

RESEARCH ARTICLE

Cryptosporidium parvum vaccine candidates are incompletely modified with O-linked-N-acetylgalactosamine or contain N-terminal N-myristate and S-palmitate

John R. Haserick^{1,2}, Joshua A. Klein³, Catherine E. Costello², John Samuelson^{1*}

1 Department of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, Boston, Massachusetts, United States of America, **2** Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts, United States of America, **3** Program for Bioinformatics, Boston University, Boston, Massachusetts, United States of America

* jsamuels@bu.edu



OPEN ACCESS

Citation: Haserick JR, Klein JA, Costello CE, Samuelson J (2017) *Cryptosporidium parvum* vaccine candidates are incompletely modified with O-linked-N-acetylgalactosamine or contain N-terminal N-myristate and S-palmitate. PLoS ONE 12(8): e0182395. <https://doi.org/10.1371/journal.pone.0182395>

Editor: Silvia N. Moreno, University of Georgia, UNITED STATES

Received: May 5, 2017

Accepted: July 17, 2017

Published: August 8, 2017

Copyright: © 2017 Haserick et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005989 and [10.6019/PXD005989](https://doi.org/10.6019/PXD005989). The data may be accessed at <http://www.ebi.ac.uk/pride/archive/projects/PXD005989>.

Funding: Support for this study came from NIH grants R01 AI110638, R01 GM031318, and UL1TR001430 (J.S.) and P41 GM104603, S10

Abstract

Cryptosporidium parvum (studied here) and *Cryptosporidium hominis* are important causes of diarrhea in infants and immunosuppressed persons. *C. parvum* vaccine candidates, which are on the surface of sporozoites, include glycoproteins with Ser- and Thr-rich domains (Gp15, Gp40, and Gp900) and a low complexity, acidic protein (Cp23). Here we used mass spectrometry to determine that O-linked GalNAc is present in dense arrays on a glycopeptide with consecutive Ser derived from Gp40 and on glycopeptides with consecutive Thr derived from Gp20, a novel *C. parvum* glycoprotein with a formula weight of ~20 kDa. In contrast, the occupied Ser or Thr residues in glycopeptides from Gp15 and Gp900 are isolated from one another. Gly at the N-terminus of Cp23 is N-myristoylated, while Cys, the second amino acid, is S-palmitoylated. In summary, *C. parvum* O-GalNAc transferases, which are homologs of host enzymes, densely modify arrays of Ser or Thr, as well as isolated Ser and Thr residues on *C. parvum* vaccine candidates. The N-terminus of an immunodominant antigen has lipid modifications similar to those of host cells and other apicomplexan parasites. Mass spectrometric demonstration here of glycopeptides with O-glycans complements previous identification *C. parvum* O-GalNAc transferases, lectin binding to vaccine candidates, and human and mouse antibodies binding to glycopeptides. The significance of these post-translational modifications is discussed with regards to the function of these proteins and the design of serological tests and vaccines.

Introduction

C. parvum infects humans and cows, while *C. hominis* only infects humans [1–3]. *C. parvum* was first identified as an opportunistic pathogen and cause of severe diarrhea in AIDS patients [4, 5]. In 1993, *C. parvum* contaminated the municipal water supply in Milwaukee, Wisconsin, U.S.A and caused a massive outbreak of diarrhea among immunocompetent persons [6, 7]. More recently *C. parvum* has been shown to be the second most important cause (after

RR025082, and S10 OD010724 (C.E.C.) and from NIH-NHLBI contract HHSN268201000031C (C.E. C.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests: The authors have declared that no competing interests exist.

rotavirus) of diarrhea and death in infants in low resource countries where the parasite is endemic [8–10]. Presently, there are no human vaccines for *C. parvum*, although numerous candidates have been identified [11, 12]. Treatment of *C. parvum* is difficult in populations with the most severe disease: infants and immunosuppressed persons [1, 13].

Oocysts of *C. parvum* have acid-fast (lipid-rich) walls, which are resistant to environmental insults and to gastrointestinal acids, proteases, and bile [5, 14]. Oocysts each contain four infectious sporozoites, which have on their surface Ser- and Thr-rich glycoproteins (e.g. Gp900 and Gp40) [15–20]. The precursor protein (Gp40/Gp15), which is specific for *C. parvum* and *C. hominis*, is cleaved by a furin-like protease into an N-segment (Gp40) and a C-segment (Gp15) [21]. Subsequently, the N-terminal signal peptide of Gp40 is removed, and a glycosyl-phosphatidylinositol (GPI) anchor is added to the C-terminus of Gp15 [22]. Gp40 contains a domain of 17 consecutive Ser residues followed by Thr-Ser-Thr, while the Ser and Thr residues of Gp15 are dispersed. Gp900, although much larger than Gp40/Gp15, is a secreted protein present in *C. parvum*, *C. hominis*, and *C. muris*. Gp900 has four sets of consecutive Thr residues, ranging in length from 33 to 155 residues, as well as dispersed Ser and Thr [16, 17]. Gp900, which is shed from the surface of sporozoites during gliding motility, tethers sporozoites to the interior surface of the oocyst wall [15, 16, 23]. In contrast, Gp15 is present at the apical end of sporozoites and on the outer surface of the oocyst wall [15]. Sporozoites also make a galactose/GalNAc-specific lectin, which interacts with Gp900 and Gp40 [24].

Polymorphisms in Gp15 and Gp40 have been used to distinguish among isolates of *C. parvum*, while recombinant Gp15 has been used to measure the serological response in epidemiological studies [25–30]. Vaccination studies have been performed using recombinant *C. parvum* proteins, bacterial vectors (e.g. *Salmonella*), or DNA encoding *C. parvum* proteins [11, 12, 31]. These vaccines either contain no O-glycans (bacterially expressed proteins) or may display host O-glycans (DNA vaccines). The presence of O-glycans (most likely O-GalNAc) on *C. parvum* glycoproteins has not previously been detected by mass spectrometry, but it has been suggested on the basis of the following five observations: (1) The *C. parvum* genome predicts four O-GalNAc transferases (O-GalNAcTs), and parasite lysates add O-GalNAc to synthetic peptides [32]. (2) A lectin that recognizes O-GalNAc (*Helix pomatia* agglutinin) (HPA) binds to the surface of sporozoites, while binding of a monoclonal antibody (4E9) to Western blots of *C. parvum* proteins is competed by HPA and reduced by treating proteins with an O-GalNAcase [25]. (3) The *Maclura pomiphora* agglutinin, which binds O-GalNAc, dramatically enriches Gp40, Gp900, and other mucin-like glycoproteins of *C. parvum* [15]. (4) Sera from patients infected with *C. parvum* bind to synthetic peptides containing O-linked GalNAc [33]. (5) O-GalNAc is added to *C. parvum* Gp40 exogenously expressed in *Toxoplasma gondii* [34].

Cp23, which is also known as the immunodominant antigen, is a small, low complexity, acidic protein of sporozoites [35]. Monoclonal antibodies to Cp23 partially protect neonatal mice against oral infection with *C. parvum*, while antibodies to Cp23 have more frequently been found in HIV/AIDS patients infected with *C. parvum* but without diarrhea [36, 37]. Recombinant Cp23 has been used to demonstrate humoral and cellular immune responses to *C. parvum* in human, cattle, and mouse infections, whereas recombinant Cp23 and DNA-based vaccines have been used to immunize mice and to elicit an innate immune response from mouse and human dendritic cells *in vitro* [29, 38–45]. Mass spectrometric studies of *C. parvum* sporozoites and oocysts have identified numerous peptides from Gp40, Gp15, Gp900, and Cp23, but none of these studies localized post-translational modifications (PTMs), which may include O-linked glycans, Asn-linked glycans (N-glycans), and fatty acyl chains [15, 22, 32, 33, 46–51]. Recently we used mass spectrometry to determine that *C. parvum* N-glycans, which are built on a predicted precursor with a single long mannose arm, appear to be processed by glucosidase-2 but not by ER mannosidases, and are not modified by Golgi

glycosyltransferases [52–54]. The resulting *N*-glycans, which are likely GlcMan₅GlcNAc₂ and Man₅GlcNAc₂, are remarkable for their simplicity, as compared to the complicated *N*-glycans identified in other protists [55]. In this report, we used mass spectrometry to characterize tryptic glycopeptides of lysates of *C. parvum* oocysts and thereby directly determine the number and some of the positions of *O*-GalNAc residues on Gp40, Gp15, Gp900, and a previously uncharacterized glycoprotein with a predicted weight of 20-kDa (named here Gp20). Mass spectrometry of hydrophobic peptides also detected the addition of myristoyl and palmitoyl groups to the first and second residues, respectively, at the N-terminus of Cp23 [56, 57].

Materials and methods

Reagents and parasites

Freshly passaged *C. parvum* oocysts were purchased from Bunch Grass Farm (Deary, ID) and handled under BSL-2 protocols, with the approval of the Boston University Institutional Biosafety Committee. All reagents and chemicals were purchased from Sigma-Aldrich (St. Louis, MO), unless noted otherwise. Solvents used for LC-MS were Optima™ grade, procured from Fisher Scientific (Thermo-Fisher Scientific, Waltham, MA).

Protein extraction and trypsin digestion

Procedures for extracting proteins from *C. parvum* oocysts and digesting them with trypsin have recently been described in detail [54], and so a summary of the methods is presented here. Briefly, 10⁹ *C. parvum* oocysts were concentrated by centrifugation, washed 3X with PBS, and resuspended with PBS containing EDTA-free cOmplete™ protease inhibitor (Roche, Basel, Switzerland). Oocyst walls were disrupted in a bead beater with 0.5-mm glass beads and centrifuged. The PBS supernatant was removed and saved, while the remaining insoluble materials and beads were extracted with a solution composed of 10 mM HEPES, 25 mM KCl, 1 mM CaCl₂, 10 mM MgCl₂, 2% CHAPS, 6 M guanidine HCl, 50 mM dithiothreitol, 1X protease inhibitor, pH 7.4. The resulting guanidine-DTT supernatant was combined with the PBS supernatant, and the insoluble material was discarded. The proteins were then precipitated, and the pellet was washed with methanol and vacuum dried. Alternatively, oocyst proteins were extracted with hot phenol, and phenol and interphase layers were kept, while the aqueous layer was discarded. Proteins were precipitated with methanol containing 100 mM NH₄OAc and dried. The pelleted proteins were resuspended in 50 mM NH₄HCO₃, pH 8.0, reduced with 50 mM DTT, alkylated with iodoacetamide, and then digested with proteomics grade trypsin (Sigma-Aldrich, St. Louis, MO). Tryptic peptides were dried and desalted using C18 ZipTip concentrators following the manufacturer's protocol (EMD Millipore, Danvers, MA).

Mass spectrometry

The LC-MS/MS methodologies and the manual interpretation of MS/MS spectra of *C. parvum* glycopeptides containing *O*-glycans were performed using the methods described for *C. parvum* *N*-glycosylated peptides [54]. Desalted and dried peptides from three biological replicates were dissolved in 2% ACN, 0.1% formic acid (FA) and separated using a NanoAcquity Ultra Performance Liquid Chromatography (UPLC) system (Waters, Milford, MA), fitted with a nanoAcquity Symmetry C18 trap column and a BEH130C18 analytical column. Solvent mixtures for the mobile phase gradient were 99:1:0.1 HPLC grade water/ACN/FA and 99:1:0.1 ACN/HPLC grade water/FA. The UPLC was coupled to a TriVersa NanoMate ion source (Advion, Ithaca, NY), operated at 1.5 kV to introduce ions into either an LTQ-Orbitrap-XL or a QE Plus mass spectrometer (Thermo-Fisher Scientific, San Jose, CA). Both mass spectrometers were operated

in the positive-ion mode. MS¹ spectra were recorded over the range m/z 350–2000. MS² HCD spectra were acquired by isolating the top 5 (LTQ-Orbitrap) or top 20 (QE+) precursor ions with a 2- m/z window and fragmenting the selected precursor ions with 15–45 V HCD energy. The lower energy MS² HCD spectra were scanned from m/z 100 to an upper m/z value, which was dependent upon the parent ion m/z . For the 45-V HCD spectra, ions below m/z 210 were excluded to avoid trapping the very abundant HexNAc oxonium ion.

Manual interpretation of mass spectra. Data obtained from LC-MS/MS experiments were first examined using Qual Browser in the Xcalibur 2.2 software suite (Thermo-Fisher Scientific). Extracted ion chromatograms were generated from MS/MS spectra for oxonium ions of interest (HexNAc, m/z 204.0866; Hex-HexNAc m/z 366.1395; HexNAc₂, m/z 407.1670). Spectra containing one or more of these ion(s) were then manually interpreted [54]. Once a sequence was obtained, it was searched against the 3,803 entries within the *C. parvum* Iowa-II predicted proteome and cross-searched within the entire NCBI nr database, using the online NCBI BLASTP algorithm (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) [58–60]. The software, Glycoworkbench v2.1, release 146, was used to help calculate glycan compositions [61]. O-glycosylated peptides utilized HexNAc almost exclusively. Due to the labile nature of O-linked glycans, b and y ions containing one or more HexNAc residues typically had very low abundances. The charge-reduced molecular ion that had undergone the loss of one or more HexNAc residues was often observed. The information obtained from manual interpretations was then used for database searches, allowing for deeper sequencing of the data and higher throughput processing of samples. The peak list, the assigned ions, and their mass errors for manually annotated spectra shown in the figures are listed in (S2 Excel File).

Database searches for glycopeptides. Automated database searches were performed using the PEAKS software suite version 8.0 (Bioinformatics Solutions Inc., Waterloo, ON, Canada), using recently described methods for N-glycans with modifications [54]. The search criteria were set as follows: trypsin as the enzyme with \leq two missed cleavages and \leq one non-specific cleavage, the error tolerances for the precursor of 6 ppm and 0.02 Da for fragment ions, carbamidomethyl cysteine as a fixed modification, and the dynamic modifications on Ser/Thr with (HexNAc to HexNAc₄) with \leq six/peptide. The peptide match threshold ($-10 \log P$) was set to 15, with estimation of the false discovery rate (FDR), a 5.6 FDR was calculated. A multi-round search was performed using the *de novo* only results from the first PEAKSDB search to find peptides with attached lipids. The second search parameters were identical to the prior PEAKSDB search, with the exception that myristate (N-term, Ser, Thr) and palmitate (Cys, Ser, Thr, Lys) were specified as dynamic modifications, and HexNAc modifications were removed. The results from the searches were exported into Excel and collated.

Re-annotation of glycopeptides from automated database searches. For multiple reasons, the PEAKS DB search algorithm failed to annotate the product ions appropriately. Therefore, the glycopeptide results from the PEAKSDB search were exported in mzIdentML 1.1 format [62], manually verified, then provided to GlycReSoft (a software package developed in-house for glycopeptide discovery and annotation). The code for GlycReSoft, which is currently in active development with periodic updates and improvements, is open source and freely available from the online repository: <https://github.com/BostonUniversityCBMS/glycresoft>. All peptides listed in the mzIdentML document and all non-redundant theoretical tryptic digest peptides for each included protein were used as templates, upon which a database of theoretical glycopeptides was constructed. Glycosylation was permitted at up to 20 putative sites. All distinct combinations-with-replacement with the putative glycan compositions were generated. For each template peptide, theoretical glycopeptides were produced by assigning glycosylation events for combinations of between 1 and k glycosylation sites, where k is the total number of potential glycosylation sites. The combinatorial complexity was reduced

by limiting the number of possibilities to the first 100 combinations, for glycopeptides having an excess of 100 possible placements.

Each dataset was deisotoped, charge state deconvolved, and searched independently against the database described above. Individual datasets of MS/MS scans in the range m/z 100–240 were filtered, and only tandem mass spectra for which the average ratio of oxonium ion signal to maximum signal exceeded 5% were considered. In addition to including normal peptide backbone fragments, the search considered spectra containing peaks that indicated either the presence of a HexNAc residue or its loss. The software also searched for the intact peptide backbone with zero or more partial losses of each potential glycan. Glycopeptide-spectrum matches were evaluated based upon joint binomial intensity-backbone coverage criteria, which included in a novel algorithm that is based in part on a binomial scoring function described previously [63]. The lists of ions assigned for each of these spectra are located in [S1 Excel File](#). [S4 Fig](#) shows a representative spectrum annotated by GlycReSoft (one of 345) submitted to the ProteomeXchange Consortium [64].

Other bioinformatic methods. The furin-like protease site that separates Gp40/Gp15 was predicted by the online tool “ProP 1.0 Server”, made available by the Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark (www.cbs.dtu.dk/services/ProP/) [65]. Signal peptides and transmembrane helices were predicted using the online tool Phobius (<http://phobius.sbc.su.se/index.html>) [66]. The GPI-anchor site of Gp15 was predicted using the BIG-PI prediction server (http://mendel.imp.ac.at/gpi/gpi_server.html) [67]. Cartoon representations of proteins and protein features, which were mapped with all the peptides across all MS/MS experiments, were generated using the online software tool Protter v.1.0 (<http://wlab.ethz.ch/protter/start/>) [68]. The assigned peptides from the PEAKSDB search results were used to map to the proteins of interest. The protein features were mapped using the results from the bioinformatics searches.

O-linked glycan release and characterization

Ser-linked or Thr-linked glycans were released from the proteins via reductive alkaline β -elimination. Briefly, purified proteins from a total oocyst lysate were first lyophilized in glass conical vials. To the dried protein extract, an aqueous solution of 0.1 M NaOH + 1 M NaBD₄ was added. The loosely capped vials were placed into an oven and kept at 45°C for 18 h. After the incubation period, the borate was removed by extensive washes with 10% acetic acid in methanol and then neat methanol. The released glycans were subsequently separated from the proteins by solid phase extraction columns. To the dried sample, LC-MS grade water containing 0.1% trifluoroacetic acid (TFA) was added; the tube was vigorously vortexed, and the contents were then passed through a C-18 Sep-Pak cartridge (Waters Corporation, Milford, MA). Three bed volumes of 0.1% TFA/water were subsequently passed through the column, and the eluent fractions were pooled and lyophilized. The released O-glycans were permethylated using previously described methods [69, 70]. A slurry of powdered NaOH in DMSO was added to the dried, released O-glycans. An equal volume of methyl iodide was added, and the reaction mixture was agitated gently while protected from light. The process was repeated three times to ensure complete permethylation. The product was extracted with chloroform/water, and the aqueous layer was removed and discarded. Washes with water were repeated until the pH of the solution was that of the LC-MS grade water being used for the washes. The chloroform layer was removed, placed into a new clean vial, dried in a speed vacuum and stored in a desiccator at -20°C until it was analyzed.

Monosaccharide composition determination by GC-MS. The permethylated sugars were identified using GC-MS with a Bruker Scion SQ interfaced to a 436-GC (Bruker

Daltonics, Billerica, MA). Separation was performed using a (30 m x 0.25 mm x 0.25 μ m) Restek™ Rxi™ - 5ms capillary column (Restek Corporation, Bellefonte, PA), using helium as the carrier gas. Samples were dissolved in hexane, and then 1 μ l of the solution was introduced via an auto-injector, using a split/split-less injection program, maintaining a constant column flow rate of 1 ml/min. The injector temperature was set to 220°C. The split-less injection sampling was set for 1 min before the split flow was started at 100 ml/min for 1 min. Then a split flow of 50 ml/min was used for the remainder of the program. The initial oven temperature of 60°C was maintained for 1 min, then ramped at 4°C/min to 250°C, with a final ramp to 300°C at 20°C/min, and held there for 10 min. Ions generated by an electron impact (EI) ionization source (70 eV) were introduced into the mass analyzer after a 5-min solvent delay. Centroid mass spectra were acquired in the positive mode, scanning the range m/z 50–500, taking 500 ms/scan. An internal standard of permethylated *myo*-inositol was added to all samples to verify retention time repeatability. Four spectra were averaged, and background was subtracted. The retention times and EI spectra of the released and permethylated glycans were compared to those of deuterio-reduced, permethylated monosaccharide standards. Data analysis was performed using the software MS Data Review 8.0 (Bruker). Retention times were compared using extracted ion chromatograms (XIC), for the ion signal at m/z 101, an ion common to GalNAc and GlcNAc. The EI spectra recorded for the standards and β -elimination products were compared at the same time points.

Results

The vast majority of peptides with *O*-HexNAc derive from Gp40, Gp15, and Gp900, which are vaccine candidates

Peptides obtained from trypsin digestion of total proteins of *C. parvum* oocysts were separated using a UPLC reversed phase C18 column that was online with a mass spectrometer. Peptides were subjected to Higher-energy C-trap Dissociation (HCD), and *O*-glycosylated peptides were recognized by the observation of a sugar-oxonium ion signal at m/z 204.0866 in the MS/MS spectra, corresponding to the fragmentation of a precursor containing a HexNAc residue. We reported previously the detection of larger sugar-oxonium ions (corresponding to Hex-HexNAc, HexNAc-HexNAc, and Hex-HexNAc-HexNAc) that all derive from *N*-glycans [54]. Since there was no enrichment for glycoproteins in the protein preparations (*e.g.* lectin chromatography), we identified the most abundant glycopeptides without selection bias. Glycopeptides with *O*-linked glycans originated from three *C. parvum* vaccine candidates (Gp15, Gp40, and Gp900), as well as Gp20, the immunogenicity of which is unknown (Fig 1, Table 1, and S1 Excel File). We also detected the presence of myristate and palmitate on an N-terminal peptide of the immunodominant antigen Cp23. In addition, we identified at least two peptides without *O*-HexNAc from each of 811 other *C. parvum* proteins. Information about these peptides and the glycopeptides described below has been deposited in the ProteomeXchange Consortium.

Dense arrays of *O*-GalNAc are present on the Ser-rich domain of Gp40

Gp40/Gp15 precursor (cgd6_1080) has an N-terminal signal peptide, a furin cleavage site that separates Gp40 (AA-22 to 220) from Gp15 (AA-221 to 324), and a C-terminal site for the addition of a GPI-anchor (Fig 1) [18–21]. A tryptic glycopeptide (AA-43 to 60) of Gp40, which contains 17 consecutive Ser residues followed by Thr-Ser-Thr, was found to be modified with 15 to 20 HexNAc residues (Table 1 and S1 Excel File). For example, the monoisotopic mass of the precursor ion m/z 1757.2272 $[M + 4H]^{4+}$ corresponds to the value calculated for the peptide ${}_{(43)}\text{DVPVEGSSSSSSSSSSSSSSSSSTSTVAPANK}_{(60)}$ with the addition of 20 HexNAc

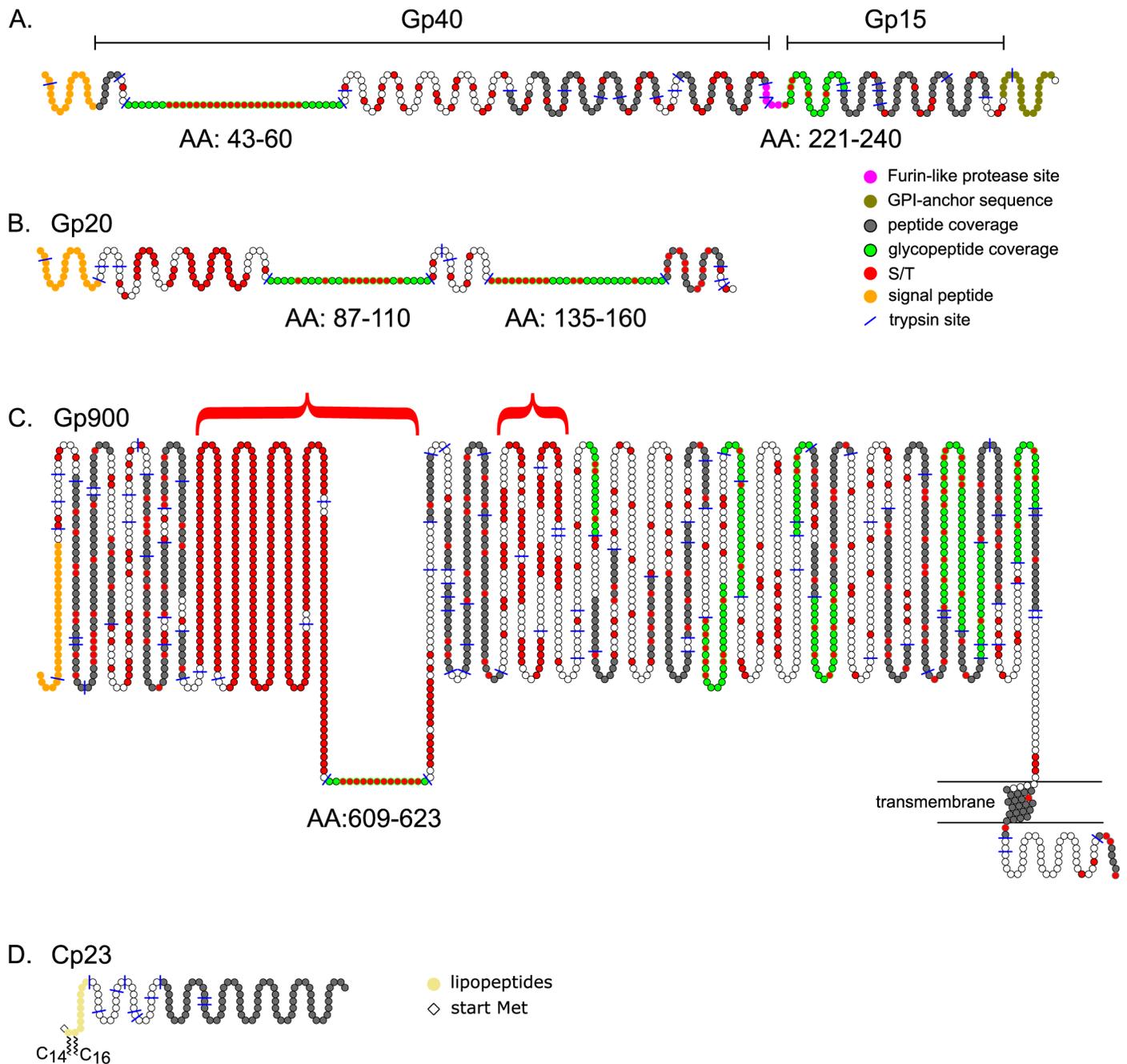


Fig 1. Schematics of *Cryptosporidium* glycoproteins characterized here by mass spectrometry. (A) Gp40/Gp15 precursor is cleaved at a furin-like protease site (pink) into Gp40 and Gp15. Mass spectrometry showed Gp40 has a Thr-rich domain (AA-43 to 60) with numerous *O*-linked HexNAc modifications (marked in green, with Ser and Thr residues marked in red). Gp15 contains a single domain (AA-221 to 240) that is glycosylated. Other peptides identified with mass spectrometry are marked in grey. Predicted N-terminal signal peptide is marked in orange, while GPI-anchor signal is marked in olive. (B) A 20-kDa glycoprotein (Gp20) contains two Thr-rich domains (AA-87 to 110 and AA-135 to 160), which contain numerous HexNAc modifications. (C) Gp900 contains two very large Thr-rich domains (red brackets), one of which contains a peptide with three HexNAc residues (AA-609 to 623). The transmembrane helix near the C-terminus is encompassed by two horizontal lines, representing a membrane. (D) The N-terminus of Cp23 is modified with *N*-myristate (C₁₄) and *S*-palmitate (C₁₆). The start Met is absent (diamond).

<https://doi.org/10.1371/journal.pone.0182395.g001>

residues (Fig 2 and S2 Excel File). The very abundant HexNAc oxonium ion (m/z 204.0866) and a very low abundance peak that fits the value for HexNAc₂ (m/z 407.1670) are present in

Table 1. Overview of tryptic glycopeptides identified by mass spectrometry.

Protein	Peptide [†]	#HexNAc
Gp40	(31)DVPVEGSSSSSSSSSSSSSSSTSTVAPANK ₍₆₂₎	15–20*
Gp15	(221)ETSEAAATVDLFAFTLDGGK ₍₂₄₀₎	1,2, 3*,4
Gp20	(87)EGEETDENTDETTTTTTTASPCK ₍₁₁₀₎	6,7*,8
	(135)SSTTTTTTAPVSSSEDNKPEDSEDEK ₍₁₆₀₎	8
Gp900	(609)KPTTTTTTTTTTTTK ₍₆₂₃₎	3
	(958)IADTSNLFVPVQTHK ₍₉₇₁₎	1
	(1197)TPTQTDSVTGKPIDPTTGLPFNPPTGH ₍₁₂₂₃₎	1
	(1197)TPTQTDSVTGK ₍₁₂₀₇₎	1
	(1243)YAVSNGIKTDNVYGLPVDEITGLPK ₍₁₂₆₇₎	1
	(1248)NGIKTDNVYGLPVDEITGLPK ₍₁₂₆₇₎	1
	(1373)GKDGLIVPPTNSINK ₍₁₃₈₇₎	1
	(1410)VIPGSLPGSLNYPSFNTPQQTDEITGK ₍₁₄₃₆₎	1
	(1646)TIPGSAASVIHTALGTPTQTDPTTGLPSPSTGLPFIIPGFNVLVDPQTGEQIK ₍₁₆₉₈₎	1
	(1658)ALGTPQTDPPTTGLPSPSTGLPFIIPGFNVLVDPQTGEQIK ₍₁₆₉₈₎	1
	(1710)EKNIIVTEAAYGLPVDPK ₍₁₇₂₆₎	1
	(1795)LIDPESGIAIDNSVSGVFATVPGTAAPK ₍₁₈₂₂₎	1
	(1813)ATVPGTAAPK ₍₁₈₂₂₎	1

The number of HexNAc residues is listed for each glycopeptide, and the most abundant glycopeptide in each set is marked with an asterisk. Missed and non-tryptic cleavages are omitted, unless these occurrences led to a change in the number of Ser or Thr residues.

[†]The complete list of glycopeptides used to generate this table can be viewed in [S1 Excel File](#).

<https://doi.org/10.1371/journal.pone.0182395.t001>

the 30-V HCD MS/MS spectrum. The observed dimer could be an artifact generated in the gas phase from the high population of HexNAc monomers. To a very large extent, glycan loss occurs prior to fragmentation of the peptide, with the result that the observed b and y ions contain zero to four HexNAc residues. The only product ion that can be used to assign the HexNAc modification to a specific amino acid is the γ_7^* ion, indicating the presence of HexNAc on the Thr closest to the C-terminus. In a second experiment, to avoid overpopulating the orbitrap analyzer with the less informative HexNAc oxonium ion, the start of the selection window was raised from m/z 100 to m/z 210, and, to ensure the generation of more peptide backbone fragments, the HCD energy was increased to 45 V ([S1 Fig](#) and [S2 Excel File](#)). The 45-V HCD MS/MS spectrum exhibited extensive fragmentation of the aglycon peptide, which resulted in product ions that composed a nearly complete b and y series (y_2 – y_{25} , y_{30}) and (b_2 – b_4 , b_6 – b_7) ([S1 Fig](#)).

Because we saw little evidence for the presence of HexNAc-HexNAc, we assume that each of the 20 potential O-glycan sites is occupied with a single HexNAc residue. We were unable to localize site occupancy in glycopeptides with 15–19 HexNAc residues, due to the labile nature of the O-glycans. Quite likely, the peptides modified with 15 to 19 HexNAc residues are a mixture of components having different occupancies.

Release of O-glycans from *C. parvum* sporulated oocyst proteins by reductive β -elimination, followed by GC/MS monosaccharide analysis versus sugar standards, showed that the HexNAc residues in the Gp40 glycopeptide and in glycopeptides of the other vaccine candidates are likely GalNAc ([S2 Fig](#)). However, we cannot rule out a small amount of GlcNAc, which was suggested by Western blots of Gp15 [[71](#)]. In support of this assignment are the previous reports that *C. parvum* has four O-GalNAcTs, and patient sera recognize synthetic glycopeptides derived from Gp40 and Gp15 with O-GalNAc [[32](#), [33](#)]. In summary, the Gp40 spectra

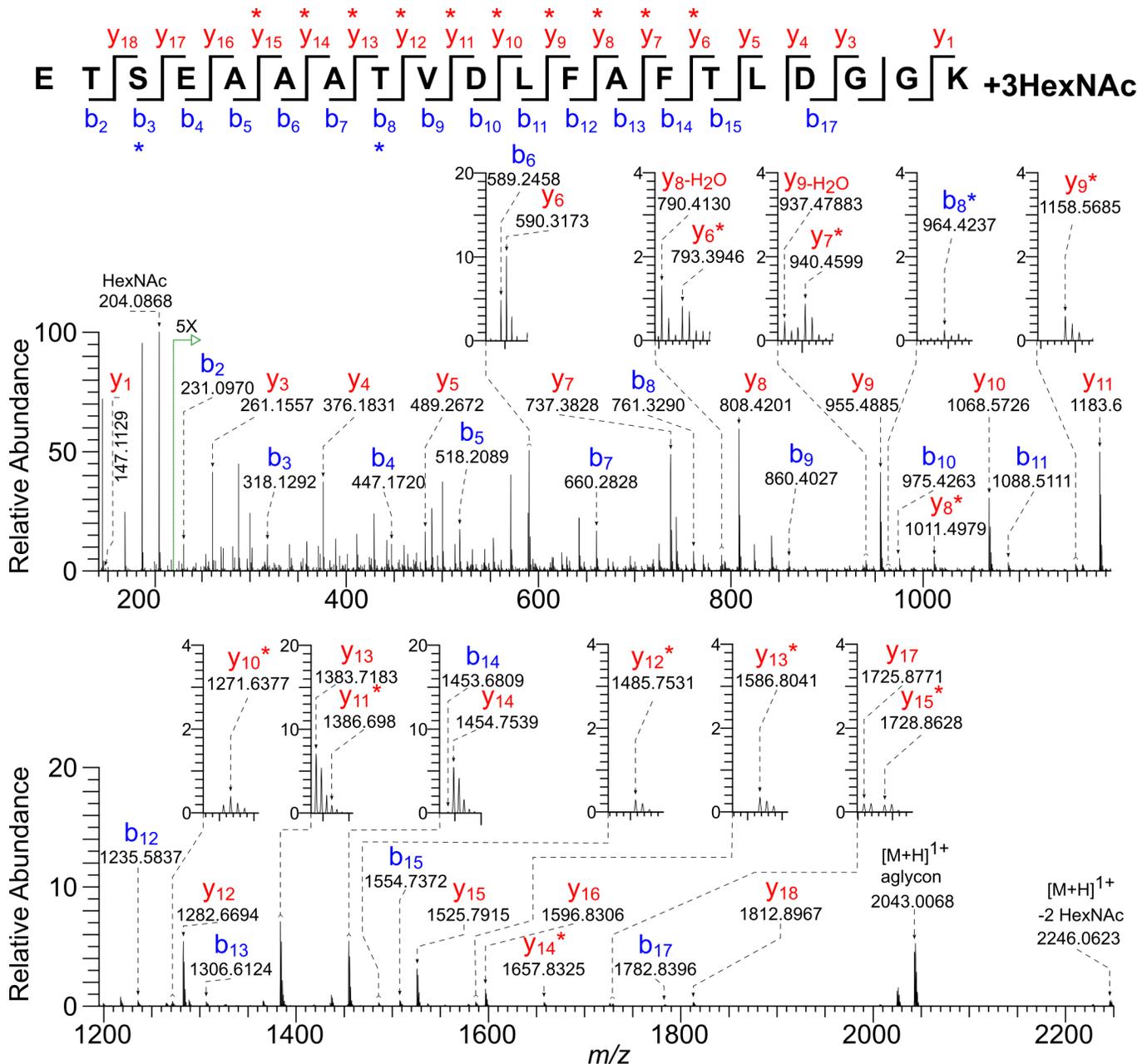


Fig 3. HCD MS/MS spectrum (@ 30 V) of a tryptic glycopeptide of Gp15 shows that three of four O-glycan sites contain the HexNAc modification. The precursor ion $[M + 2H]^{2+}$ m/z 1326.6164 of the most abundant Gp15 glycopeptide has a monoisotopic mass corresponding to that calculated for the peptide ${}_{(221)}\text{ETSEAAATVDLFAFTLDGGK}_{(240)}$ with the addition of three HexNAc residues (Δ 0.3 ppm). There is a prominent HexNAc oxonium ion (m/z 204.0868) and full series of b and y ions, some of which contain a single HexNAc residue (*).

<https://doi.org/10.1371/journal.pone.0182395.g003>

The b_3^* ion indicates that either Thr-222 or Ser-223 is modified. Thus there is evidence for distribution of the three HexNAc residues over the four available sites in this peptide. In the glycopeptide with four HexNAc modifications, all possible O-glycan sites are occupied. Analyses of the fragmentation patterns of numerous other peptides (S1 Excel File), both tryptic and non-tryptic, suggest that Thr-222 is preferentially modified over Ser-223, while Thr-228 and Thr-235 are nearly always modified.

Dense arrays of O-GalNAc are present on Thr-rich glycopeptides of Gp20

Gp20 (cgd7_1280), is a small, acidic, secreted protein with four domains with consecutive Thr residues, two of which are described here (Fig 1). The first Gp20 glycopeptide $_{(87)}\text{EGEETDEN TDETTTTTTTASPKPK}_{(110)}$ has 10 potential O-glycan sites and was found to be decorated with six to eight HexNAc residues (Table 1 and S1 Excel File). For example, the peak corresponding to the precursor ion of the most abundant Gp20 glycopeptide has a monoisotopic $[M + 4H]^{4+}$ m/z 1001.9305 equal to the value calculated for the peptide modified by seven HexNAc residues (Fig 4 and S2 Excel File). The 30-V HCD MS/MS spectrum includes a HexNAc oxonium ion (m/z 204.0868) and numerous b and y ions retaining zero to two HexNAc residues (marked with asterisks). Because the vast majority of HexNAc residues were lost prior to peptide fragmentation, it was not possible to define the seven occupied sites or to determine whether the occupancy was heterogeneous. A second Gp20 glycopeptide $_{(135)}\text{SSTTTTTTTT APVSSDKNKPEDSEDEK}_{(160)}$ with 12 potential O-glycan sites has a monoisotopic mass equal to that of the peptide with the addition of eight HexNAc residues (Table 1 and S1 Excel File). Again glycan loss prior to peptide backbone fragmentation made it impossible to localize the occupied O-glycans sites. Two other Thr-rich domains of Gp20 are present in a 55-amino acid tryptic peptide that was not identified. Regardless, the two Gp20 spectra show that the *C. parvum* O-GalNAcTs are capable of nearly saturating arrays of Thr residues.

A glycopeptide of Gp900 with consecutive Thr residues is lightly modified by O-GalNAc, while numerous Gp900 glycopeptides contain a single O-HexNAc residue

Gp900 (cgd7_4020), which has an N-terminal signal peptide and a transmembrane domain near its C-terminus, is by far the largest of the *C. parvum* vaccine candidates (1912 amino acids minus the signal peptide) (Fig 1) [16, 17]. One reason for the large size of Gp900 is the presence of a vast array of consecutive Thr residues, which extends from AA-304 to 640. A second Thr-rich region extends from AA-797 to 908. Because of the paucity of tryptic sites in the Thr-rich arrays of Gp900, and the likelihood that the Thr stretches are also heavily O-glycosylated, these regions were not observed by mass spectrometry, with one exception (Table 1, S1 Excel File, and S3 Fig). The precursor ion m/z 732.0284 $[M + 3H]^{3+}$ has a monoisotopic mass equal to that calculated for the peptide $_{(609)}\text{KPTTTTTTTTTTTTTTK}_{(623)}$ with the addition of only three HexNAc residues, despite the presence of 12 available sites (S3 Fig and S2 Excel File). The 30-V HCD MS/MS spectrum includes a HexNAc oxonium ion (m/z 204.0868) and numerous b and y ions containing zero to two HexNAc residues (marked with asterisks). Here again, because of the lability of the glycans, it was not possible to precisely define the occupied sites or to determine whether the occupancy was heterogeneous.

Many of the most abundant glycopeptides of Gp900 have a single HexNAc modification at an isolated Ser or Thr residue (Table 1 and S1 Excel File). For example, the precursor ion m/z 895.4646 $[M + 4H]^{4+}$ has a monoisotopic mass corresponding to the value calculated for the peptide $_{(1712)}\text{NIVTEAAYGLPVDPK}_{(1726)}$ plus a single HexNAc residue (Fig 5 and S2 Excel File). The b_4^* , b_6^* , and b_7^* ions show that Thr-1715 is modified. The mass spectra of 12 unique peptides from Gp900, each with a single HexNAc modification, together with the spectra from Gp15, suggest that the *C. parvum* O-GalNAcTs are capable of modifying isolated Ser and Thr residues, in addition to stretches of consecutive Ser residues in Gp40 and Thr in Gp20 and Gp900.

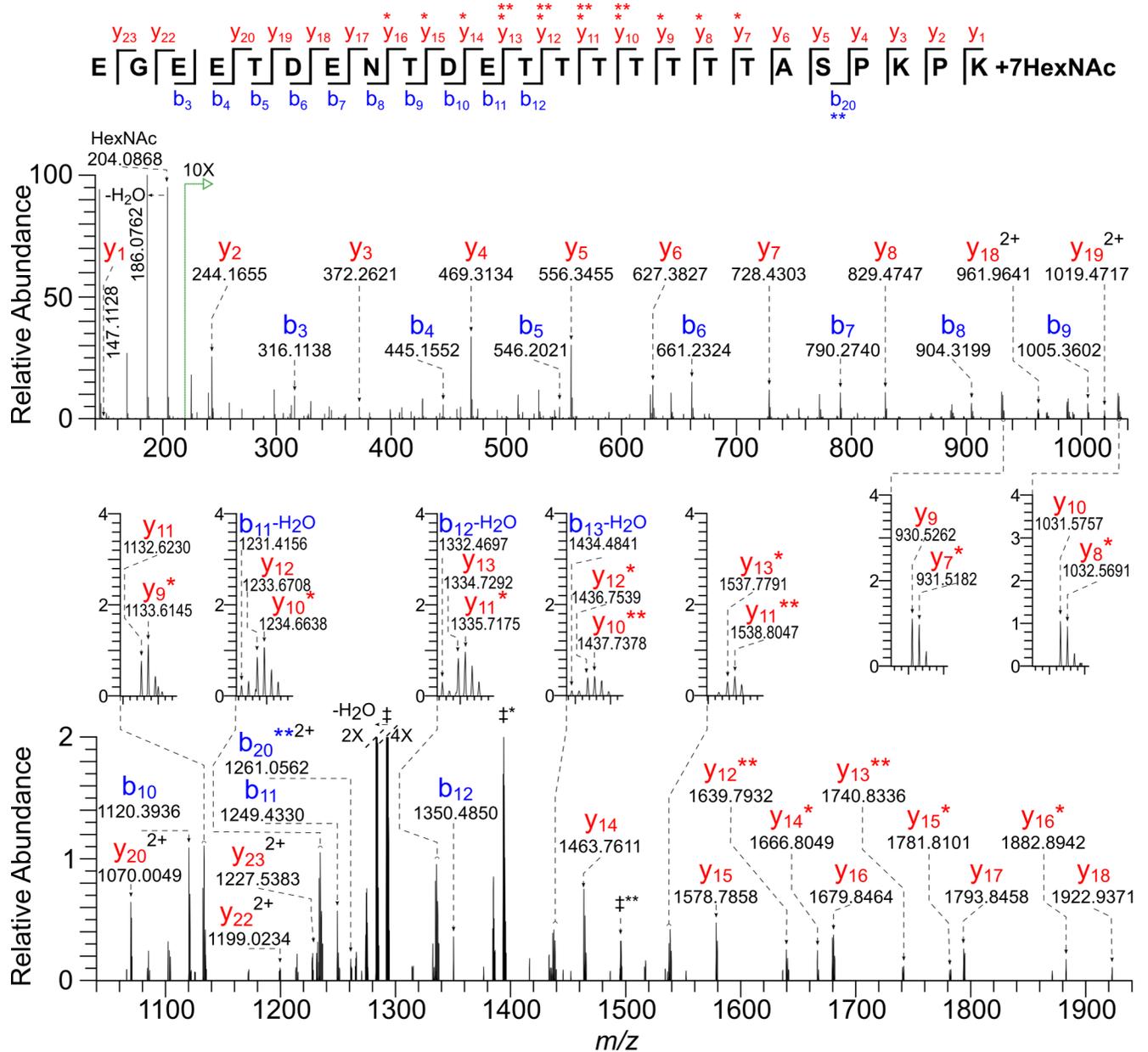


Fig 4. HCD MS/MS spectrum (@ 30V) of a tryptic glycopeptide of Gp20 shows near saturation of consecutive Thr residues with the O-HexNAc modification. The precursor ion $[M + 4H]^{4+}$ m/z 1001.9305 has a monoisotopic mass corresponding to that calculated for the peptide ${}_{(87)}\text{EGEETDENTDETTTITTTTASP}K_{(110)}$ plus seven HexNAc residues (Δ 0.5 ppm). There is a prominent HexNAc oxonium ion (m/z 204.0868) and full series of b and y ions, some of which contain one (*) or two (**) HexNAc residues. All ions are singly charged, except where indicated. In addition, charge-reduced ions, all 2+, which correspond to species that have undergone consecutive losses of HexNAc residues, are labeled as follows: ‡** = $[M + 2H]^{2+} - \text{HexNAc}_5$ (m/z 1495.1470), ‡* = $[M + 2H]^{2+} - \text{HexNAc}_6$ (m/z 1393.6224), ‡ = $[M + 2H]^{2+}$ aglycon peptide (m/z 1292.0851).

<https://doi.org/10.1371/journal.pone.0182395.g004>

At the N-terminus of Cp23 myristoyl modifies Gly1, while palmitoyl modifies Cys2

The immunodominant antigen Cp23 (cgd4_3620) contains no signal peptide but has an N-terminal sequence (2) GCSSSKPETK (11) similar to those modified by fatty acyl chains in the

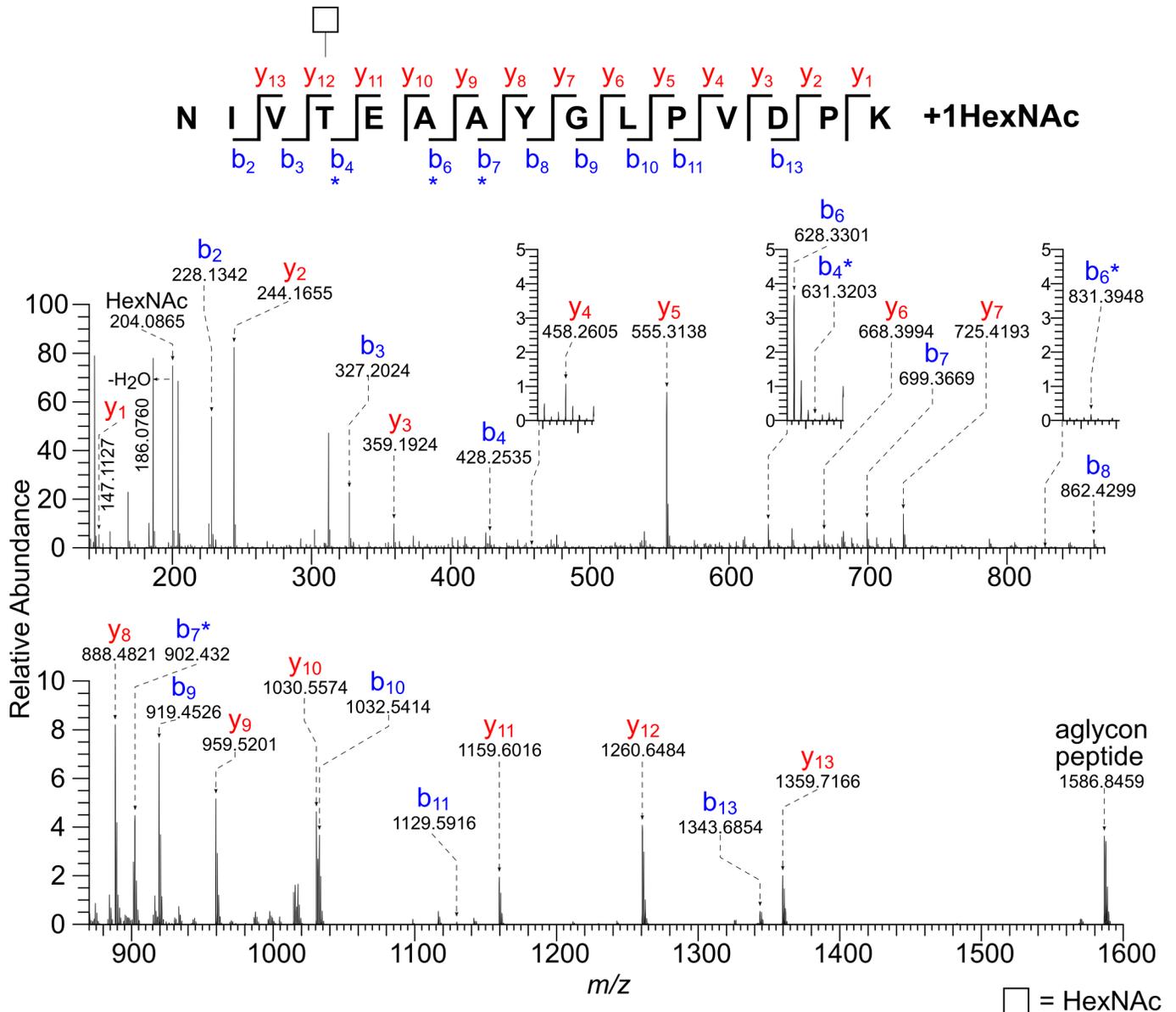


Fig 5. HCD MS/MS spectrum (@ 30V) of a tryptic glycopeptide of Gp900 shows a single O-HexNAc modification on a Thr residue. The precursor ion $[M + 4H]^{4+}$ m/z 895.4646 has a monoisotopic mass corresponding to that calculated for the peptide $(1712)_{NIVTEAAYGLPVDPK}$ (1726) plus a single HexNAc residue (Δ 0.1 ppm). There is a prominent HexNAc oxonium ion (m/z 204.0865) and full series of b and y ions, some of which contain a HexNAc residue (*). The b_4^* , b_6^* , and b_7^* ions show that Thr-1715 is modified.

<https://doi.org/10.1371/journal.pone.0182395.g005>

host and other apicomplexans (Fig 1) [52, 72–78]. Consistent with this resemblance, numerous hydrophobic peptides were identified by mass spectrometry containing the N-terminus of Cp23 minus Met-1, with no modification, substituted by either myristate or palmitate, or both (S1 Table.) [56]. For example, the precursor ion $[M + 2H]^{2+}$ m/z 736.4573 has a monoisotopic mass equal to that calculated for the peptide GCSSSKPETK with the addition myristate and palmitate (Fig 6 and S2 Excel File). Fragmentation using 30-V HCD showed the presence of myristate (m/z 211.2056) and palmitate (m/z 239.2370), as well as the charge reduced $[M + H]^{1+}$ molecular ion with loss of palmitate (m/z 1233.6782) or myristate (m/z 1261.7001). The presence of a complete y-ion series and a partial b-ion series allowed us to assign myristate to the N-terminal Gly

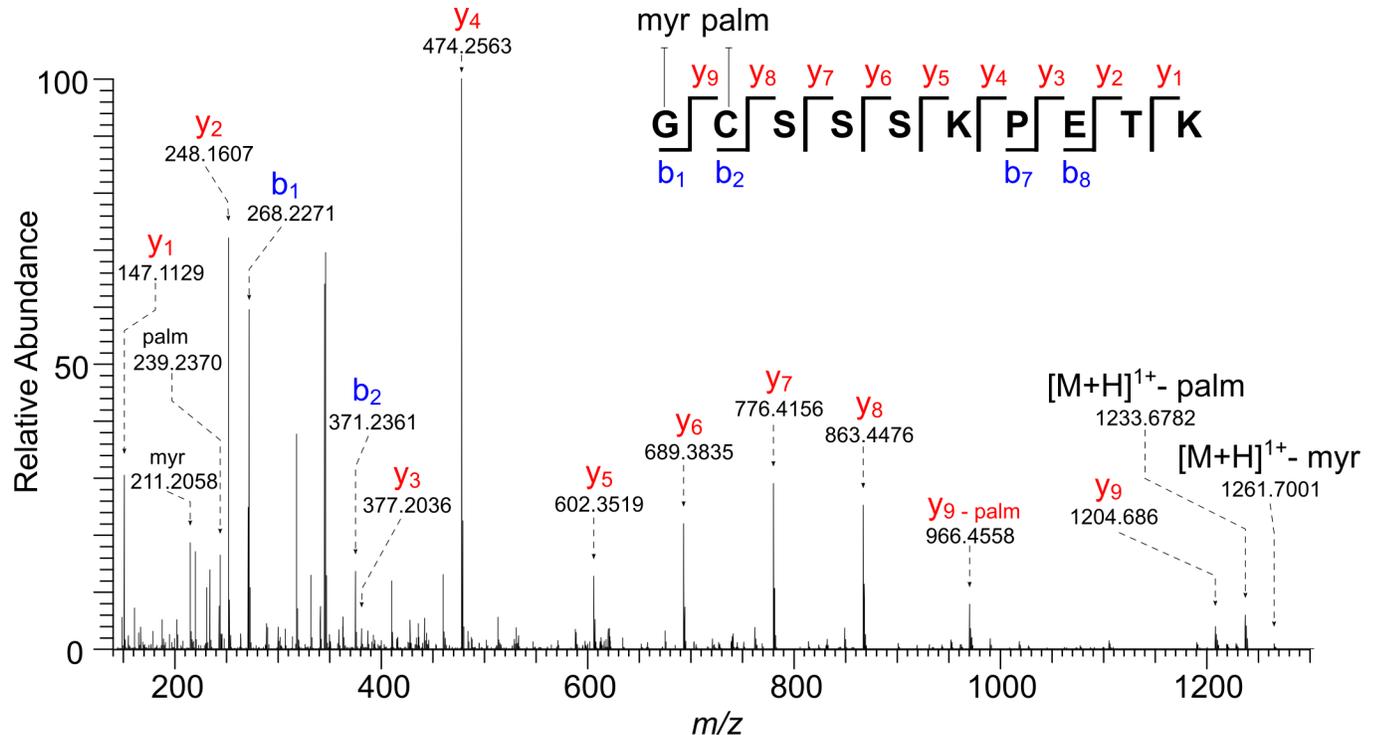


Fig 6. HCD MS/MS spectrum (@30 V) of an N-terminal peptide of Cp23 (minus Met-1) shows Gly-1 is modified with myristate and Cys-2 is modified with palmitate. The precursor ion $[M + 2H]^{2+}$ m/z 736.4573 has a monoisotopic mass corresponding to that calculated for the peptide $^{(2)}GCSSSKPETK_{(11)}$ plus myristate and palmitate (Δ 1.1 ppm). Fragment ions could be assigned to myristate (m/z 211.2056) and palmitate (m/z 239.2370), as well as the charge-reduced $[M + H]^{1+}$ molecular ions that have undergone loss of palmitate (m/z 1233.6782) or myristate (m/z 1261.7001). All the appropriate b/y ions contain the lipid modification, unless otherwise indicated.

<https://doi.org/10.1371/journal.pone.0182395.g006>

and palmitate to the Cys. We believe the example given is what is present on the native protein. The peptides where palmitate is absent and Cys is carbamidomethylated or palmitate modifies Ser residues arose during sample processing [56, 57].

Discussion

Mass spectrometry here directly demonstrated that addition of O-linked HexNAc (presumably O-GalNAc) is a widespread modification of *C. parvum* vaccine candidates (Gp15, Gp40, and Gp900) [11, 12]. Previous evidence for the addition of O-GalNAc to these proteins has been obtained through the use of synthetic glycopeptides, lectins, patient sera, or a monoclonal anti-carbohydrate antibody to *C. parvum* [25, 33]. O-GalNAc modifications saturate Ser-rich sequences of Gp40, and they nearly saturate Thr-rich sequences of Gp20, a previously uncharacterized protein. Nearly all of peptides with O-glycans derive from just four proteins (Gp40, Gp15, Gp900, and Gp20), even though >800 proteins were identified by mass spectrometry. Limitations of our observations include 1) failure to observe most of the Thr-rich domains of Gp900 and two of the Thr-rich domains of Gp20, 2) inability to assign O-glycans sites on many of the peptides due to loss of the O-linked glycan residues during HCD fragmentation, and 3) limited sampling of glycoproteins with O-GalNAc. In particular, the GalNAc-binding *Maclura pomifera* agglutinin previously enriched six mucin-like glycoproteins in addition to Gp40, Gp15, and Gp900 from lysed oocysts [15]. Other glycoproteins with O-GalNAc beyond those described here and the six mucin-like glycoproteins are certainly likely.

These results show that the four *O*-GalNAcTs of *C. parvum*, each of which has a lectin domain in addition to its glycosyltransferase domain, efficiently continue to glycosylate regions of glycoproteins that are already glycosylated [32, 79, 80]. Indeed the four *O*-GalNAcTs of *C. parvum* are able to make the same kind of modifications to arrays of Ser and Thr and to isolated Ser and Thr as the 20 *O*-GalNAcTs of the host. In studies beyond the scope of those performed here, the activity of each *O*-GalNAcT might be determined by 1) recombinant expression of enzymes with peptide substrates or 2) knockouts of the genes encoding these enzymes, a technology that is now available in *C. parvum* grown in mice [81]. *O*-glycans of *C. parvum* differ from those of the host in that *O*-GalNAc is not extended by other sugars [80]. *C. parvum* then is the equivalent of the “SimpleCell” lines engineered to express truncated *O*-GalNAc (knockout of *cosmc* gene), which have been used to map occupied *O*-glycan sites [82–84]. Similarly, the unextended *C. parvum* *O*-glycans are recognized by anti-Tn antibodies, which bind to unextended *O*-GalNAc [33, 85].

Properties that distinguish *C. parvum* Gp15 and Gp40 include glycosylation (discrete *O*-GalNAc residues versus densely clustered *O*-GalNAc residues), localization on sporozoites (apically associated versus diffusely covering surface), localization on oocyst walls (outer surface versus inner surface), and structure (GPI-anchored versus secreted) [15–23, 25]. We infer that the densely clustered *O*-GalNAc residues make the Ser-rich regions of Gp40 and Thr-rich regions of Gp20 and Gp900 rigid and extended rather than unstructured [86–88]. These extended regions of *O*-glycosylation may contribute to the tethering function of Gp40 and Gp900, which attach sporozoites to the inner layer of the oocyst wall [15]. *O*-glycosylation on these glycoproteins that coat the sporozoite surface may also affect host cell invasion and/or the innate and acquired immune responses to infecting parasites [25, 26, 28–30]. For example, the *C. parvum* galactose/GalNAc lectin binds to *O*-GalNAc on Gp40 and Gp900 [24]. In contrast, addition of myristate to N-terminal Gly and palmitate to Cys likely directs Cp23 from the cytosol to membranes of *C. parvum* and thus is likely important for its function, as has been extensively studied in *Toxoplasma*, *Plasmodium*, and the host [50, 72–78]. Chemical biology experiments or mutation of sites for addition of myristate and palmitate on Cp23 would be useful to test the roles of fatty acyl modifications in *C. parvum*. It is not clear how vaccination with recombinant Cp23, which is presumably a cytosolic protein, even if membrane bound, produces a protective immune response.

Serological screens for *C. parvum* use recombinant proteins made in bacterial systems that fail to add *O*-GalNAc (Gp40 and Gp15) or fatty acyl chains (Cp23) [26, 28–30, 36–38]. Because the host antibody response includes antibodies to glycopeptides with *O*-glycans including some with *O*-GalNAc on sites described here [33], these serological screens are likely lacking sensitivity and might be improved by expressing *C. parvum* proteins in SimpleCells that add unextended *O*-GalNAc to glycoproteins [82–84]. Similarly, vaccination with recombinant proteins produced in bacteria produces an immune response to the unmodified peptides, whereas acquired immunity to *C. parvum* infections includes responses to the *O*-glycans on Gp40 and Gp15 [33] and possibly lipid-modifications of Cp23. Again, the *cosmc* knockout might be used to produce recombinant Gp40 or Gp15 coated with unextended *O*-GalNAc for vaccination. Production of Cp23 in mammalian cells that add fatty acyl chains might increase the sensitivity of serological screens for this antigen and generate a better vaccine.

Supporting information

S1 Excel File. An Excel spreadsheet, summarizing the 358 spectra identified containing HexNAc(s). The list of unique peptides was used to make Table 1 in the main text, as shown as the representative peptides (sheet 1). The Gp40 peptides differ in the number of HexNAc

residues present and the number of spectra observed. There are 14 unique peptide sequences with varying number of HexNAc residues, and N-terminal degradation products represented by the Gp15 peptide in Table 1. There are two unique glycopeptide sequences represented by two of the Thr-rich peptides of Gp20. The first peptide has two unique sequences, resulting from an atypical trypsin cleavage between the terminal (107) PKPK(110), with cleavage after Lys-108. Sheet 2 contains detailed information about each manually verified PEAKS spectrum match. These spectra were re-annotated using the GlycoReSoft software developed in-house, as described in the methods section. These re-annotated spectra have been deposited in the PRIDE repository as.svg image files, Their concatenated name refers to the Source File-raw and the Scan# of the individual spectrum (columns O and P, respectively, in sheet 2 of S1 Excel File). An example re-annotated spectrum is provided in S4 Fig.

(XLSX)

S2 Excel File. An Excel spreadsheet of peak lists and ion assignments for the mass spectra shown in Figs 2 to 6, S1, and S3.

(XLSX)

S1 Fig. HCD MS/MS spectrum (@ 45V) of a tryptic glycopeptide of Gp40 gives complete sequence of the peptide. The precursor ion $[M + 4H]^{4+}$ m/z 1757.2272 corresponds to the monoisotopic mass equal to that of the peptide (43) DVPVEGSSSSSSSSSSSSSSSTSTVAPANK (60) with the addition of 20 HexNAc residues (Δ 0.6 ppm). The selection window was set to start at m/z 210 in order to exclude the very abundant oxonium ion (m/z 204.0866). There is extensive fragmentation of the aglycon peptide ((y₂—y₂₅, y₃₀) and (b₂—b₄, b₆—b₇) ions). Fig 2 shows the 30-V HCD MS/MS spectrum of the same peptide.

(TIFF)

S2 Fig. GC-MS analysis shows the O-glycan released by β -elimination from *C. parvum* glycoproteins is GalNAc. GC/MS data obtained for the deuteroreduced and permethylated glycan released by reductive β -elimination of *C. parvum* oocyst glycoproteins are compared with results observed for the GlcNAc and GalNAc standards, which were similarly treated. Extracted ion chromatograms of m/z 101 are shown on the left. Electron impact mass spectra of the standards and the product from reductive β -elimination of *C. parvum* oocyst glycoproteins are shown on the right. The sugar released from *C. parvum* is assigned as GalNAc, because the retention time (34 min) and EI mass spectrum both match those of the standard GalNAc. The EI mass spectra of the components eluting at positions marked with one asterisk (*) and two asterisks (**) do not correspond to sugar derivatives.

(TIFF)

S3 Fig. HCD MS/MS spectrum (@ 30V) of a tryptic glycopeptide of Gp900 shows partial glycosylation of a Thr-rich repeat. The precursor ion $[M + 3H]^{3+}$ m/z 732.0284 has a monoisotopic mass corresponding to the peptide (609) KPTTTTTTTTTTTTTK (623) with the addition of three HexNAc residues (Δ -0.1 ppm). There is a prominent HexNAc oxonium ion (m/z 204.0868) and a full series of b and y ions, some of which contain one (*) or two (**) HexNAc residues. The loss of HexNAc residues made it impossible to localize occupied sites. Charge-reduced aglycon peptide ions are observed in the spectrum: $= [M + 2H]^{2+}$ m/z 792.9193 and $\dagger = [M + H]^{1+}$ m/z 1584.8348.

(TIFF)

S4 Fig. A representative GlycReSoft re-annotated spectrum, one of 345 tandem mass spectra deposited into the PRIDE repository. This figure presents an example of an MS/MS spectrum initially assigned by PEAKS DB, then manually verified and re-annotated using the in-

house software GlycReSoft. GlycReSoft is capable of discovery and annotation, but only the annotation capabilities were utilized here.

(TIFF)

S1 Table. Lipid modifications of the N-terminus of Cp23.

(DOCX)

Acknowledgments

Thanks to Joseph Zaia for comments and supervision of J.A.K. We also thank Thermo Fisher Scientific for generously loaning the QE-Plus mass spectrometer.

Author Contributions

Conceptualization: John R. Haserick, Joshua A. Klein, Catherine E. Costello, John Samuelson.

Data curation: John R. Haserick, Joshua A. Klein.

Formal analysis: John R. Haserick, Joshua A. Klein, Catherine E. Costello, John Samuelson.

Funding acquisition: Catherine E. Costello, John Samuelson.

Investigation: John R. Haserick, Joshua A. Klein.

Methodology: John R. Haserick, Joshua A. Klein.

Project administration: Catherine E. Costello, John Samuelson.

Resources: Catherine E. Costello, John Samuelson.

Software: Joshua A. Klein.

Supervision: Catherine E. Costello.

Writing – original draft: John R. Haserick, Joshua A. Klein, Catherine E. Costello, John Samuelson.

Writing – review & editing: John R. Haserick, Catherine E. Costello, John Samuelson.

References

1. Checkley W, White AC Jr., Jaganath D, Arrowood MJ, Chalmers RM, Chen XM, et al. A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. *Lancet Infect Dis.* 2015; 15(1):85–94. [https://doi.org/10.1016/S1473-3099\(14\)70772-8](https://doi.org/10.1016/S1473-3099(14)70772-8) PMID: 25278220; PubMed Central PMCID: PMC4401121.
2. Esch KJ, Petersen CA. Transmission and epidemiology of zoonotic protozoal diseases of companion animals. *Clin Microbiol Rev.* 2013; 26(1):58–85. <https://doi.org/10.1128/CMR.00067-12> PMID: 23297259; PubMed Central PMCID: PMC4353666.
3. Fayer R, Xiao L. *Cryptosporidium* and cryptosporidiosis: CRC press; 2007.
4. Tzipori S, Widmer G. A hundred-year retrospective on cryptosporidiosis. *Trends Parasitol.* 2008; 24(4):184–9. <https://doi.org/10.1016/j.pt.2008.01.002> PMID: 18329342; PubMed Central PMCID: PMC2716703.
5. Garcia LS, Bruckner DA, Brewer TC, Shimizu RY. Techniques for the recovery and identification of *Cryptosporidium* oocysts from stool specimens. *J Clin Microbiol.* 1983; 18(1):185–90. PMID: 6193138; PubMed Central PMCID: PMC270765.
6. MacKenzie WR, Schell WL, Blair KA, Addiss DG, Peterson DE, Hoxie NJ, et al. Massive outbreak of waterborne cryptosporidium infection in Milwaukee, Wisconsin: recurrence of illness and risk of secondary transmission. *Clin Infect Dis.* 1995; 21(1):57–62. PMID: 7578760.
7. Baldursson S, Karanis P. Waterborne transmission of protozoan parasites: review of worldwide outbreaks—an update 2004–2010. *Water Res.* 2011; 45(20):6603–14. <https://doi.org/10.1016/j.watres.2011.10.013> PMID: 22048017.

8. Sarkar R, Kattula D, Francis MR, Ajjampur SS, Prabakaran AD, Jayavelu N, et al. Risk factors for cryptosporidiosis among children in a semi urban slum in southern India: a nested case-control study. *Am J Trop Med Hyg.* 2014; 91(6):1128–37. <https://doi.org/10.4269/ajtmh.14-0304> PMID: 25331810; PubMed Central PMCID: PMC4257634.
9. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multi-center Study, GEMS): a prospective, case-control study. *Lancet.* 2013; 382(9888):209–22. [https://doi.org/10.1016/S0140-6736\(13\)60844-2](https://doi.org/10.1016/S0140-6736(13)60844-2) PMID: 23680352.
10. Platts-Mills JA, Babji S, Bodhidatta L, Gratz J, Haque R, Havt A, et al. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health.* 2015; 3(9):e564–75. [https://doi.org/10.1016/S2214-109X\(15\)00151-5](https://doi.org/10.1016/S2214-109X(15)00151-5) PMID: 26202075.
11. Mead JR. Prospects for immunotherapy and vaccines against *Cryptosporidium*. *Hum Vaccin Immunother.* 2014; 10(6):1505–13. <https://doi.org/10.4161/hv.28485> PMID: 24638018; PubMed Central PMCID: PMC4185963.
12. Ludington JG, Ward HD. Systemic and Mucosal Immune Responses to *Cryptosporidium*-Vaccine Development. *Curr Trop Med Rep.* 2015; 2(3):171–80. PMID: 26279971; PubMed Central PMCID: PMC4535728.
13. Cabada MM, White AC Jr. Treatment of cryptosporidiosis: do we know what we think we know? *Curr Opin Infect Dis.* 2010; 23(5):494–9. <https://doi.org/10.1097/QCO.0b013e32833de052> PMID: 20689422.
14. Bushkin GG, Motari E, Carpentieri A, Dubey JP, Costello CE, Robbins PW, et al. Evidence for a structural role for acid-fast lipids in oocyst walls of *Cryptosporidium*, *Toxoplasma*, and *Eimeria*. *MBio.* 2013; 4(5):e00387–13. <https://doi.org/10.1128/mBio.00387-13> PMID: 24003177; PubMed Central PMCID: PMC3760245.
15. Chatterjee A, Banerjee S, Steffen M, O'Connor RM, Ward HD, Robbins PW, et al. Evidence for mucin-like glycoproteins that tether sporozoites of *Cryptosporidium parvum* to the inner surface of the oocyst wall. *Eukaryot Cell.* 2010; 9(1):84–96. <https://doi.org/10.1128/EC.00288-09> PMID: 19949049; PubMed Central PMCID: PMC2805294.
16. Barnes DA, Bonnin A, Huang JX, Gousset L, Wu J, Gut J, et al. A novel multi-domain mucin-like glycoprotein of *Cryptosporidium parvum* mediates invasion. *Mol Biochem Parasitol.* 1998; 96(1–2):93–110. PMID: 9851610.
17. Petersen C, Gut J, Doyle PS, Crabb JH, Nelson RG, Leech JH. Characterization of a > 900,000-M(r) *Cryptosporidium parvum* sporozoite glycoprotein recognized by protective hyperimmune bovine colostrum immunoglobulin. *Infect Immun.* 1992; 60(12):5132–8. PMID: 1452347; PubMed Central PMCID: PMC258288.
18. Priest JW, Kwon JP, Arrowood MJ, Lammie PJ. Cloning of the immunodominant 17-kDa antigen from *Cryptosporidium parvum*. *Mol Biochem Parasitol.* 2000; 106(2):261–71. PMID: 10699255.
19. Cevallos AM, Zhang X, Waldor MK, Jaison S, Zhou X, Tzipori S, et al. Molecular cloning and expression of a gene encoding *Cryptosporidium parvum* glycoproteins gp40 and gp15. *Infect Immun.* 2000; 68(7):4108–16. PMID: 10858228; PubMed Central PMCID: PMC101706.
20. Strong WB, Gut J, Nelson RG. Cloning and sequence analysis of a highly polymorphic *Cryptosporidium parvum* gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and 45-kilodalton zoite surface antigen products. *Infect Immun.* 2000; 68(7):4117–34. PMID: 10858229; PubMed Central PMCID: PMC101708.
21. Wanyiri JW, O'Connor R, Allison G, Kim K, Kane A, Qiu J, et al. Proteolytic processing of the *Cryptosporidium* glycoprotein gp40/15 by human furin and by a parasite-derived furin-like protease activity. *Infect Immun.* 2007; 75(1):184–92. <https://doi.org/10.1128/IAI.00944-06> PMID: 17043102; PubMed Central PMCID: PMC1828422.
22. Priest JW, Mehlert A, Moss DM, Arrowood MJ, Ferguson MA. Characterization of the glycosylphosphatidylinositol anchor of the immunodominant *Cryptosporidium parvum* 17-kDa antigen. *Mol Biochem Parasitol.* 2006; 149(1):108–12. <https://doi.org/10.1016/j.molbiopara.2006.04.006> PMID: 16759714.
23. O'Connor RM, Wanyiri JW, Cevallos AM, Priest JW, Ward HD. *Cryptosporidium parvum* glycoprotein gp40 localizes to the sporozoite surface by association with gp15. *Mol Biochem Parasitol.* 2007; 156(1):80–3. <https://doi.org/10.1016/j.molbiopara.2007.07.010> PMID: 17719100; PubMed Central PMCID: PMC2020432.
24. Bhat N, Joe A, PereiraPerrin M, Ward HD. *Cryptosporidium* p30, a galactose/N-acetylgalactosamine-specific lectin, mediates infection in vitro. *J Biol Chem.* 2007; 282(48):34877–87. <https://doi.org/10.1074/jbc.M706950200> PMID: 17905738.

25. Cevallos AM, Bhat N, Verdon R, Hamer DH, Stein B, Tzipori S, et al. Mediation of *Cryptosporidium parvum* infection in vitro by mucin-like glycoproteins defined by a neutralizing monoclonal antibody. *Infect Immun*. 2000; 68(9):5167–75. PMID: [10948140](#); PubMed Central PMCID: PMCPMC101770.
26. Ajjampur SS, Sarkar R, Allison G, Banda K, Kane A, Muliylil J, et al. Serum IgG response to *Cryptosporidium* immunodominant antigen gp15 and polymorphic antigen gp40 in children with cryptosporidiosis in South India. *Clin Vaccine Immunol*. 2011; 18(4):633–9. <https://doi.org/10.1128/CVI.00464-10> PMID: [21288997](#); PubMed Central PMCID: PMCPMC3122581.
27. Hira KG, Mackay MR, Hempstead AD, Ahmed S, Karim MM, O'Connor RM, et al. Genetic diversity of *Cryptosporidium* spp. from Bangladeshi children. *J Clin Microbiol*. 2011; 49(6):2307–10. <https://doi.org/10.1128/JCM.00164-11> PMID: [21471344](#); PubMed Central PMCID: PMCPMC3122776.
28. Lazarus RP, Ajjampur SS, Sarkar R, Geetha JC, Prabakaran AD, Velusamy V, et al. Serum Anti-Cryptosporidial gp15 Antibodies in Mothers and Children Less than 2 Years of Age in India. *Am J Trop Med Hyg*. 2015; 93(5):931–8. <https://doi.org/10.4269/ajtmh.15-0044> PMID: [26304924](#); PubMed Central PMCID: PMCPMC4703283.
29. McDonald AC, Mac Kenzie WR, Addiss DG, Gradus MS, Linke G, Zembrowski E, et al. *Cryptosporidium parvum*-specific antibody responses among children residing in Milwaukee during the 1993 waterborne outbreak. *J Infect Dis*. 2001; 183(9):1373–9. <https://doi.org/10.1086/319862> PMID: [11294669](#).
30. Sarkar R, Ajjampur SS, Muliylil J, Ward H, Naumova EN, Kang G. Serum IgG responses and seroconversion patterns to *Cryptosporidium* gp15 among children in a birth cohort in south India. *Clin Vaccine Immunol*. 2012; 19(6):849–54. <https://doi.org/10.1128/CVI.00051-12> PMID: [22518011](#); PubMed Central PMCID: PMCPMC3370436.
31. Roche JK, Rojo AL, Costa LB, Smeltz R, Manque P, Woehlbier U, et al. Intranasal vaccination in mice with an attenuated *Salmonella enterica* Serovar 908htr A expressing Cp15 of *Cryptosporidium*: impact of malnutrition with preservation of cytokine secretion. *Vaccine*. 2013; 31(6):912–8. <https://doi.org/10.1016/j.vaccine.2012.12.007> PMID: [23246541](#); PubMed Central PMCID: PMCPMC3563240.
32. Bhat N, Wojczyk BS, DeCicco M, Castrodad C, Spitalnik SL, Ward HD. Identification of a family of four UDP-polypeptide N-acetylgalactosaminyl transferases in *Cryptosporidium* species. *Mol Biochem Parasitol*. 2013; 191(1):24–7. <https://doi.org/10.1016/j.molbiopara.2013.08.002> PMID: [23954365](#); PubMed Central PMCID: PMCPMC3856541.
33. Heimbürg-Molinari J, Priest JW, Live D, Boons GJ, Song X, Cummings RD, et al. Microarray analysis of the human antibody response to synthetic *Cryptosporidium* glycopeptides. *Int J Parasitol*. 2013; 43(11):901–7. <https://doi.org/10.1016/j.ijpara.2013.05.012> PMID: [23856596](#); PubMed Central PMCID: PMCPMC3937990.
34. O'Connor RM, Kim K, Khan F, Ward HD. Expression of Cpgp40/15 in *Toxoplasma gondii*: a surrogate system for the study of *Cryptosporidium* glycoprotein antigens. *Infect Immun*. 2003; 71(10):6027–34. PubMed Central PMCID: PMCPMC201096. <https://doi.org/10.1128/IAI.71.10.6027-6034.2003> PMID: [14500524](#)
35. Perryman LE, Jasmer DP, Riggs MW, Bohnet SG, McGuire TC, Arrowood MJ. A cloned gene of *Cryptosporidium parvum* encodes neutralization-sensitive epitopes. *Mol Biochem Parasitol*. 1996; 80(2):137–47. PMID: [8892291](#).
36. Arrowood MJ, Mead JR, Mahrt JL, Sterling CR. Effects of immune colostrum and orally administered antisporezoite monoclonal antibodies on the outcome of *Cryptosporidium parvum* infections in neonatal mice. *Infect Immun*. 1989; 57(8):2283–8. PMID: [2744847](#); PubMed Central PMCID: PMCPMC313443.
37. Wanyiri JW, Kanyi H, Maina S, Wang DE, Steen A, Ngugi P, et al. Cryptosporidiosis in HIV/AIDS patients in Kenya: clinical features, epidemiology, molecular characterization and antibody responses. *Am J Trop Med Hyg*. 2014; 91(2):319–28. <https://doi.org/10.4269/ajtmh.13-0254> PMID: [24865675](#); PubMed Central PMCID: PMCPMC4125256.
38. Smith LM, Priest JW, Lammie PJ, Mead JR. Human T and B cell immunoreactivity to a recombinant 23-kDa *Cryptosporidium parvum* antigen. *J Parasitol*. 2001; 87(3):704–7. [https://doi.org/10.1645/0022-3395\(2001\)087\[0704:HTABCI\]2.0.CO;2](https://doi.org/10.1645/0022-3395(2001)087[0704:HTABCI]2.0.CO;2) PMID: [11426740](#).
39. Benitez AJ, McNair N, Mead JR. Oral immunization with attenuated *Salmonella enterica* serovar Typhimurium encoding *Cryptosporidium parvum* Cp23 and Cp40 antigens induces a specific immune response in mice. *Clin Vaccine Immunol*. 2009; 16(9):1272–8. <https://doi.org/10.1128/CVI.00089-09> PMID: [19605593](#); PubMed Central PMCID: PMCPMC2745010.
40. Ehigiator HN, Romagnoli P, Priest JW, Secor WE, Mead JR. Induction of murine immune responses by DNA encoding a 23-kDa antigen of *Cryptosporidium parvum*. *Parasitol Res*. 2007; 101(4):943–50. <https://doi.org/10.1007/s00436-007-0565-0> PMID: [17487508](#).
41. Bedi B, Mead JR. *Cryptosporidium parvum* antigens induce mouse and human dendritic cells to generate Th1-enhancing cytokines. *Parasite Immunol*. 2012; 34(10):473–85. <https://doi.org/10.1111/j.1365-3024.2012.01382.x> PMID: [22803713](#).

42. Borad AJ, Allison GM, Wang D, Ahmed S, Karim MM, Kane AV, et al. Systemic antibody responses to the immunodominant p23 antigen and p23 polymorphisms in children with cryptosporidiosis in Bangladesh. *Am J Trop Med Hyg.* 2012; 86(2):214–22. <https://doi.org/10.4269/ajtmh.2012.11-0273> PMID: 22302851; PubMed Central PMCID: PMC3269270.
43. Du XL, Xu JM, Hou M, Yu RB, Ge JJ, Zhu HS, et al. Simultaneous detection of serum immunoglobulin G antibodies to *Cryptosporidium parvum* by multiplex microbead immunoassay using 3 recognized specific recombinant *C. parvum* antigens. *Diagn Microbiol Infect Dis.* 2009; 65(3):271–8. <https://doi.org/10.1016/j.diagmicrobio.2009.07.020> PMID: 19733995.
44. Liu K, Zai D, Zhang D, Wei Q, Han G, Gao H, et al. Divalent Cp15–23 vaccine enhances immune responses and protection against *Cryptosporidium parvum* infection. *Parasite Immunol.* 2010; 32(5):335–44. <https://doi.org/10.1111/j.1365-3024.2009.01191.x> PMID: 20500662.
45. Shayan P, Ebrahimzadeh E, Mokhber-Dezfouli MR, Rahbari S. Recombinant *Cryptosporidium parvum* p23 as a target for the detection of *Cryptosporidium*-specific antibody in calf sera. *Parasitol Res.* 2008; 103(5):1207–11. <https://doi.org/10.1007/s00436-008-1117-y> PMID: 18677624.
46. Snelling WJ, Lin Q, Moore JE, Millar BC, Tosini F, Pozio E, et al. Proteomics analysis and protein expression during sporozoite excystation of *Cryptosporidium parvum* (Coccidia, Apicomplexa). *Mol Cell Proteomics.* 2007; 6(2):346–55. <https://doi.org/10.1074/mcp.M600372-MCP200> PMID: 17124246.
47. Sanderson SJ, Xia D, Prieto H, Yates J, Heiges M, Kissinger JC, et al. Determining the protein repertoire of *Cryptosporidium parvum* sporozoites. *Proteomics.* 2008; 8(7):1398–414. <https://doi.org/10.1002/pmic.200700804> PMID: 18306179; PubMed Central PMCID: PMC32770187.
48. Madrid-Aliste CJ, Dybas JM, Angeletti RH, Weiss LM, Kim K, Simon I, et al. EPIC-DB: a proteomics database for studying Apicomplexan organisms. *BMC Genomics.* 2009; 10:38. <https://doi.org/10.1186/1471-2164-10-38> PMID: 19159464; PubMed Central PMCID: PMC32652494.
49. Singh P, Mirdha BR, Srinivasan A, Rukmangadachar LA, Singh S, Sharma P, et al. Identification of invasion proteins of *Cryptosporidium parvum*. *World J Microbiol Biotechnol.* 2015; 31(12):1923–34. <https://doi.org/10.1007/s11274-015-1936-9> PMID: 26492887.
50. Resh MD. Fatty acylation of proteins: The long and the short of it. *Prog Lipid Res.* 2016; 63:120–31. <https://doi.org/10.1016/j.plipres.2016.05.002> PMID: 27233110; PubMed Central PMCID: PMC4975971.
51. Winter G, Gooley AA, Williams KL, Slade MB. Characterization of a major sporozoite surface glycoprotein of *Cryptosporidium parvum*. *Funct Integr Genomics.* 2000; 1(3):207–17. <https://doi.org/10.1007/s101420000028> PMID: 11793239.
52. Samuelson J, Banerjee S, Magnelli P, Cui J, Kelleher DJ, Gilmore R, et al. The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. *Proc Natl Acad Sci U S A.* 2005; 102(5):1548–53. <https://doi.org/10.1073/pnas.0409460102> PMID: 15665075; PubMed Central PMCID: PMC326545090.
53. Bushkin GG, Ratner DM, Cui J, Banerjee S, Duraisingh MT, Jennings CV, et al. Suggestive evidence for Darwinian Selection against asparagine-linked glycans of *Plasmodium falciparum* and *Toxoplasma gondii*. *Eukaryot Cell.* 2010; 9(2):228–41. <https://doi.org/10.1128/EC.00197-09> PMID: 19783771; PubMed Central PMCID: PMC32823003.
54. Haserick JR, Leon DR, Samuelson J, Costello CE. Asparagine-Linked Glycans of *Cryptosporidium parvum* Contain a Single Long Arm, Are Barely Processed in the Endoplasmic Reticulum (ER) or Golgi, and Show a Strong Bias for Sites with Threonine. *Mol Cell Proteomics.* 2017; 16(4 suppl 1):S42–S53. Epub February 8, 2017. <https://doi.org/10.1074/mcp.M116.066035> PMID: 28179475; PubMed Central PMCID: PMC326539390.
55. Schiller B, Hykollari A, Yan S, Paschinger K, Wilson IB. Complicated N-linked glycans in simple organisms. *Biol Chem.* 2012; 393(8):661–73. <https://doi.org/10.1515/hsz-2012-0150> PMID: 22944671; PubMed Central PMCID: PMC326589692.
56. Ji Y, Leymarie N, Haeussler DJ, Bachschmid MM, Costello CE, Lin C. Direct detection of S-palmitoylation by mass spectrometry. *Anal Chem.* 2013; 85(24):11952–9. <https://doi.org/10.1021/ac402850s> PMID: 24279456; PubMed Central PMCID: PMC326589692.
57. Ji Y, Bachschmid MM, Costello CE, Lin C. S- to N-Palmitoyl Transfer During Proteomic Sample Preparation. *J Am Soc Mass Spectrom.* 2016; 27(4):677–85. <https://doi.org/10.1007/s13361-015-1319-3> PMID: 26729453; PubMed Central PMCID: PMC326589692.
58. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25(17):3389–402. PMID: 9254694; PubMed Central PMCID: PMC326589692.
59. Aurrecochea C, Barreto A, Basenko EY, Brestelli J, Brunk BP, Cade S, et al. EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic Acids Res.* 2017; 45(D1):D581–D91. <https://doi.org/10.1093/nar/gkw1105> PMID: 27903906; PubMed Central PMCID: PMC326589692.

60. Heiges M, Wang H, Robinson E, Aurrecochea C, Gao X, Kaluskar N, et al. CryptoDB: a *Cryptosporidium* bioinformatics resource update. *Nucleic Acids Res.* 2006; 34(Database issue):D419–22. <https://doi.org/10.1093/nar/gkj078> PMID: 16381902; PubMed Central PMCID: PMCPMC1347441.
61. Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res.* 2008; 7(4):1650–9. <https://doi.org/10.1021/pr7008252> PMID: 18311910.
62. Jones AR, Eisenacher M, Mayer G, Kohlbacher O, Siepen J, Hubbard SJ, et al. The mzIdentML data standard for mass spectrometry-based proteomics results. *Mol Cell Proteomics.* 2012; 11(7):M111014381. <https://doi.org/10.1074/mcp.M111.014381> PMID: 22375074; PubMed Central PMCID: PMCPMC3394945.
63. Risk BA, Edwards NJ, Giddings MC. A peptide-spectrum scoring system based on ion alignment, intensity, and pair probabilities. *J Proteome Res.* 2013; 12(9):4240–7. <https://doi.org/10.1021/pr400286p> PMID: 23875887; PubMed Central PMCID: PMCPMC4117239.
64. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol.* 2014; 32(3):223–6. <https://doi.org/10.1038/nbt.2839> PMID: 24727771; PubMed Central PMCID: PMCPMC3986813.
65. Duckert P, Brunak S, Blom N. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel.* 2004; 17(1):107–12. <https://doi.org/10.1093/protein/gzh013> PMID: 14985543.
66. Kall L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* 2007; 35(Web Server issue):W429–32. <https://doi.org/10.1093/nar/gkm256> PMID: 17483518; PubMed Central PMCID: PMCPMC1933244.
67. Eisenhaber B, Bork P, Eisenhaber F. Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. *Protein Eng.* 1998; 11(12):1155–61. PMID: 9930665.
68. Omasits U, Ahrens CH, Muller S, Wollscheid B, Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics.* 2014; 30(6):884–6. <https://doi.org/10.1093/bioinformatics/btt607> PMID: 24162465.
69. Ciucanu I, Kerek F. A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research.* 1984; 131(2):209–17. [https://doi.org/10.1016/0008-6215\(84\)85242-8](https://doi.org/10.1016/0008-6215(84)85242-8)
70. Ciucanu I, Costello CE. Elimination of oxidative degradation during the per-O-methylation of carbohydrates. *J Am Chem Soc.* 2003; 125(52):16213–9. <https://doi.org/10.1021/ja035660t> PMID: 14692762.
71. Tilley M, Upton SJ, Fayer R, Barta JR, Chrisp CE, Freed PS, et al. Identification of a 15-kilodalton surface glycoprotein on sporozoites of *Cryptosporidium parvum*. *Infect Immun.* 1991; 59(3):1002–7. PMID: 1705238; PubMed Central PMCID: PMCPMC258359.
72. Child MA, Hall CI, Beck JR, Ofori LO, Albrow VE, Garland M, et al. Small-molecule inhibition of a depalmitoylase enhances *Toxoplasma* host-cell invasion. *Nat Chem Biol.* 2013; 9(10):651–6. <https://doi.org/10.1038/nchembio.1315> PMID: 23934245; PubMed Central PMCID: PMCPMC3832678.
73. Beck JR, Fung C, Straub KW, Coppens I, Vashisht AA, Wohlschlegel JA, et al. A *Toxoplasma* palmitoyl acyl transferase and the palmitoylated armadillo repeat protein TgARO govern apical rhoptry tethering and reveal a critical role for the rhoptries in host cell invasion but not egress. *PLoS Pathog.* 2013; 9(2):e1003162. <https://doi.org/10.1371/journal.ppat.1003162> PMID: 23408890; PubMed Central PMCID: PMCPMC3567180.
74. Frenal K, Kemp LE, Soldati-Favre D. Emerging roles for protein S-palmitoylation in *Toxoplasma* biology. *Int J Parasitol.* 2014; 44(2):121–31. <https://doi.org/10.1016/j.ijpara.2013.09.004> PMID: 24184909.
75. Foe IT, Child MA, Majmudar JD, Krishnamurthy S, van der Linden WA, Ward GE, et al. Global Analysis of Palmitoylated Proteins in *Toxoplasma gondii*. *Cell Host Microbe.* 2015; 18(4):501–11. <https://doi.org/10.1016/j.chom.2015.09.006> PMID: 26468752; PubMed Central PMCID: PMCPMC4694575.
76. Paul P, Chowdhury A, Das Talukdar A, Choudhury MD. Homology modeling and molecular dynamics simulation of N-myristoyltransferase from *Plasmodium falciparum*: an insight into novel antimalarial drug design. *J Mol Model.* 2015; 21(3):37. <https://doi.org/10.1007/s00894-015-2586-4> PMID: 25663521.
77. Wetzel J, Herrmann S, Swapna LS, Prusty D, John Peter AT, Kono M, et al. The role of palmitoylation for protein recruitment to the inner membrane complex of the malaria parasite. *J Biol Chem.* 2015; 290(3):1712–28. <https://doi.org/10.1074/jbc.M114.598094> PMID: 25425642; PubMed Central PMCID: PMCPMC4340414.
78. Wright MH, Clough B, Rackham MD, Rangachari K, Brannigan JA, Grainger M, et al. Validation of N-myristoyltransferase as an antimalarial drug target using an integrated chemical biology approach. *Nat Chem.* 2014; 6(2):112–21. <https://doi.org/10.1038/nchem.1830> PMID: 24451586; PubMed Central PMCID: PMCPMC4739506.

79. Gill DJ, Clausen H, Bard F. Location, location, location: new insights into O-GalNAc protein glycosylation. *Trends in cell biology*. 2011; 21(3):149–58. <https://doi.org/10.1016/j.tcb.2010.11.004> PMID: [21145746](https://pubmed.ncbi.nlm.nih.gov/21145746/)
80. Hurtado-Guerrero R. Recent structural and mechanistic insights into protein O-GalNAc glycosylation. *Biochem Soc Trans*. 2016; 44(1):61–7. <https://doi.org/10.1042/BST20150178> PMID: [26862189](https://pubmed.ncbi.nlm.nih.gov/26862189/).
81. Vinayak S, Pawlowic MC, Sateriale A, Brooks CF, Studstill CJ, Bar-Peled Y, et al. Genetic modification of the diarrhoeal pathogen *Cryptosporidium parvum*. *Nature*. 2015; 523(7561):477–80. <https://doi.org/10.1038/nature14651> PMID: [26176919](https://pubmed.ncbi.nlm.nih.gov/26176919/); PubMed Central PMCID: [PMCPMC4640681](https://pubmed.ncbi.nlm.nih.gov/PMC4640681/).
82. Wang Y, Ju T, Ding X, Xia B, Wang W, Xia L, et al. Cosmc is an essential chaperone for correct protein O-glycosylation. *Proceedings of the National Academy of Sciences*. 2010; 107(20):9228–33.
83. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J*. 2013; 32(10):1478–88. <https://doi.org/10.1038/emboj.2013.79> PMID: [23584533](https://pubmed.ncbi.nlm.nih.gov/23584533/); PubMed Central PMCID: [PMCPMC3655468](https://pubmed.ncbi.nlm.nih.gov/PMC3655468/).
84. Yang Z, Halim A, Narimatsu Y, Jitendra Joshi H, Steentoft C, Schjoldager KT, et al. The GalNAc-type O-Glycoproteome of CHO cells characterized by the SimpleCell strategy. *Mol Cell Proteomics*. 2014; 13(12):3224–35. <https://doi.org/10.1074/mcp.M114.041541> PMID: [25092905](https://pubmed.ncbi.nlm.nih.gov/25092905/); PubMed Central PMCID: [PMCPMC4256479](https://pubmed.ncbi.nlm.nih.gov/PMC4256479/).
85. Ju T, Otto VI, Cummings RD. The Tn antigen-structural simplicity and biological complexity. *Angew Chem Int Ed Engl*. 2011; 50(8):1770–91. <https://doi.org/10.1002/anie.201002313> PMID: [21259410](https://pubmed.ncbi.nlm.nih.gov/21259410/).
86. Hanisch FG. O-glycosylation of the mucin type. *Biol Chem*. 2001; 382(2):143–9. <https://doi.org/10.1515/BC.2001.022> PMID: [11308013](https://pubmed.ncbi.nlm.nih.gov/11308013/).
87. Strous GJ, Dekker J. Mucin-type glycoproteins. *Crit Rev Biochem Mol Biol*. 1992; 27(1–2):57–92. <https://doi.org/10.3109/10409239209082559> PMID: [1727693](https://pubmed.ncbi.nlm.nih.gov/1727693/).
88. Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*. 2015; 31(6):857–63. <https://doi.org/10.1093/bioinformatics/btu744> PMID: [25391399](https://pubmed.ncbi.nlm.nih.gov/25391399/); PubMed Central PMCID: [PMCPMC4380029](https://pubmed.ncbi.nlm.nih.gov/PMC4380029/).