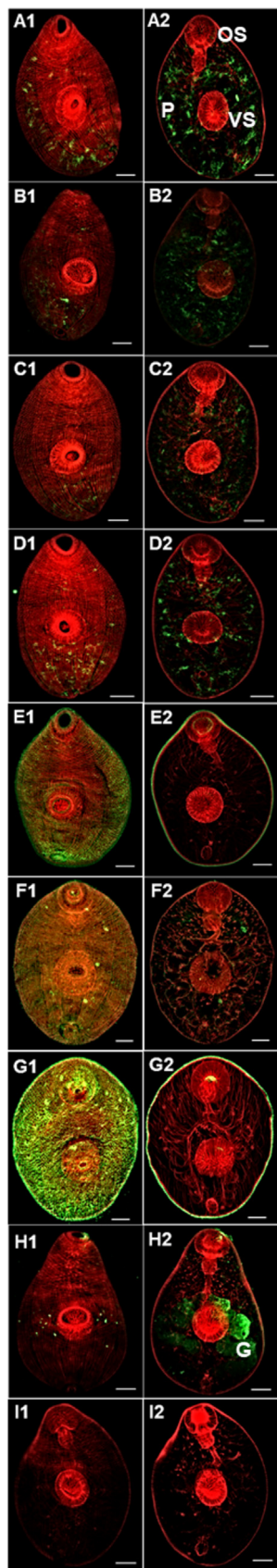


FIG. 4. LC-MS/MS analysis of four Pentose-containing O-glycans. O-glycans released by reductive β -elimination were subjected to LC-MS/MS in both positive- and negative-ion mode as the substitution of pentose (Pent) is more labile in negative-ion mode. In the positive-ion mode (A and B), Z ions at m/z 360 $[M+Na]^+$ indicate the substitution of Pent on GalNAc alditol. In the negative-ion mode (C and D), most Z ions are devoided of Pent (e.g., m/z 407 ($Z_{2\alpha}/Z_{1\beta}$ ions) and m/z 569 ($Z_{3\alpha}/Y_{1\beta}$ ions) in C; m/z 790 ($Y_{4\alpha}/Y_{1\beta}$), 475/952 ($Y_{5\alpha}/Y_{1\beta}$), and 568/1137 ($Z_{6\alpha}/Y_{1\beta}$) in D). The Y and/or Z ions indicate that one terminus contains Gal, whilst the terminus of the 6-arm contains one Pent residue. Together with the B1 ion at m/z 226 (A), B2 at m/z 388 (B) and B2 and B6 ions at m/z 364 and 1094, respectively (D), the fragmentation suggests a glycan with one arm carrying a linear triLacNAc structure and another arm with one Pent. This is confirmed by a cross-ring cleavage at m/z 628 (A4 ions, D) of internal GlcNAc. The Y ions at m/z 449 (A and B) and m/z 407 (C) suggest a GlcNAc-GalNAc core. The fragment ion at m/z 844 (C) and m/z 1209 (D) are more likely derived from parent ions losing substituted Pent and $C_3H_8O_4$ (108 μ), which are diagnostic ions for core 1 or 3-type O-glycans (GlcNAc β 1,3GalNAc α 1; Karlsson *et al.* 2004, 10.1016/j.jasms.2004.01.002). Finally, the presence of Z ions at m/z 360 (A) and no isomeric structure of all pentosylated glycans suggest that the Pent residue is linked to the GalNAc alditol.

structures and, thus, we should not exclude that those larger O-glycans (Hex $_2$ - $_8$ HexNAc $_0$ - $_8$ Pent $_0$ - $_1$; [Supplementary Material 1](#)) evidenced by glycomic analysis might be attached to the NEJs glycoproteins.

The function of these novel NEJs pentosylated O-glycans remains unknown. To date, xylose (Xyl) is the only pentose monosaccharide identified in parasitic glycans. However, only xylosylated N-glycans have been described in helminths

extracts, that is, β 1,2-xylosylated N-glycans (51), which are also common carbohydrates of plants (88, 89). These structures are important since parasite and plant β 1,2-xylosylated motifs are recognized as targets for IgE antibodies and are suggested to contribute to the induction of strong Th2-biased allergy-like responses in murine models (<http://hdl.handle.net/10810/25071>) (90). Whilst more analyses are required to determine the specific type of pentose associated to NEJs glycans, Xyl



residues are not commonly observed modifying canonical O-glycans and have only been previously described for glycans present in proteoglycans, such as O-xylosyl glycans (in which a Xyl is added directly to a Ser residue of a protein) and O-glucosyl glycans (37).

Protein O-glycosylation occurs in the Golgi, where O-glycans are normally added to sites exposed on the surface of the folded protein in a stepwise fashion. *F. hepatica* mucins have been mainly associated with the mucopolysaccharide layer of metacercariae and with the parasite tegumental surface, where they provide protection and moderate key interactions with the host (85, 91). Electron microscopical analysis of the parasite tegument reveals a highly active syncytial layer with extensive networks of mitochondria and Golgi apparatus, which was suggested to reflect the continuous turnover of the thick surface glycocalyx (23). Turnover of the glycocalyx enables the parasite to shed bound antibodies and prevent the activation of the classical pathway of complement (23, 27). Cancela *et al.* (92) found that mucin expression is upregulated in NEJs and suggested a role for these glycoproteins in the invasion of the intestinal mucosa.

Lectin Staining Suggests Compartmentalization of Glycans in *F. hepatica* NEJs Tissues

Lectin immunofluorescence staining suggests distinct distribution and accessibility of glycan structures on the surface and within the NEJs (Fig. 5). LacNAc and terminal β GlcNAc-binding lectins, namely ECL and WGA, were detected in the internal parenchymatous tissue that surrounds various organs, including those forming the digestive, nervous, and reproductive systems. The lectins PNA (mainly core 1 O-glycans, and terminal Gal) and GSL-I (terminal α Gal and Tn-antigen) also primarily bound to parenchyma cells but were especially found in cells lining the intestine (Fig. 5, panels B and A, respectively). In contrast, ConA, which has an affinity to α -linked Man residues, strongly and preferentially bound to the tegumental surface of the NEJs, covering their spines, which is in line with the abundance of oligo- and paucimannose structures in the extracts. This observation was further supported using anti-Man3/Man5 monoclonal antibodies that recognize mannotriose/mannopentose (*i.e.*, Man3 and Man5 glycan motifs) as these bound primarily to the NEJs surface and, in lesser amounts, to the gut region (Fig. 5,

FIG. 5. **Lectin and anti-glycan staining of *F. hepatica* NEJs.** A panel of lectins and anti-glycan antibodies were used to characterize the glycan distribution on *F. hepatica* NEJs tissues. A1 and 2: GSL I; B1 and 2: PNA; C1 and 2: ECL; D1 and 2: WGA; E1 and 2: ConA; F1 and 2: anti-Man3/Man5; G1 and 2: anti-PC; H1 and 2: Non-related positive control anti-FhCL3 highlighting the NEJs bifurcated gut (G); I1 and 2: control pre-immune sera. *Left panel:* Outside surface of NEJs. *Right panel:* Inside of NEJs. Other main NEJs features are noted: OS, Oral sucker; VS, ventral sucker; P, Parenchyma. Scale bars: 25 μ M.

panels E and F). As *N*-glycans containing Man3/Man5 moieties are the most frequent structures attached to NEJs glycoproteins, it is not surprising that the intensity of anti-Man3/Man5 immunostaining overlapped with that observed for ConA.

Strong anti-PC fluorescence was observed on the surface of the NEJs, specifically linked to the outer tegument (Fig. 5, panels G), suggesting PC is used to decorate glycoconjugates on the surface of NEJs, likely including glycolipids. Such intense presence of PC-decorated glycoconjugates on the juvenile parasite surface might reinforce their importance for the parasite's interaction with the host and escape of detrimental immune responses. Our results differ from the study by Garcia-Campos *et al.* (25), which identified Man, LacNAc and terminal Gal/GalNAc expressed on the NEJs surface, possibly due to technical differences between the studies. However, both studies associate PNA and GSL-I binding with the NEJs Som extract demonstrating that *F. hepatica* NEJs produce numerous *O*-glycans but perhaps keep them internally hidden. Lastly, immunostaining with the control antibody anti-rFhCL3 showed the expected staining highlighting the NEJs bifurcated gut (93), whilst no fluorogenic signal was detected in parasites stained with the rabbit pre-immune antibodies (Fig. 5, panels H and I, respectively).

Complexity and Heterogeneity of *F. hepatica* NEJs Glycopeptides

Our glycomic analysis of *F. hepatica* NEJs showed that the parasite glycan repertoire is significantly more complex than anticipated by the *in silico* study of McVeigh *et al.* (30), and consists of novel *N*- and *O*-glycans containing poly-LacNAc, PC and multi-PC terminals, pentosylation of *O*-GalNAc cores, as well as the strong indication of branches formed by LacdiNAc motifs. By extension, these results suggest that

F. hepatica possesses a wider portfolio of glycosyltransferases/exoglycosidases than previously considered and/or that these enzymes possess distinct specialized activities. While the glycan study discussed above has revealed new information on the glycosylation of *F. hepatica* proteins this is a general broad view of the glycan composition of the Som and ES preparations. Consequently, to provide a deeper biological and functional context to the glycan structures identified in this study, we characterized and determined the *N*- and *O*-glycopeptides to which these glycans are attached and mapped those to their respective proteins within the Som and ES extracts. This approach allowed us to compile and compare information on site-specific glycosylation and determine their complexity and global heterogeneity (94).

The data generated from the *N*- and *O*-glycopeptide analyses allowed us to (a) identify which proteins in the Som and ES preparations are glycosylated, (b) map the *N*- and *O*-glycosylated sites within each glycoprotein, and (c) inform which glycan(s) structures are used to decorate each glycosite. As a result, a total of 123 glycoproteins were identified in the NEJs extracts, the majority of which were detected in the Som extract (N = 114) compared to the ES (N = 71). Half of these glycoproteins, *i.e.*, N = 62, are present in both Som and ES preparations. The majority of NEJs proteins (N = 112) are *N*-glycosylated while five are *O*-glycosylated and six harbor both *N*- and *O*-glycans. Nine glycoproteins were detected exclusively in the ES extract, eight of which are *N*-glycoproteins, and one is an *O*-glycoprotein (Fig. 6 and Supplementary Material 6).

Glycosylation of *F. hepatica* NEJs Generates Rich Protein Variability

By mapping individual glycopeptides and their associated glycan forms to their respective glycoproteins, we identified

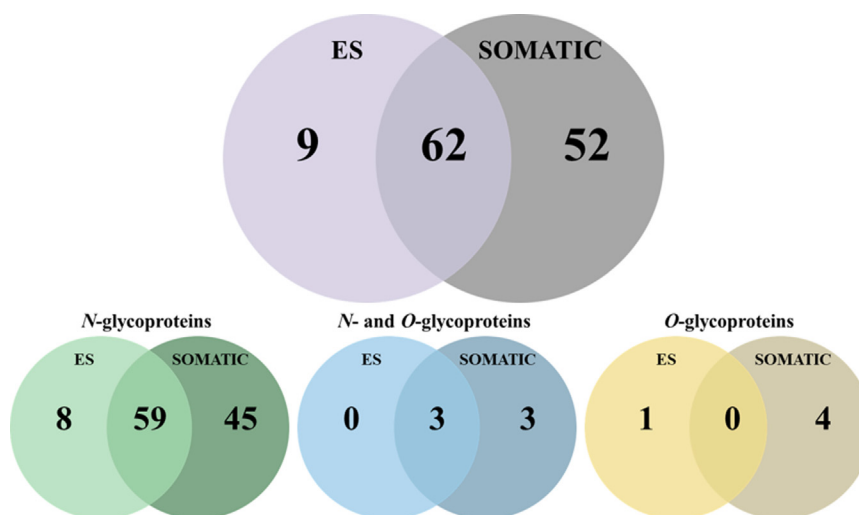


FIG. 6. Distribution of the glycoproteins characterized in this study in the *F. hepatica* NEJs ES and Somatic extracts. Bottom row, separation of the glycoproteins according to their glycan type.

340 unique *N*- and 16 *O*-glycopeptides associated with the *F. hepatica* NEJs glycoproteins (Fig. 1). We found that each NEJs glycoprotein has, on average, three occupied glycosites and that the ES glycoproteins tend to contain less glycosylated sites compared to the glycoproteins in the Som preparation (Fig. 7C). Moreover, based on the average glycopeptides identified per protein, *N*-glycoproteins contain more glycosylated sites (on average three glycopeptides/protein) than the *O*-glycosylated proteins (average of two glycopeptides/protein).

We observed vast micro-heterogeneity within the *F. hepatica* NEJs glycoproteins, whereby a specific glycosite could be modified with a variety of glycan forms, resulting in many versions of the same glycopeptide that diverge only in terms of their glycosylation (95). While the NEJs glycoproteins are mostly single-site glycosylated (63%), most of these glycosites have more than one glycan form associated with them (Fig. 7A). The heterogeneity of NEJs glycoproteins manifested in various ways; for example, some proteins possess several glycosites but relatively little glycan heterogeneity between them, whilst others have one glycosite modified with a multitude of glycan forms. There is also significant variation between the copies of the same protein in found in the Som extracts and in the ES (Supplementary Material 9). The degree of micro-heterogeneity of the 356 glycosites is highlighted in Figure 7, A–C. Our data show that *N*-glycosylation is vastly

more heterogeneous than *O*-glycosylation, with up to 22 glycosylated versions observed for certain *N*-glycopeptides (Fig. 7A). In contrast, most of the *O*-glycosylated sites identified were modified by only a single *O*-glycan form (Fig. 7B). Altogether, our data indicate that the NEJs of the parasite *F. hepatica* greatly vary their repertoire of proteins by adding an assortment of sugars to the same polypeptide sequence. On average, 14 glycan forms can be used to modify the same NEJ protein, with significantly more variation found in *N*- compared to *O*-proteins (Fig. 7E).

The assembly of complex glycans is subject to variation and micro-heterogeneity; however, it is not known whether this variation is random or not. Certainly, glycan heterogeneity can affect protein structure and function, including binding specificities to receptors and antibodies (96) but the impact of such observed glycoprotein variability on liver fluke biology is currently unknown. Interestingly, *F. hepatica* is characterized by high genetic polymorphism (97) that could contribute greatly to glycan heterogeneity and confer this parasite with selective advantages. We can speculate that it allows for evasion, modulation, and distraction of the host immune response, whereby copies with less immunogenic sugars might escape or drive a specific type of immune response aiding parasite survival.

Micro-heterogeneity may also play a role in enabling this parasite to infect and survive within a wide range of different

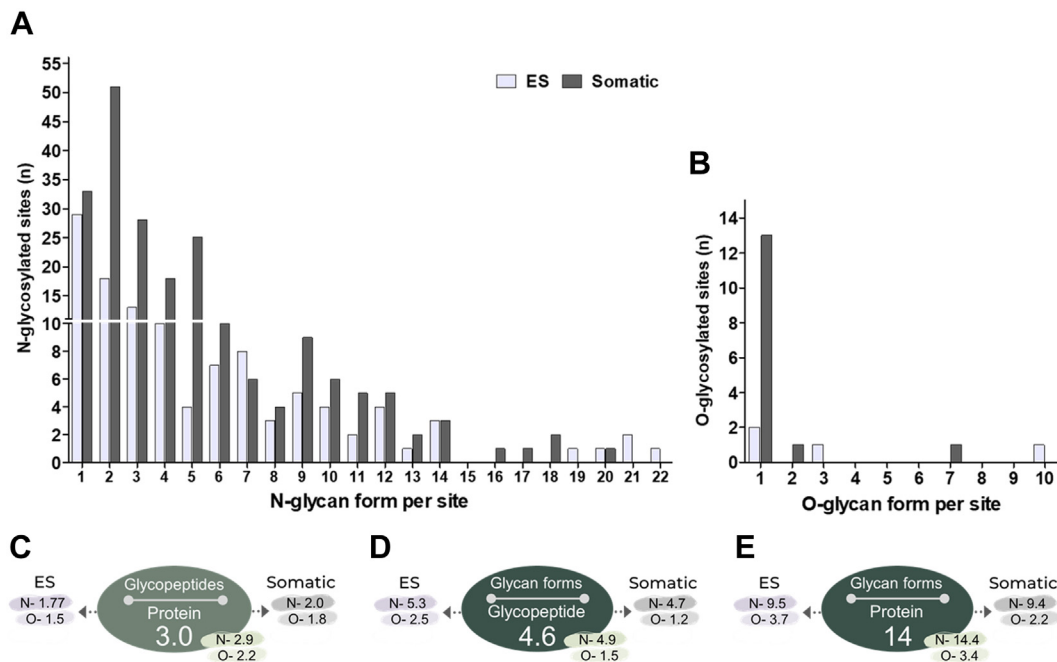


FIG. 7. Heterogeneity of *F. hepatica* NEJs glycoproteins based on the different glycan forms that modify the identified glycosylated sites within the glycopeptides. A, *N*-glycosites and (B) *O*-glycosites distributed according to the number of different *N*-glycan forms that can modify them and the differences in ES and Somatic samples. C, average number of glycopeptides per NEJs' glycosylated protein. D, average number of glycan forms that can modify the same glycopeptide in secreted (ES) and Somatic extracts. E, average number of glycan forms that can modify individual NEJs' glycoproteins and the breakdown in different extracts, ES and Somatic extracts, and *N*- and *O*-glycosylated proteins.

mammalian hosts and to negotiate different immune assaults while migrating through various tissues. Of note, internal (parasite organs and cells) and external (host tissues) environments play a defining role in the expression of enzymes involved in glycosylation and monosaccharide (substrate) availability that could greatly affect the glycan type and configuration (87). The global heterogeneity (or macro-heterogeneity) of NEJs glycoproteins encompass an array of 53 *N*-glycans and 16 *O*-glycan structures that are employed in various combinations to decorate all 123 NEJs glycoproteins identified (Supplementary Material 9). Further specific investigations are required to elucidate the required enzymatic network for *F. hepatica* NEJs to create such a complex glycan array to decorate its somatic and secreted glycoproteins.

Glycosylation Drives *F. hepatica* NEJs Glycoprotein Localization

The glycan attached to a protein likely affects more than just its biology but also the cellular compartment where it will be expressed. Here, we observed that oligomannose *N*-glycans are more frequently found in Som proteins, whereas paucimannose and hybrid/complex *N*-glycans are relatively more abundant within the ES. In addition, as noted earlier, fewer ES glycoproteins were modified by PC-glycans (Fig. 8A). In terms of *O*-glycosylation, the Som proteins were associated with the small Tn-antigen, which tended to be pentosylated, while core 1- and 2-type *O*-glycans are predominantly present in the ES (Fig. 8B).

The presence of diverse glycan structures at the different glycosites of a given protein is curious because these sites would be expected to be subjected to the same assembly line of glycosyltransferases and glycosidases during processing in the ER and Golgi apparatus. Glycosite accessibility and the structure around the signature glycosite motifs might explain such heterogeneity (87, 98) and thus we assessed these properties within the glycoproteins identified in the present study. The relative surface accessibility (*i.e.*, high exposure, low exposure or buried), type of secondary structure where it is placed (*i.e.*, coil, helix, or strand) and disorder level of the region (*i.e.*, low, medium or high) were used to evaluate how accessible the individual glycosites were to the enzymatic glycosylation machinery (87, 99). Our results revealed that the majority of sites (60%) are in high exposed regions within coil structures, and only 11% of the glycosylated sites are in buried regions (Fig. 9, Supplementary Material 6). The surface accessibility analysis suggests that the probability of a glycosite being glycosylated is higher when located on the coil and surface of the proteins.

Glycosite accessibility on the protein surface has also been associated with efficient glycan processing and consequently with the presence of more complex structures, including, for example, core-fucosylation of *N*-glycans. On the other hand, relatively occluded sites carry under-processed oligomannose (Man₅₋₆GlcNAc₂) and afucosylated-core *N*-glycans, likely due

to their restricted access to enzymes during processing in the Golgi apparatus (99). Accordingly, the analysis of the glycan forms associated with buried glycosites of NEJs glycoproteins revealed that 67% are in helix or strand structures and that the vast majority of these (93%) are *N*-glycosites modified by oligomannose (62.5%) or paucimannose (25%) structures. Curiously, the few hidden glycosites that contain hybrid/complex *N*-glycans (12.5%) carry short-extended arms and afucosylated structures (*i.e.*, Man₃GlcNAc₃₋₄), which do require a certain level of processing. The presence of these buried glycosites in proteins of both Som and ES preparations suggests that they are important at the structural level, driving the correct folding of the protein, as opposed to contributing to the overall antigenicity of the protein. In contrast, exposed glycosites may be expected to play roles in protein-protein interaction, biological activity, antigenicity, and host-parasite communication; an example of this is the enzyme myeloperoxidase (MPO) in neutrophils, which when carrying hyper-truncated *N*-glycans on exposed glycosites endow on higher enzyme activity, increased thermal stability, polypeptide accessibility, and increased potential for ceruloplasmin-mediated inhibition (99).

Conservation of *F. hepatica* NEJs Glycosites Across Eukaryotes

To better comprehend the observed heterogeneity within the glycosites of NEJs proteins, we used a hierarchical cluster analysis to evaluate the motifs associated with *N*- and *O*-glycosylation. Our data show that the majority of *N*-glycosylated sites analysed contained the previously characterized canonical motif N-x-T/S, where “x” is any amino acid but proline, consistent with studies in nematodes (46, 100). However, *F. hepatica* NEJs glycoproteins display a clear preference for a Thr (T) residue in the position +2, with the motif N-x-T being present in 70% of the *N*-glycopeptides. In contrast an *O*-glycan-specific motif could not be identified within the current NEJ glycoproteins. Our results indicate that *F. hepatica* and mammalian oligosaccharyltransferase (OST), the enzyme responsible for the first step of *N*-glycosylation, have similar specificity. Of note, a distinct consensus sequence (N-x-T-x-x-x-L) has been identified and assigned to 22 *N*-glycosites belonging to 19 NEJ glycoproteins (8 in ES and 18 in Som; Supplementary Material 9). In our analysis, this leucine-containing motif does not apparently affect the glycosites’ accessibility or glycosylation, which is quite heterogeneous and is found in proteins such as cathepsin L3_4, legumain 2, and several uncharacterized proteins.

As analysis of glycosylation is often reliant on *in silico* prediction tools originally devised for mammalian proteins, once the canonical consensus motifs for *N*-glycosylation were confirmed in *F. hepatica* Som and ES glycoproteins, we sought to ascertain how appropriate these tools are for predicting glycosylation sites within parasite proteins. Comparative analyses between our glycopeptide analyses with the

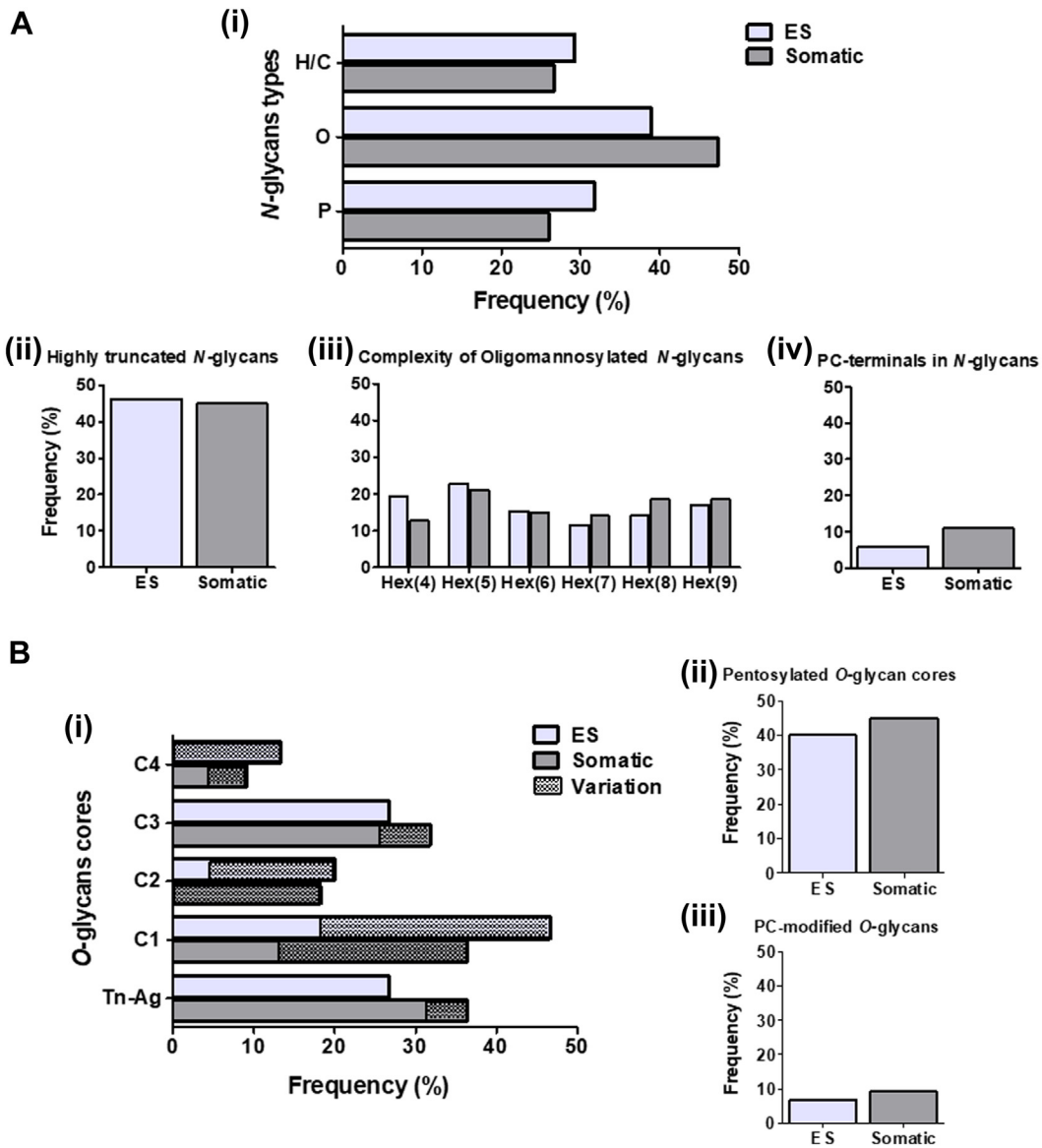


FIG. 8. Glycan complexity of *F. hepatica* NEJs glycoproteins expressed in different compartments. A, (i) frequency with which each N-glycan-type are found attached to glycoproteins in different NEJs extracts. P: paucimannose; O: oligomannose; H/C: hybrid/complex-type; the graphics show the frequency with which (ii) truncated, (iii) oligomannose and (iv) PC-modified N-glycans occur in the NEJs ES and Somatic extracts. B, (i) Frequency with which each O-glycan core and Tn-antigen (Tn-Ag) are found attached to glycoproteins present in ES and Somatic extracts. C1-C4: Core 1, 2, 3 and 4, respectively. Frequency of (ii) pentosylated or (iii) PC-modified O-glycans within the NEJs ES and Somatic extracts.

online N- and O-glycosylation prediction tools revealed that ~40% of all predicted N-glycosites are in fact N-glycosylated. Of note, for 10 (8%) N-glycoproteins the glycosylated sites predicted differed from those actually glycosylated, whilst three N-glycoproteins were predicted as non-glycosylated. The comparative data for O-glycosylation was poor, with only ~2% of the predicted O-glycosylated proteins being truly glycosylated (Supplementary Material 6). These discrepancies further highlight the limitations of methods and information being used for the current *in silico* prediction tools and that further in-depth analyses are required to elucidate and refine

helminth-related glycosite signatures. However, it is important to note that the observed inconsistency between predicted and occupied sites could have been affected by limitations in our glycopeptide analyses as well. A more thorough analysis would be necessary to screen each individual glycoprotein for its glycosylated sites.

F. hepatica NEJs Express Unique Glycoproteins With Potentially Important Host Invasion Functions

Several proteomic investigations of NEJs proteins have been reported (21, 101–103) but these have overlooked glycan/

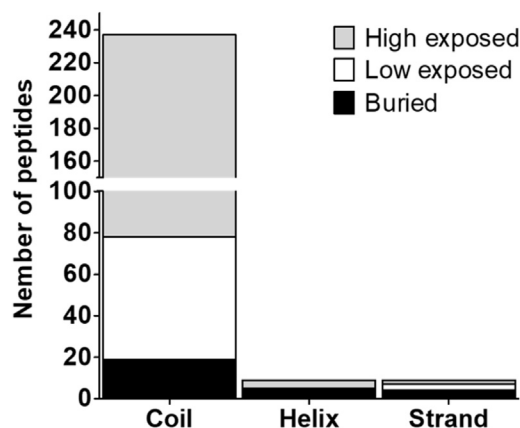


FIG. 9. Accessibility of the glycosylated sites present in *F. hepatica* NEJs glycoproteins. Glycopeptide location was categorized according to the secondary protein structure, namely, coil, helix, or strand, and relative surface accessibility (high exposed, low exposed, and buried) represented on a grey scale by many glycopeptides.

glycoprotein analysis due to the lack of appropriate tools and sample type. Here, our evaluation of the glycan profile of major NEJ glycoproteins revealed a distinct list of 25 proteins uniquely modified by glycans containing motifs such as PC, multi-PC terminal, and pent (Table 2). These included the

virulence-associated cathepsin peptidases, and a number of major uncharacterized and CUB proteins, all of which were identified in Som and/or ES preparations. Although the glycosylation state/glycan profile of these proteins are unknown in other systems, some of these listed NEJs glycoproteins (Table 2) have had their presence and function described in other organisms, and investigation of protein function conservation across organisms may shed light on how the NEJs establish infection. Key examples include, (a) the M60 domain-containing metalloprotease that targets mucosal glycoproteins such as mucins, a main component of the protective barrier that prevents gut infection by bacteria and parasites, allowing commensal and pathogenic organisms to colonize the host digestive tract (104); (b) tenascins that keep intestinal cells separated by binding to epidermal growth factor (EGF) receptors on the surface of these cells (105, 106); (c) hemicentin-1 involved in regulating cellular positioning and attachment (107, 108). These proteins could be involved in aiding the NEJs in their invasion across the intestinal wall by breaking up the mucus in the intestine, allowing the NEJs to access and attach to the epithelial cell layer, which could be manipulated by the parasite tenascins (109).

Apart from SCP domain-containing proteins, very little is known about the other glycoproteins listed that display the distinctive pentosylated O-glycans. SCP proteins are

TABLE 2
F. hepatica NEJs glycoproteins carrying unusual N- or O-glycan structures

Name (accession)	Extract	Glycan
Uncharacterized protein (BN1106_s574B000140)	ES	O-PC
Cathepsin B2 (FhCB2) (BN1106_s4482B000044)	ES/Som	Multiple PC
Cathepsin B11 (FhCB11) (maker-scaffold10x_889_pilon-snap-gene-0.33)	ES/Som	Multiple PC
Legumain 2 (BN1106_s2087B000065)	ES/Som	Multiple PC
CUB domain-containing protein (BN1106_s7704B000009)	ES/Som	Multiple PC
Uncharacterized protein - CUB (BN1106_s5172B000090)	ES/Som	Multiple PC
Acid sphingomyelinase phosphodiesterase 3a (BN1106_s1285B000159)	ES/Som	Multiple PC
Uncharacterized protein (AOA4E0R0K3)	ES/Som	PC
Uncharacterized protein - CD59-like (BN1106_s5246B000010)	ES/Som	PC
Uncharacterized protein - CUB (BN1106_s114B000622)	ES/Som	PC
CUB domain-containing protein (maker-scaffold10x_280_pilon-snap-gene-0.116)	ES/Som	PC
Laminin subunit beta (BN1106_s1357B000110)	ES/Som	PC
Low-density lipoprotein receptor (Ldl) (BN1106_s60B000536)	ES/Som	PC
Basement membrane-specific heparan sulfate proteoglycan core protein (BN1106_s25B000189)	ES/Som	PC
SCP domain-containing protein/Peptidase inhibitor 16 (BN1106_s10890B000012)	ES/Som	O-GalNAc-Pent
Tenascin (BN1106_s115B000510)	Som	O-PC
Uncharacterized protein (AOA4E0RHQ9)	Som	O-GalNAc-Pent
Uncharacterized protein (BN1106_s1308B000129)	Som	O-GalNAc-Pent
Uncharacterized protein (maker-scaffold10x_2189_pilon-snap-gene-0.8)	Som	O-GalNAc-Pent
CUB domain-containing protein (AOA4E0RX49)	Som	O-GalNAc-Pent
Cathepsin L4 (FhCL4_1) (BN1106_s6995B000048)	Som	Multiple PC
Hemicentin-1 (BN1106_s9952B000009)	Som	Multiple PC
Peptidase M60 domain-containing protein (BN1106_s2007B000402)	Som	Multiple PC
Uncharacterized protein (BN1106_s3656B000030)	Som	Multiple PC
Putative sodium/potassium-transporting ATPase subunit beta-3 (BN1106_s3547B000116)	Som	PC

ES, Excretory-secretory extract; O-GalNAc-Pent, canonical O-glycan core (GalNAc) modified by a pentose; O-PC, PC motif associated to an O-glycan; PC, Phosphorylcholine; Som, Somatic extract.

Complete glycan structures associated to each glycoprotein can be verified in the [Supplementary material 6](#).

cysteine-rich secretory proteins associated with helminth parasitism due to their abundance, secretion, and immunomodulatory properties (110, 111). The significant and specific up-regulation of SCP expression by NEJs suggests they are essential for liver fluke establishment in mammalian hosts (110). Future characterization of these proteins and their unique O-glycan might reveal specific mechanisms regarding such sugar structures and their involvement in host invasion and immune evasion.

Several parasite proteins were observed to carry N-glycans that are highly truncated or with indefinite linkages. For example, Man₄GlcNAc₂Fuc₁ was identified as attached to 22 glycoproteins, mainly to uncharacterized proteins containing Cubilin domains (N = 12). Interestingly, this sugar is predominantly associated with the secreted versions of these proteins or, in different cases, is attached to predicted membrane proteins (*i.e.*, phospholipases, tetraspanin), suggesting key roles in host-parasite interactions such as cell glycan-receptor recognition.

F. hepatica NEJs Secreted Glycoproteins

Most of the excreted-secreted glycoproteins are also present in the Som extract (Fig. 6); however, those identified in the Som and the ES preparations differed significantly in terms of glycosylation (Supplementary Material 9) and varied both in the number of occupied glycosites and in the occupancy in terms of glycan forms and complexity. This selective and distinct glycosylation of non-secreted and secreted forms of the same protein opens up an intriguing avenue to be explored in *F. hepatica* biology. It is worth noting that the glycosylation of proteins secreted by NEJs are less heterogeneous than that observed for the Som molecules. Generally, we found that a given glycosite of a protein in the Som extract would be modified with double the number of N-glycan forms compared to the same site present in the ES protein copies (Supplementary Material 9). Lee *et al.* (98) described how differences in the glycosylation profile of a protein can be explained by the differential solvent accessibility of glycosites, which was further associated with a distinct subcellular location of the protein. The mechanisms of *F. hepatica* protein trafficking and processing of glycans during or after protein glycosylation are still unknown but are likely related to the type and extent of glycosylation. The NEJs glycoproteins studied here are predicted to localize in various compartments including the extracellular space (1%), cytoplasm (32%), plasma membrane (35%), and unknown (32%) (Supplementary Material 6), which indicates defined subcellular trafficking and distinct processing and sorting.

We verified that 88 (72%) of the NEJs glycoprotein sequences are preceded by an N-terminal signal peptide for guiding secretion and, unexpectedly, included ~75% of the Som glycoproteins (Supplementary Material 6). The absence of a signal peptide in 19 (14%) of the ES glycoproteins indicates that many liver fluke proteins are secreted *via* alternative pathways (although it is also important to note that poor

annotation of some sequences might have compromised the analysis). Nevertheless, many glycoproteins identified in the ES extract have been previously reported as cargo molecules of EVs released by *F. hepatica* NEJs and adults (29, 60). Encasing the glycoproteins modified by more immunogenic glycan structures within EVs may represent a key *F. hepatica* immune evasion strategy and an effective way to guarantee that these glycoconjugated molecules reach more distant tissues and cells within the host, where they exert their primary function. The presence of numerous glycosylases in these EVs (60) could suggest that the intra-vesicle environment allows further glycan processing before releasing these molecules into the host.

A limited number of nine glycoproteins were found to be exclusive to the ES products of NEJs. Out of the total 71 ES glycoproteins identified in this study, 30 were not associated with *F. hepatica* NEJs secretions by proteomic analysis previously performed by our group (25, 34); Supplementary Materials 6 and 9), suggesting that glycopeptide analysis can complement LC-MS/MS methods for protein identification in complex samples. Of interest, 22 glycoproteins exclusively associated with the NEJs Som extract in this study have been previously associated to ES samples. Such discrepancies might reflect differences in the glycosylation state of such proteins when secreted and, therefore, the absence of glycosylated peptides to be identified in the ES. However, it is worth noting that the low abundance of some proteins in the ES may limit their glycopeptide detection.

Complementary gene ontology (GO) enrichment analysis of the 71 ES glycoproteins showed an association with the terms related to metabolic process, proteolysis, and regulation of catalytic activity. Important GO terms associated with proteolytic activity (GO:0006508, proteolysis; GO:0016787, hydrolase; GO:0008233, peptidase activity; GO:0016798, hydrolase activity on glycosyl bonds; GO:0008234, cysteine-type peptidase activity; GO:0050790, regulation of catalytic activity) and carbohydrate metabolic processes (GO:0005975) were amongst the most enriched terms for ES molecules and reflect the results of the pathway enrichment analysis for all NEJs glycoproteins (Supplementary material 10), highlighting a predominance of glycoproteins that function as hydrolases, glycosyltransferases and glycosidases in the NEJs secretions. Considering the dearth of knowledge concerning the glycosylation state of parasite proteins, our data provides unique information to progress the development of liver fluke diagnostics and vaccines, especially those that involve some well-known targets such as cathepsin L and B peptidases, cystatin/stefin, serpin, and saposins, which we now know are glycosylated in various ways (Supplementary material 6).

F. hepatica Virulence-Associated Peptidases are Highly Glycosylated

We have previously described the importance of cathepsin L and B cysteine peptidases in virulence, invasion, migration,

and feeding by *F. hepatica* parasites. The parasite employs a range of cathepsin L and B peptidases that concertedly degrade cells and macromolecules to invade and migrate through host tissues, as exemplified by their abundance within the NEJ ES proteome (21, 34, 97, 101, 102). Similarly, abundant within the NEJs ES proteome are the asparaginyl endopeptidases (aka legumains) that are important for the regulation of cathepsin peptidase activity by the trans-activation of the pro-cathepsin zymogen peptidase *via* cleavage and removal of the inhibitory propeptide domain (101, 112). Interrogation of *F. hepatica* genomic and transcriptomic datasets has revealed that the cysteine peptidases belong to multi-membered gene families that are differentially transcribed during the liver fluke life cycle (101). Specific members of these peptidase families that are known to be abundantly secreted by the infective NEJs parasites include the cathepsin B peptidases, FhCB1, FhCB2, and FhCB3 and the cathepsin L peptidase, FhCL3, which is encoded by four genes. The present analysis of the Som and ES preparations glycoprotein data revealed nine glycosylated cathepsin peptidases, including a number of cathepsin B peptidases

(FhCB1, FhCB2, FhCB3, FhCB9, FhCB11), FhCL4 and a previously uncharacterized cysteine peptidase and two legumains (legumain-2 and legumain-like) (Supplementary material 6). Interestingly, while FhCL3s were also glycosylated, this was only observed for two of the four members of the FhCL3 group, namely FhCL3_2 and FhCL3_4.

The NEJs cathepsin peptidases are mostly decorated by paucimannose (some of them highly truncated) and oligomannose *N*-glycans, which we found located at up to three glycosites within the mature domain and distant from the active site pocket. *O*-glycans were only associated with FhCL3_4 (Fig. 10). Complex structures were only observed in FhCL4, FhCB2, FhCB11 and legumain-2, which were modified by multiple PC residues (Fig. 11 and Table 2). Our results complement previous data by Garcia-Campos *et al.* (25) that showed that NEJs FhCB1 and FhCL3_4 carry paucimannose *N*-glycans (*i.e.*, Man₂GlcNAc₂). However, we found that the FhCB1 glycosite (Asn80) is quite heterogeneous and may be decorated with glycans that are different from Man₂GlcNAc₂. Similar heterogeneity is also observed for FhCL3_4; several glycan forms of the protein can result depending on the *N*-

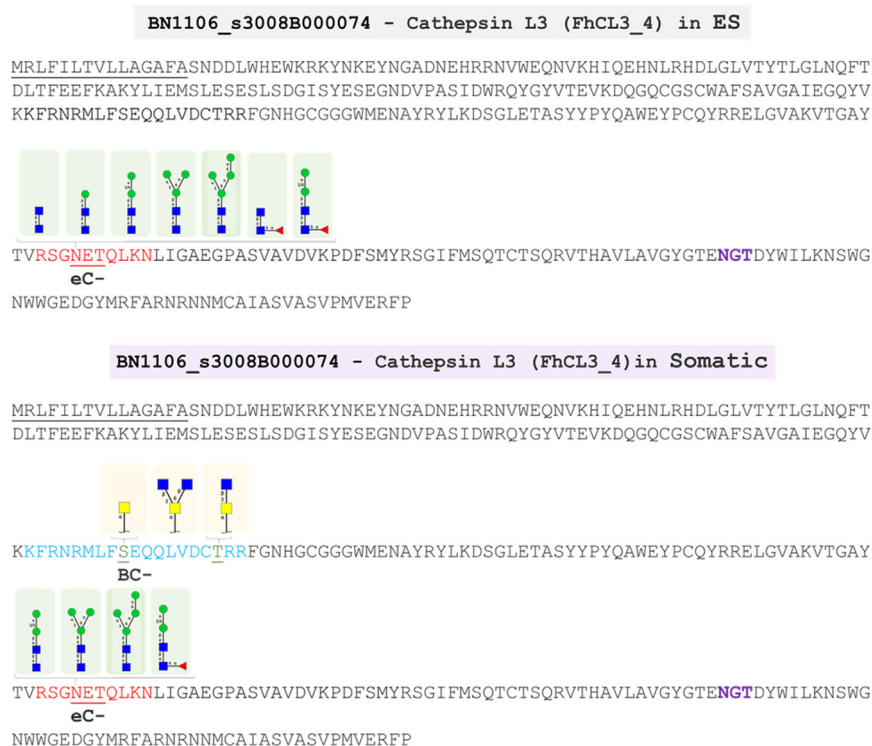


FIG. 10. Variability in glycosylation within the *F. hepatica* NEJs cysteine peptidase cathepsin L3_4. *F. hepatica* Cathepsin L3 (FhCL3_4) was present in both ES and Somatic extracts and exhibited a variable glycan profile. Underlined amino acids highlight the signal peptide. Residues in bold and purple were predicted *N*-glycosylated sites (NetNGlyc 1.0 Server) but were not shown to be glycosylated. Residues in red depict the *N*-glycopeptide identified by our glycopeptide analysis; the underlined residues mark the *N*-glucosite occupied and not predicted. Residues in blue depict the *O*-glycopeptide identified by glycopeptide analysis; the *underlined green* residues mark the two possible *O*-glycosylated sites that were not predicted. Accessibility of the glycosylated glycosites is highlighted: *e.g.* low exposed; B: buried; C, coil; -, region of low disorder. The *N*- and *O*-glycan forms identified at the respective glycosylated sites are highlighted in *green* and *yellow*, respectively. The brackets indicate the different glycan forms that can occupy the identified glycosylated site, in different copies of the same protein in each extract analyzed.

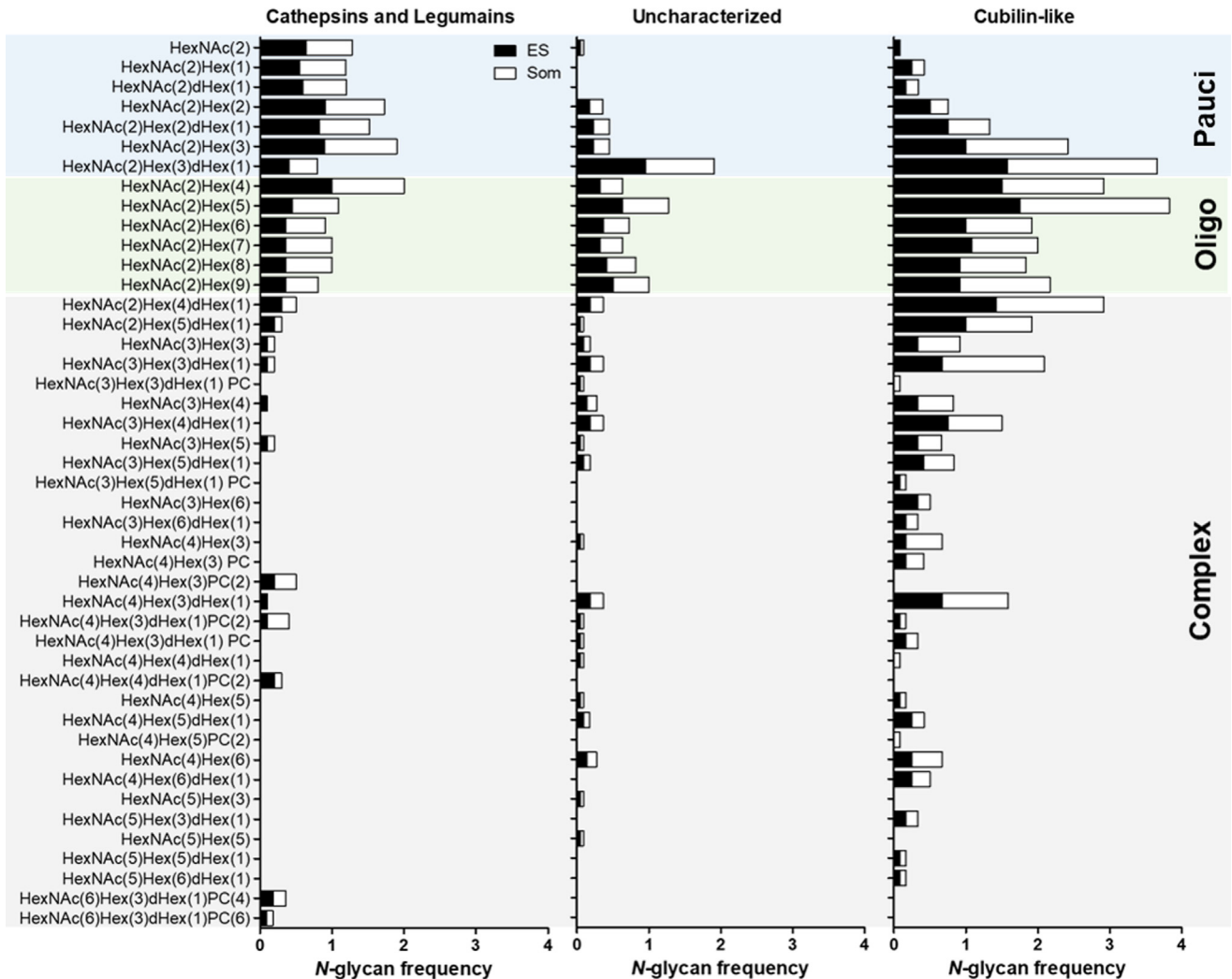


FIG. 11. Frequency of different N-glycans in glycoproteins calculated by dividing the total occurrence (frequency) of a determined N-glycan form by the number of proteins in each group. N-glycan types: Pauci: paucimannose; Oligo: oligomannose; Complex: hybrid or complex structures.

glycan form attached to Asn225 site, which differs depending on whether the protein is secreted or not (ES dataset: seven forms; Som dataset: four forms; Fig. 10). Conversely, the O-glycosite (Ser156/T164) is only occupied in the FhCL3_4 glycoprotein identified within the Som extract and appears to be modified by one or two glycan forms (HexNAc₃ or HexNAc and HexNAc₂).

Glycosylation has been shown to impact the activation and activity of cysteine proteases (3, 113). While detailed structural and biochemical analyses would be required to confirm the functional impact of the glycosylation on these peptidases, we have identified glycosylated sites within key domains for several *F. hepatica* peptidases. In legumain 2, the glycosylated site (Asn103) is located within one of the hemoglobinase domains, while FhCB1 glycosylation occurs at the junction between the propeptide and mature domain, suggesting it could

play a role in enzymatic activation, perhaps by differential interaction with the transactivating enzyme legumain.

Many *F. hepatica* NEJ Glycoproteins Remain Uncharacterized

A high proportion of the gene models predicted from the *F. hepatica* genome remain uncharacterized. Also, in this study, 25 of the NEJ glycoproteins identified in both Som and ES preparations have no known function (Supplementary material 6), albeit *in silico* analysis revealed they are possibly membrane proteins, which may explain this lack of information, as membrane proteins are difficult to isolate and characterize (114). Most of the uncharacterized NEJs glycoproteins ($n = 22$) were identified within the Som extract, indicating that they either participate in internal parasite processes or on the NEJs' surface. Interestingly, 13 of these

glycoproteins were also observed in the proteomic dataset derived from NEJs ES products and may be part of the cargo of NEJs-specific EVs (Supplementary materials 6 and 11). While some of these glycoproteins display unique glycan structures, they are predominantly decorated with oligo-mannose glycans (Fig. 11).

Bioinformatic interrogation of these uncharacterized proteins revealed that three proteins contain CD59-like complement regulatory domains, suggesting their potential role in immune modulation of host immune cells. Five uncharacterized glycoproteins were found to have cubilin (CUB) domains. Cubilin proteins, also known as intrinsic factor receptors, facilitate the uptake of nutrients such as albumin, transferrin and vitamin D from the host (115). In total, 10 NEJ glycoproteins in this study were found to contain CUB domains, comparable with previous proteomics studies of the NEJs Som and ES proteomes (34, 102). Our data shows that these NEJs CUB-domain containing proteins are multi-glycosylated, containing on average three occupied glycosites, which are heterogeneous and can be decorated with quite complex *N*-glycans containing multi-PC terminal, as well as *O*-glycans modified with pentose (Table 2 and Supplementary material 9). Interestingly, four of these proteins also contain an ovochymase-related domain (BN1106_s5172B000090, BN1106_s114B000614, BN1106_S1110B000106, and A0A4E0RX49), which has previously been associated with serine peptidase (chymotrypsin-like) activity, indicating a potentially new functional role for these proteins in hydrolytic reactions (116).

CONCLUSIONS

Profiling the glycosylation of individual *F. hepatica* NEJs proteins is an essential step to fully understand the biological processes behind infection, invasion, virulence, and immune escape that allows these parasites to survive in host tissues. This is especially pertinent for the early-stage infective NEJs stage investigated here, when the parasite is microscopic, depending on glycogen reserves, and most susceptible to the soluble and cellular effectors of the innate and adaptive host immune system. Here we shed light on the essential biology of *F. hepatica* by demonstrating the parasite forte for both common and unique mechanisms for protein glycosylation beyond what the genomic and transcriptomic data could provide. However, we acknowledge that the approach taken in this study may have not identified some glycoproteins, glycosites, and/or glycans associated to *F. hepatica* NEJs, primarily due to analytical limitations of the methods and sample availability. New or improved glycopeptide analytical techniques might allow the mapping of these and other structures in the near future.

The data we generated in the present study will not only instigate the research community to pursue the functions of these glycans in the many interactions between parasite and host but also prompt the characterization of genes and enzymes (glycosidases and glycosyltransferases) involved in the

processes required to form these structures. This information can lead to the discovery of targets to be explored to uncover new strategies to control fascioliasis. We have shown that, *via* glycosylation, NEJs create great protein heterogeneity and variability and, unexpectedly, we found that by our current approach, the ES extract is far more complex than anticipated by previous proteomic analysis. Information on the glycosylation status and profile of individual antigens released by NEJs will direct the development of diagnosis and vaccines, as observed with other parasite studies (12, 25, 83, 117–119).

DATA AVAILABILITY

The data generated as part of this study are provided in the main article and Supplementary materials, with the source data accessible *via* the relevant repositories. The annotated glycomic data were submitted to the Unicarb-DR databatase (<https://unicarb-dr.glycosmos.org/references/524>) according to MIRAGE guidelines. Raw data were uploaded on Glycopost (<https://glycopost.glycosmos.org/preview/357229270647997907659a>), accessed on July, 2023. The NEJs proteomic and glycoproteomic data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifiers PXD007255 and PXD045488, respectively.

Supplemental data—This article contains [supplemental data](#) (25, 34, 60).

Acknowledgments—The authors acknowledge the facilities and scientific and technical assistance of the Centre for Microscopy & Imaging at the University of Galway (<https://imaging.universityofgalway.ie/imaging/>).

Funding and additional information—This work was supported by a Science Foundation Ireland (SFI) Professorship grant (17/RP/5368) awarded to Prof John P. Dalton (<https://www.sfi.ie/>), and by the Irish Research Council (RCS1904) awarded to Dr Carolina De Marco Verissimo.

Author contributions—C. D. M. V. and J. P. D. conceptualization; C. D. M. V., K. C., J. N., E. M., C. J., and N. G. K. methodology; C. D. M. V. validation; C. D. M. V. formal analysis; C. D. M. V., K. C., J. N., and C. J. data curation; C. D. M. V. writing-original draft preparation; C. D. M. V. and J. P. D. financial acquisition.

Conflict of interest—The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations—The abbreviations used are: ACN, acetonitrile; Asn, Asparagine; B4GALT2, Beta-1,4-Galactosyltransferase 2; C1-C4, Core 1, 2, 3 and 4; CB, cathepsin B; CL,

cathepsin L; CL3, cathepsin L3; ConA, Concanavalin A; CUB, cubilin; ECL, Erythrina cristagalli; ER, endoplasmic reticulum; ES, excretory/secretory extract; EVs, extracellular vesicles; FA, formic acid; FASP, filter-aided sample preparation; Fh, *Fasciola hepatica*; FITC, fluorescein isothiocyanate; Gal, Galactose; GalNAc, N-acetylgalctosamine; GALNT10, Polypeptide N-acetylgalactosaminyltransferase 10; GlcNAc, N-acetylglucosamine; GOA, Gene Ontology Annotation; GSL-I, *Griffonia simplicifolia*; HCD, higher-energy collision dissociation; Hex, Hexose; HexNAc, N-Acetylhexosamine; HILIC, hydrophilic interaction liquid chromatography; LacdiNAc, N,N-Diacetyllactosamine; LacNAc, N-Acetyllactosamine; LC-ESI-MS/MS, liquid chromatograph-electrospray ionization tandem mass spectrometry; Man, Mannose; NEJs, Newly excysted juveniles; PC, phosphorylcholine; PE, phosphatidylethanolamine; Pent, Pentose; PNA, *Arachis hypogaea*; PVDF, polyvinylidene difluoride membrane; Ser/Thr, Serine/Threonine; Som, somatic extract; TFA, trifluoroacetic acid; WGA, *Triticum vulgare*.

Received July 24, 2023, and in revised form, November 17, 2023
Published, MCPRO Papers in Press, November 21, 2023, <https://doi.org/10.1016/j.mcpro.2023.100684>

REFERENCES

- Varki, A., Freeze, H. H., and Gagneux, P. (2009) Evolution of glycan diversity. In: *Essentials of Glycobiology*, 2nd ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., et al. (2022) *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Riley, N. M., Hebert, A. S., Westphall, M. S., and Coon, J. J. (2019) Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. *Nat. Commun.* **10**, 1311
- Gagneux, P., Hennet, T., and Varki, A. (2022) Biological functions of glycans. In: *Essentials of Glycobiology*, 4th ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Cummings, R. D. (2019) Stuck on sugars – how carbohydrates regulate cell adhesion, recognition, and signaling. *Glycoconj. J.* **36**, 241–257
- Nyame, A. K., Kawar, Z. S., and Cummings, R. D. (2004) Antigenic glycans in parasitic infections: implications for vaccines and diagnostics. *Arch. Biochem. Biophys.* **426**, 182–200
- van Die, I., and Cummings, R. D. (2010) Chapter 15-glycomics in unraveling glycan-driven immune responses by parasitic helminths. In: Cummings, R. D., Pierce, J. M., eds. *Handbook of Glycomics*, Academic Press, San Diego: 367–396
- van Die, I., and Cummings, R. D. (2010) Glycan gimmickry by parasitic helminths: a strategy for modulating the host immune response? *Glycobiology* **20**, 2–12
- Verissimo, C. M., Graeff-Teixeira, C., Jones, M. K., and Morassutti, A. L. (2019) Glycans in the roles of parasitological diagnosis and host-parasite interplay. *Parasitology* **146**, 1217–1232
- Smit, C. H., van Diepen, A., Nguyen, D. L., Wuhrer, M., Hoffmann, K. F., Deelder, A. M., et al. (2015) Glycomic analysis of life stages of the human parasite *Schistosoma mansoni* reveals developmental expression profiles of functional and antigenic glycan motifs. *Mol. Cell. Proteomics* **14**, 1750–1769
- Reason, A. J., Ellis, L. A., Appleton, J. A., Wisniewski, N., Grieve, R. B., McNeil, M., et al. (1994) Novel tyvelose-containing tri- and tetra-antennary N-glycans in the immunodominant antigens of the intracellular parasite *Trichinella spiralis*. *Glycobiology* **4**, 593–603
- North, S. J., Botchway, K., Doonan, J., Lumb, F. E., Dell, A., Harnett, W., et al. (2019) Site-specific glycoproteomic characterization of ES-62: the major secreted product of the parasitic worm *Acanthocheilonema viteae*. *Glycobiology* **29**, 562–571
- Caravedo, M. A., and Cabada, M. M. (2020) Human fascioliasis: current epidemiological status and strategies for diagnosis, treatment, and control. *Res. Rep. Trop. Med.* **11**, 149–158
- Howell, A., Baylis, M., Smith, R., Pinchbeck, G., and Williams, D. (2015) Epidemiology and impact of *Fasciola hepatica* exposure in high-yielding dairy herds. *Prev. Vet. Med.* **121**, 41–48
- Charlier, J., Vercruysse, J., Morgan, E., van Dijk, J., and Williams, D. J. (2014) Recent advances in the diagnosis, impact on production and prediction of *Fasciola hepatica* in cattle. *Parasitology* **141**, 326–335
- Mas-Coma, S., Valero, M. A., and Bargues, M. D. (2009) Chapter 2. *Fasciola*, Iymnaeids and human fascioliasis, with a global overview on disease transmission, epidemiology, evolutionary genetics, molecular epidemiology and control. *Adv. Parasitol.* **69**, 41–146
- Gonzalez-Miguel, J., Becerro-Recio, D., and Siles-Lucas, M. (2021) Insights into *Fasciola hepatica* juveniles: crossing the fasciolosis rubicon. *Trends Parasitol.* **37**, 35–47
- Cwiklinski, K., O'Neill, S. M., Donnelly, S., and Dalton, J. P. (2016) A prospective view of animal and human fasciolosis. *Parasite Immunol.* **38**, 558–568
- Cwiklinski, K., De Marco Verissimo, C., McVeigh, P., Donnelly, S., and Dalton, J. P. (2021) Applying “Omics” Technologies to Understand *Fasciola Spp. Biology*. CABI International, Egham, United Kingdom
- Ryan, S., Shiels, J., Taggart, C. C., Dalton, J. P., and Weldon, S. (2020) *Fasciola hepatica*-derived molecules as regulators of the host immune response. *Front. Immunol.* **11**, 2182
- Robinson, M. W., Menon, R., Donnelly, S. M., Dalton, J. P., and Ranganathan, S. (2009) An integrated transcriptomics and proteomics analysis of the secretome of the helminth pathogen *Fasciola hepatica*: proteins associated with invasion and infection of the mammalian host. *Mol. Cell. Proteomics* **8**, 1891–1907
- Hillyer, G. V., Soler de Galanes, M., Rodriguez-Perez, J., Bjorland, J., Silva de Lagrava, M., Ramirez Guzman, S., et al. (1992) Use of the Falcon assay screening test–enzyme-linked immunosorbent assay (FAST-ELISA) and the enzyme-linked immunoelectrotransfer blot (EITB) to determine the prevalence of human fascioliasis in the bolivian altiplano. *Am. J. Trop. Med. Hyg.* **46**, 603–609
- Hanna, R. E. B. (1980) *Fasciola hepatica*: glycocalyx replacement in the juvenile as a possible mechanism for protection against host immunity. *Exp. Parasitol.* **50**, 103–114
- Hanna, R. E. B., and Trudgett, A. G. (1983) *Fasciola hepatica*: development of monoclonal antibodies and their use to characterize a glycocalyx antigen in migrating flukes. *Parasite Immunol.* **5**, 409–425
- Garcia-Campos, A., Ravid, A., Nguyen, D. L., Cwiklinski, K., Dalton, J. P., Hokke, C. H., et al. (2016) Tegument glycoproteins and cathepsins of newly excysted juvenile *Fasciola hepatica* carry mannosidic and paucimannosidic N-glycans. *PLoS Negl. Trop. Dis.* **10**, e0004688
- Garcia-Campos, A., Baird, A. W., and Mulcahy, G. (2017) Migration of *Fasciola hepatica* newly excysted juveniles is inhibited by high-mannose and oligomannose-type N-glycan-binding lectins. *Parasitology* **144**, 1708–1717
- De Marco Verissimo, C., Jewhurst, H. L., Dobó, J., Gál, P., Dalton, J. P., and Cwiklinski, K. (2022) *Fasciola hepatica* is refractory to complement killing by preventing attachment of mannose binding lectin (MBL) and inhibiting MBL-associated serine proteases (MASPs) with serpins. *PLoS Pathog.* **18**, e1010226
- Ravid, A., Cwiklinski, K., Aldridge, A. M., Clarke, P., Thompson, R., Gerlach, J. Q., et al. (2016) *Fasciola hepatica* surface tegument: glycoproteins at the interface of parasite and host. *Mol. Cell. Proteomics* **15**, 3139–3153
- Murphy, A., Cwiklinski, K., Lalor, R., O'Connell, B., Robinson, M. W., Gerlach, J., et al. (2020) *Fasciola hepatica* extracellular vesicles isolated from excretory-secretory products using a gravity flow method modulate dendritic cell phenotype and activity. *PLoS Negl. Trop. Dis.* **14**, e0008626
- McVeigh, P., Cwiklinski, K., Garcia-Campos, A., Mulcahy, G., O'Neill, S. M., Maule, A. G., et al. (2018) In silico analyses of protein glycosylating genes in the helminth *Fasciola hepatica* (liver fluke) predict protein-linked glycan simplicity and reveal temporally-dynamic expression profiles. *Sci. Rep.* **8**, 11700

31. Jensen, P. H., Karlsson, N. G., Kolarich, D., and Packer, N. H. (2012) Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc.* **7**, 1299–1310
32. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H., and Campbell, M. P. (2013) Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J. Am. Soc. Mass Spectrom.* **24**, 895–906
33. Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362
34. Cwiklinski, K., Jewhurst, H., McVeigh, P., Barbour, T., Maule, A. G., Tort, J., et al. (2018) Infection by the helminth parasite *Fasciola hepatica* requires rapid regulation of metabolic, virulence, and invasive factors to adjust to its mammalian host. *Mol. Cell. Proteomics* **17**, 792–809
35. Cheng, A., Grant, C. E., Noble, W. S., and Bailey, T. L. (2019) MoMo: discovery of statistically significant post-translational modification motifs. *Bioinformatics* **35**, 2774–2782
36. Meng, X., Xing, S., Perez, L. M., Peng, X., Zhao, Q., Redoña, E. D., et al. (2017) Proteome-wide analysis of lysine 2-hydroxyisobutyrylation in developing Rice (*Oryza sativa*) Seeds. *Sci. Rep.* **7**, 17486
37. Varki, A., Freeze, H. H., and Manzi, A. E. (2009) Overview of glycoconjugate analysis. *Curr. Protoc. Protein Sci.* **Chapter 12**, 12.1.1–12.1.10
38. Paschinger, K., and Wilson, I. B. H. (2016) Analysis of zwitterionic and anionic N-linked glycans from invertebrates and protists by mass spectrometry. *Glycoconj. J.* **33**, 273–283
39. Paschinger, K., and Wilson, I. B. (2019) Comparisons of N-glycans across invertebrate phyla. *Parasitology* **146**, 1733–1742
40. Kwame Nyame, A., Debose-Boyd, R., Long, T. D., Tsang, V. C. W., and Cummings, R. D. (1998) Expression of Lex antigen in *Schistosoma japonicum* and *S.haematobium* and immune responses to Lex in infected animals: lack of lex expression in other trematodes and nematodes. *Glycobiology* **8**, 615–624
41. Wilson, I. B., and Paschinger, K. (2016) Sweet secrets of a therapeutic worm: mass-spectrometric N-glycomic analysis of *Trichuris suis*. *Anal. Bioanal. Chem.* **408**, 461–471
42. Paschinger, K., Gonzalez-Sapienza, G. G., and Wilson, I. B. (2012) Mass spectrometric analysis of the immunodominant glycan epitope of *Echinococcus granulosus* antigen Ag5. *Int. J. Parasitol.* **42**, 279–285
43. Meevissen, M. H. J., Wuhrer, M., Doenhoff, M. J., Schramm, G., Haas, H., Deelder, A. M., et al. (2010) Structural characterization of glycans on omega-1, a major *Schistosoma mansoni* egg glycoprotein that drives Th2 responses. *J. Proteome Res.* **9**, 2630–2642
44. Meevissen, M. H., Balog, C. I., Koeleman, C. A., Doenhoff, M. J., Schramm, G., Haas, H., et al. (2011) Targeted glycoproteomic analysis reveals that kappa-5 is a major, uniquely glycosylated component of *Schistosoma mansoni* egg antigens. *Mol. Cell. Proteomics* **10**, M110.005710
45. Wuhrer, M., Balog, C. I., Catalina, M. I., Jones, F. M., Schramm, G., Haas, H., et al. (2006) IPSE/alpha-1, a major secretory glycoprotein antigen from schistosome eggs, expresses the Lewis X motif on core-difucosylated N-glycans. *FEBS J.* **273**, 2276–2292
46. Wang, C., Gao, W., Yan, S., Zhu, X.-Q., Suo, X., Liu, X., et al. (2021) N-glycome and N-glycoproteome of a hematophagous parasitic nematode *Haemonchus*. *Comput. Struct. Biotechnol. J.* **19**, 2486–2496
47. Williams, S. E., Noel, M., Lehoux, S., Cetinbas, M., Xavier, R. J., Sadreyev, R. I., et al. (2022) Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat. Commun.* **13**, 275
48. Nwosu, C. C., Aldredge, D. L., Lee, H., Lerno, L. A., Zivkovic, A. M., German, J. B., et al. (2012) Comparison of the human and bovine milk N-glycome via high-performance microfluidic chip liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **11**, 2912–2924
49. Kukuruzinska, M. A., and Lennon, K. (1998) Protein N-glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.* **9**, 415–448
50. Khoo, K.-H., Chatterjee, D., Caulfield, J. P., Morris, H. R., and Dell, A. (1997) Structural mapping of the glycans from the egg glycoproteins of *Schistosoma mansoni* and *Schistosoma japonicum*: identification of novel core structures and terminal sequences. *Glycobiology* **7**, 663–677
51. Cummings, R. D., and Nyame, A. K. (1999) Schistosome glycoconjugates. *Biochim. Biophys. Acta* **1455**, 363–374
52. Wang, F., Li, Y., Yang, C., Mu, Y., Wang, Y., Zhang, W., et al. (2019) Mannan-binding lectin suppresses peptidoglycan-induced TLR2 activation and inflammatory responses. *Mediators Inflamm* **2019**, 1349784
53. Sica, A., and Mantovani, A. (2012) Macrophage plasticity and polarization: *in vivo* veritas. *J. Clin. Invest.* **122**, 787–795
54. Davicino, R. C., Eliçabe, R. J., Di Genaro, M. S., and Rabinovich, G. A. (2011) Coupling pathogen recognition to innate immunity through glycan-dependent mechanisms. *Int. Immunopharmacol.* **11**, 1457–1463
55. Gomord, V., Fitchette, A. C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D., et al. (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol. J.* **8**, 564–587
56. Fitchette, A.-C., Cabanes-Macheteau, M., Marvin, L., Martin, B., Satiat-Jeunemaitre, B., Gomord, V., et al. (1999) Biosynthesis and immunolocalization of Lewis a-containing N-glycans in the plant cell. *Plant Physiol.* **121**, 333–344
57. Tezuka, K., Hayashi, M., Ishihara, H., Onozaki, K., Nishimura, M., and Takahashi, N. (1993) Occurrence of heterogeneity of N-linked oligosaccharides attached to sycamore (*Acer pseudoplatanus* L.) laccase after excretion. *Biochem. Mol. Biol. Int.* **29**, 395–402
58. Strasser, R., Bondili, J. S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., et al. (2007) A unique β 1,3-Galactosyltransferase is indispensable for the biosynthesis of N-glycans containing lewis a structures in *arabidopsis thaliana*. *Plant Cell* **19**, 2278–2292
59. Irwin, J., Morrissey, P., Ryan, J., Walshe, A., O'Neill, S., Carrington, S., et al. (2004) Glycosidase activity in the excretory-secretory products of the liver fluke, *Fasciola Hepatica*. *Parasitology* **129**, 465–472
60. Cwiklinski, K., de la Torre-Escudero, E., Trelis, M., Bernal, D., Dufresne, P. J., Brennan, G. P., et al. (2015) The extracellular vesicles of the helminth pathogen, *Fasciola hepatica*: biogenesis pathways and cargo molecules involved in parasite pathogenesis. *Mol. Cell. Proteomics* **14**, 3258–3273
61. Tang, F., Zhou, M., Qin, K., Shi, W., Yashinov, A., Yang, Y., et al. (2020) Selective N-glycan editing on living cell surfaces to probe glycoconjugate function. *Nat. Chem. Biol.* **16**, 766–775
62. Sanderson, R. D., Bandari, S. K., and Vlodavsky, I. (2019) Proteases and glycosidases on the surface of exosomes: newly discovered mechanisms for extracellular remodeling. *Matrix Biol.* **75–76**, 160–169
63. Haslam, S. M., Coles, G. C., Morris, H. R., and Dell, A. (2000) Structural characterization of the N-glycans of *Dictyocaulus viviparus*: discovery of the LewisX structure in a nematode. *Glycobiology* **10**, 223–229
64. Morelle, W., Haslam, S. M., Olivier, V., Appleton, J. A., Morris, H. R., and Dell, A. (2000) Phosphorylcholine-containing N-glycans of *Trichinella spiralis*: identification of multiantennary lactiNac structures. *Glycobiology* **10**, 941–950
65. Nakane, T., Angata, K., Sato, T., Kaji, H., and Narimatsu, H. (2019) Identification of mammalian glycoproteins with type-I LactiNac structures synthesized by the glycosyltransferase B3GALNT2. *J. Biol. Chem.* **294**, 7433–7444
66. Wilson, I. B. H., Yan, S., Jin, C., Dutkiewicz, Z., Rendić, D., Palmberger, D., et al. (2023) Increasing complexity of the N-glycome during *Caenorhabditis* development. *Mol. Cell. Proteomics* **22**, 100505
67. Sloan, T., Dooge, D., and Joyce, P. (1991) Identification of phosphorylcholine containing antigens of *Fasciola hepatica*—successful tolerance against this epitope in experimental animals. *Parasite Immunol.* **13**, 447–455
68. Petralia, L. M., van Diepen, A., Lokker, L. A., Nguyen, D. L., Sartono, E., Khatri, V., et al. (2022) Mass spectrometric and glycan microarray-based characterization of the filarial nematode *Brugia malayi* glycome reveals anionic and zwitterionic glycan antigens. *Mol. Cell. Proteomics* **21**, 100201
69. Harnett, W., Harnett, M. M., Leung, B. P., Gracie, A., and McInnes, I. B. (2004) The anti-inflammatory potential of the filarial nematode secreted product, ES-62. *Curr. Top. Med. Chem.* **4**, 553–559
70. Ahmed, U. K., Maller, N. C., Iqbal, A. J., Al-Riyami, L., Harnett, W., and Raynes, J. G. (2016) The carbohydrate-linked phosphorylcholine of the parasitic nematode product ES-62 modulates complement activation. *J. Biol. Chem.* **291**, 11939–11953
71. Deehan, M. R., Goodridge, H. S., Blair, D., Lochnit, G., Dennis, R. D., Geyer, R., et al. (2002) Immunomodulatory properties of *Ascaris suum* glycosphingolipids - phosphorylcholine and non-phosphorylcholine-dependent effects. *Parasite Immunol.* **24**, 463–469

72. Harnett, W., Frame, M., Nor, Z., MacDonald, M., and Houston, K. (1994) Some preliminary data on the nature/structure of the PC-glycan of the major excretory-secretory product of *Acanthocheilonema viteae* (ES-62). *Parasite* **1**, 179–181
73. Deehan, M. R., Harnett, W., and Harnett, M. M. (2001) A filarial nematode-secreted phosphorylcholine-containing glycoprotein uncouples the B cell antigen receptor from extracellular signal-regulated kinase-mitogen-activated protein kinase by promoting the surface Ig-mediated recruitment of Src homology 2 domain-containing tyrosine phosphatase-1 and Pac-1 mitogen-activated kinase-phosphatase. *J. Immunol.* **166**, 7462–7468
74. Goodridge, H. S., Harnett, W., Liew, F. Y., and Harnett, M. M. (2003) Differential regulation of interleukin-12 p40 and p35 induction via Erk mitogen-activated protein kinase-dependent and-independent mechanisms and the implications for bioactive IL-12 and IL-23 responses. *Immunology* **109**, 415–425
75. Goodridge, H. S., Marshall, F. A., Else, K. J., Houston, K. M., Egan, C., Al-Riyami, L., et al. (2005) Immunomodulation via novel use of TLR4 by the filarial nematode phosphorylcholine-containing secreted product, ES-62. *J. Immunol.* **174**, 284–293
76. Haslam, S. M., and Dell, A. (2003) Hallmarks of *Caenorhabditis elegans* N-glycosylation: complexity and controversy. *Biochimie* **85**, 25–32
77. Marshall, F. A., Grierson, A. M., Garside, P., Harnett, W., and Harnett, M. M. (2005) ES-62, an immunomodulator secreted by filarial nematodes, suppresses clonal expansion and modifies effector function of heterologous antigen-specific T cells *in vivo*. *J. Immunol.* **175**, 5817–5826
78. Rodriguez, E., Noya, V., Cervi, L., Chiribao, M. L., Brossard, N., Chiale, C., et al. (2015) Glycans from *Fasciola hepatica* modulate the host immune response and TLR-induced maturation of dendritic cells. *PLoS Negl. Trop. Dis.* **9**, e0004234
79. Walsh, K. P., Brady, M. T., Finlay, C. M., Boon, L., and Mills, K. H. (2009) Infection with a helminth parasite attenuates autoimmunity through TGF-beta-mediated suppression of Th17 and Th1 responses. *J. Immunol.* **183**, 1577–1586
80. Donnelly, S., Stack, C. M., O'Neill, S. M., Sayed, A. A., Williams, D. L., and Dalton, J. P. (2008) Helminth 2-Cys peroxiredoxin drives Th2 responses through a mechanism involving alternatively activated macrophages. *FASEB J.* **22**, 4022–4032
81. Donnelly, S., O'Neill, S. M., Sekiya, M., Mulcahy, G., and Dalton, J. P. (2005) Thioredoxin peroxidase secreted by *Fasciola hepatica* induces the alternative activation of macrophages. *Infect. Immun.* **73**, 166–173
82. Wilkinson, H., and Saldova, R. (2020) Current methods for the characterization of O-glycans. *J. Proteome Res.* **19**, 3890–3905
83. Van Dam, G. J., Bergwerff, A. A., Thomas-Oates, J. E., Rotmans, J. P., Kamerling, J. P., Vliegenthart, J. F., et al. (1994) The immunologically reactive O-linked polysaccharide chains derived from circulating cathodic antigen isolated from the human blood fluke *Schistosoma mansoni* have Lewis x as repeating unit. *Eur. J. Biochem.* **225**, 467–482
84. Bergwerff, A. A., van Dam, G. J., Rotmans, J. P., Deelder, A. M., Kamerling, J. P., and Vliegenthart, J. F. (1994) The immunologically reactive part of immunopurified circulating anodic antigen from *Schistosoma mansoni* is a threonine-linked polysaccharide consisting of $\rightarrow 6$ -(beta-D-GlcpA-(1 \rightarrow 3))-beta-D-GalpNAc-(1 \rightarrow repeating units. *J. Biol. Chem.* **269**, 31510–31517
85. Freire, T., Casaravilla, C., Carmona, C., and Osinaga, E. (2003) Mucin-type O-glycosylation in *Fasciola hepatica*: characterisation of carcinoma-associated Tn and sialyl-Tn antigens and evaluation of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase activity. *Int. J. Parasitol.* **33**, 47–56
86. Hanisch, F. G., Koldovsky, U., and Borchard, F. (1993) Monoclonal antibody 2B5 defines a truncated O-glycan, GlcNAc beta 1-3Gal beta 1-4GlcNAc beta 1-6 (GalNAc), on mucins from deep gastric and duodenal glands as well as metaplasia and neoplasia of gastric differentiation. *Cancer Res.* **53**, 4791–4796
87. Schachter, H. (1986) Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. *Biochem. Cell Biol.* **64**, 163–181
88. Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., et al. (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core α (1, 3)-fucose and core xylose. *Glycobiology* **13**, 427–434
89. Platts-Mills, T. A., Hilger, C., Jappe, U., van Hage, M., Gadermaier, G., Spillner, E., et al. (2021) Carbohydrate epitopes currently recognized as targets for IgE antibodies. *Allergy* **76**, 2383–2394
90. Faveeuw, C., Malleveay, T., Paschinger, K., Wilson, I. B., Fontaine, J., Mollicone, R., et al. (2003) Schistosome N-glycans containing core α 3-fucose and core β 2-xylose epitopes are strong inducers of Th2 responses in mice. *Eur. J. Immunol.* **33**, 1271–1281
91. Dixon, K. E. (1965) The structure and histochemistry of the cyst wall of the metacercaria of *Fasciola hepatica* L. *Parasitology* **55**, 215–226
92. Cancela, M., Santos, G. B., Carmona, C., Ferreira, H. B., Tort, J. F., and Zaha, A. (2015) *Fasciola hepatica* mucin-encoding gene: expression, variability and its potential relevance in host-parasite relationship. *Parasitology* **142**, 1673–1681
93. Cwiklinski, K., Drysdale, O., López Corrales, J., Corripio-Miyar, Y., De Marco Verissimo, C., Jewhurst, H., et al. (2022) Targeting secreted protease/anti-protease Balance as a vaccine strategy against the helminth *Fasciola hepatica*. *Vaccines (Basel)* **10**, 155
94. Zhu, Z., and Desaire, H. (2015) Carbohydrates on proteins: site-specific glycosylation analysis by mass spectrometry. *Annu. Rev. Anal. Chem. (Palo Alto Calif)* **8**, 463–483
95. Zacchi, L. F., and Schulz, B. L. (2016) N-glycoprotein macroheterogeneity: biological implications and proteomic characterization. *Glycoconj. J.* **33**, 359–376
96. Higel, F., Seidl, A., Sörgel, F., and Friess, W. (2016) N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *Eur. J. Pharm. Biopharm.* **100**, 94–100
97. Cwiklinski, K., Dalton, J. P., Dufresne, P. J., La Course, J., Williams, D. J., Hodgkinson, J., et al. (2015) The *Fasciola hepatica* genome: gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biol.* **16**, 71
98. Lee, L. Y., Lin, C.-H., Fanayan, S., Packer, N. H., and Thaysen-Andersen, M. (2014) Differential site accessibility Mechanistically explains subcellular-specific N-glycosylation determinants. *Front. Immunol.* **5**, 404
99. Tjondro, H. C., Ugonotti, J., Kawahara, R., Chatterjee, S., Loke, I., Chen, S., et al. (2021) Hyper-truncated Asn355- and Asn391-glycans modulate the activity of neutrophil granule myeloperoxidase. *J. Biol. Chem.* **296**, 100144
100. Kaji, H., Kamiie, J.-I., Kawakami, H., Kido, K., Yamauchi, Y., Shinkawa, T., et al. (2007) Proteomics reveals N-linked glycoprotein diversity in *Caenorhabditis elegans* and suggests an atypical translocation mechanism for integral membrane proteins. *Mol. Cell. Proteomics* **6**, 2100–2109
101. Cwiklinski, K., and Dalton, J. P. (2022) Exploiting comparative omics to understand the pathogenic and virulence-associated protease: anti-protease relationships in the zoonotic parasites *Fasciola hepatica* and *Fasciola Gigantica*. *Genes (Basel)* **13**, 1854
102. Di Maggio, L. S., Tirloni, L., Pinto, A. F., Diedrich, J. K., Yates Iii, J. R., Benavides, U., et al. (2016) Across intra-mammalian stages of the liver fluke *Fasciola hepatica*: a proteomic study. *Sci. Rep.* **6**, 32796
103. Di Maggio, L. S., Tirloni, L., Pinto, A. F. M., Diedrich, J. K., Yates, J. R., 3rd, Carmona, C., et al. (2019) A proteomic comparison of excretion/secretion products in *Fasciola hepatica* newly excysted juveniles (NEJ) derived from *Lymnaea viatrix* or *Pseudosuccinea columella*. *Exp. Parasitol.* **201**, 11–20
104. Nakjang, S., Ndeh, D. A., Wipat, A., Bolam, D. N., and Hirt, R. P. (2012) A novel extracellular metallopeptidase domain shared by animal host-associated mutualistic and pathogenic microbes. *PLoS One* **7**, e30287
105. Dubourg, A., Xia, D., Winpenny, J. P., Al Naimi, S., Bouzid, M., Sexton, D. W., et al. (2018) Giardia secretome highlights secreted tenascins as a key component of pathogenesis. *Gigascience* **7**, 1–13
106. Emery, S. J., Mirzaei, M., Vuong, D., Pascovici, D., Chick, J. M., Lacey, E., et al. (2016) Induction of virulence factors in *Giardia duodenalis* independent of host attachment. *Sci. Rep.* **6**, 20765
107. Sherwood, D. R., and Plastino, J. (2018) Invading, leading and navigating cells in *Caenorhabditis elegans*: insights into cell movement *in Vivo*. *Genetics* **208**, 53–78
108. Vogel, B. E., and Hedgecock, E. M. (2001) Hemicentin, a conserved extracellular member of the immunoglobulin superfamily, organizes epithelial and other cell attachments into oriented line-shaped junctions. *Development* **128**, 883–894

109. Linden, S., Sutton, P., Karlsson, N., Korolik, V., and McGuckin, M. (2008) Mucins in the mucosal barrier to infection. *Mucosal Immunol.* **1**, 183–197
110. Cantacessi, C., Hofmann, A., Young, N. D., Broder, U., Hall, R. S., Loukas, A., *et al.* (2012) Insights into SCP/TAPS proteins of liver flukes based on large-scale bioinformatic analyses of sequence datasets. *PLoS One* **7**, e31164
111. Gadahi, J. A., Wang, S., Bo, G., Ehsan, M., Yan, R., Song, X., *et al.* (2016) Proteomic analysis of the excretory and secretory proteins of *Haemonchus contortus* (HcESP) binding to goat PBMCs *in vivo* revealed stage-specific binding profiles. *PLoS One* **11**, e0159796
112. Dalton, J. P., Brindley, P. J., Donnelly, S., and Robinson, M. W. (2009) The enigmatic asparaginyl endopeptidase of helminth parasites. *Trends Parasitol.* **25**, 59–61
113. Kotze, A., and McClure, S. (2001) *Haemonchus contortus* utilises catalase in defence against exogenous hydrogen peroxide *in vitro*. *Int. J. Parasitol.* **31**, 1563–1571
114. Carpenter, E. P., Beis, K., Cameron, A. D., and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* **18**, 581–586
115. Christensen, E. I., Nielsen, R., and Birn, H. (2013) From bowel to kidneys: the role of cubilin in physiology and disease. *Nephrol. Dial. Transplant.* **28**, 274–281
116. Lindsay, L. L., Yang, J. C., and Hedrick, J. L. (1999) Ovochymase, a *Xenopus laevis* egg extracellular protease, is translated as part of an unusual polyprotease. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11253–11258
117. Nyame, A. K., Yoshino, T. P., and Cummings, R. D. (2002) Differential expression of LacdiNAc, fucosylated LacdiNAc, and Lewis x glycan antigens in intramolluscan stages of *Schistosoma mansoni*. *J. Parasitol.* **88**, 890–897
118. Ferguson, B. J., Newland, S. A., Gibbs, S. E., Tourlomousis, P., dos Santos, P. F., Patel, M. N., *et al.* (2015) The *Schistosoma mansoni* T2 ribonuclease omega-1 modulates inflammasome-dependent IL-1 β secretion in macrophages. *Int. J. Parasitol.* **45**, 809–813
119. van Dam, G. J., Odermatt, P., Acosta, L., Bergquist, R., de Dood, C. J., Kornelis, D., *et al.* (2015) Evaluation of banked urine samples for the detection of circulating anodic and cathodic antigens in *Schistosoma mekongi* and *S. japonicum* infections: a proof-of-concept study. *Acta Trop.* **141**, 198–203