O-fucosylated glycoproteins form assemblies in close proximity to the nuclear pore complexes of *Toxoplasma gondii*

Giulia Bandini[a], John R. Haserick[a,b], Edwin Motari[a,b,1], Dinkorma T. Ouologuem[c], Sebastian Lourido[d], David S. Roos[c], Catherine E. Costello[b], Phillips W. Robbins[a,2], and John Samuelson[a,2]

[a]Department of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, Boston, MA 02118; [b]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118; [c]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104; and [d]Whitehead Institute for Biomedical Research, Cambridge, MA 02142

Contributed by Phillips W. Robbins, August 19, 2016 (sent for review June 8, 2016; reviewed by Richard Cummings and John A. Hanover)

Toxoplasma gondii **is an intracellular parasite that causes disseminated infections in fetuses and immunocompromised individuals. Although gene regulation is important for parasite differentiation and pathogenesis, little is known about protein organization in the nucleus. Here we show that the fucose-binding *Aleuria aurantia* lectin (AAL) binds to numerous punctate structures in the nuclei of tachyzoites, bradyzoites, and sporozoites but not oocysts. AAL also binds to *Hammondia* and *Neospora* nuclei but not to more distantly related apicomplexans. Analyses of the AAL-enriched fraction indicate that AAL binds *O*-linked fucose added to Ser/Thr residues present in or adjacent to Ser-rich domains (SRDs). Sixty-nine Ser-rich proteins were reproducibly enriched with AAL, including nucleoporins, mRNA-processing enzymes, and cell-signaling proteins. Two endogenous SRDs-containing proteins and an SRD-YFP fusion localize with AAL to the nuclear membrane. Superresolution microscopy showed that the majority of the AAL signal localizes in proximity to nuclear pore complexes. Host cells modify secreted proteins with *O*-fucose; here we describe the *O*-fucosylation pathway in the nucleocytosol of a eukaryote. Furthermore, these results suggest *O*-fucosylation is a mechanism by which proteins involved in gene expression accumulate near the NPC.**

toxoplasma | fucose | nuclear glycosylation | nuclear pore complex

The apicomplexan parasite *Toxoplasma gondii* causes disseminated infections in humans, and these infections can lead to severe damage in immunocompromised individuals and fetuses (1, 2). There is no human vaccine against *T. gondii*, and recently the price of pyrimethamine, the drug used to treat toxoplasmosis in the United States, has increased more than 50-fold (2).

T. gondii has a complex life cycle, and the parasite's ability to differentiate through its life stages in response to stresses and environmental conditions is fundamental for its pathogenicity and transmission (3). Transcriptome analyses have revealed that a large percentage of mRNAs show life stage-specific expression (4) and/or cell cycle regulation (5). Recent studies have increased our understanding of gene expression in *T. gondii* by identifying the AP2 family of transcription factors (6–8) and by describing posttranslational modifications (PTMs) of histones and some of the enzymes responsible for them (9–11). However, little is known about protein organization at the nuclear periphery, a subnuclear compartment that plays a critical role in transcriptional regulation in many eukaryotes. In particular, the gene-gating model (12) suggests that the nuclear pore complex (NPC) has a role in transcriptional regulation and chromatin organization as well as in protein and mRNA transport (13, 14).

In *T. gondii* chromodomain protein 1 localizes with heterochromatin at the nuclear periphery (15), and centromeres sequester to an apical nuclear region (16). Although the nuclear localization signal (NLS) and importin-α system are present, key nuclear import and export molecules are not easily identified (17–19). Furthermore, the NPC composition is divergent, so that only nucleoporins containing phenylalanine-glycine (FG) repeats (FG-Nups) and a putative Nup54 can be predicted by primary sequence homology (20).

Here we report the discovery of numerous assemblies of *O*-fucosylated proteins that associate with the nuclear membrane near the NPCs. These results improve our understanding of the architecture of the *T. gondii* nuclear periphery and highlight *O*-fucosylation as a PTM involved in assemblies associated with the NPC.

Results

The Fucose-Binding *Aleuria aurantia* Lectin Labels the Nuclei of *T. gondii* in a Stage-Specific, Species-Specific, but Cell Cycle-Independent Manner. Fucose is a common monosaccharide in many eukaryote and prokaryote glycoconjugates (21). Because only one fucose-containing glycoconjugate, a pentasaccharide on S-phase kinase-associated protein 1 (Skp1) (22), was known in *T. gondii*, we used *Aleuria aurantia* lectin (AAL), a lectin that binds to terminal fucose, to search for fucosylated glycans in its various life stages (Fig. 1*A*). Unexpectedly, AAL strongly labels the nuclei of tachyzoites, bradyzoites, and sporozoites but fails to bind to the nuclei of oocysts (Fig. 1*B*). Binding of anti-centrin antibodies to centrosomes shows that oocyst walls are permeabilized (23). Transient expression of GAP40-YFP, which highlights the inner membrane complex (24), shows that AAL binds to *T. gondii* nuclei throughout the tachyzoite cell cycle (Fig. S1).

The apicomplexan phylogenetic tree based on housekeeping genes is cartooned in Fig. 1*C*. AAL binds to nuclei of *T. gondii*, *Hammondia hammondi*, and *Neospora caninum*, which are

Significance

We describe here the discovery that assemblies of *O*-fucosylated proteins localize to the nuclear membrane of *Toxoplasma gondii*, particularly in proximity to the nuclear pore complexes (NPCs). *O*-fucose is added to Ser and Thr residues found in some of the Phe-Gly (FG) domain-containing proteins that characterize the NPC channel as well as in Ser-rich sequences in many proteins predicted to have roles in transcription, mRNA processing, and cell signaling. *O*-fucosylation of nucleocytosolic proteins has not been described previously in any eukaryote and appears to be unique to *T. gondii* and closely related apicomplexans.

Author contributions: G.B., P.W.R., and J.S. designed research; G.B., J.R.H., E.M., and D.T.O. performed research; G.B., J.R.H., S.L., D.S.R., C.E.C., P.W.R., and J.S. analyzed data; and G.B., J.R.H., P.W.R., and J.S. wrote the paper.

Reviewers: R.C., Harvard Medical School; and J.A.H., National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

The authors declare no conflict of interest.

Data deposition: The proteomic data reported in this paper have been deposited in the ProteomeXchange Consortium database (dataset identifier PXD004426).

[1]Present address: Brunswick Labs, Southborough, MA 01772.

[2]To whom correspondence may be addressed. Email: robbinspw2@gmail.com or jsamuels@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1613653113/-/DCSupplemental.

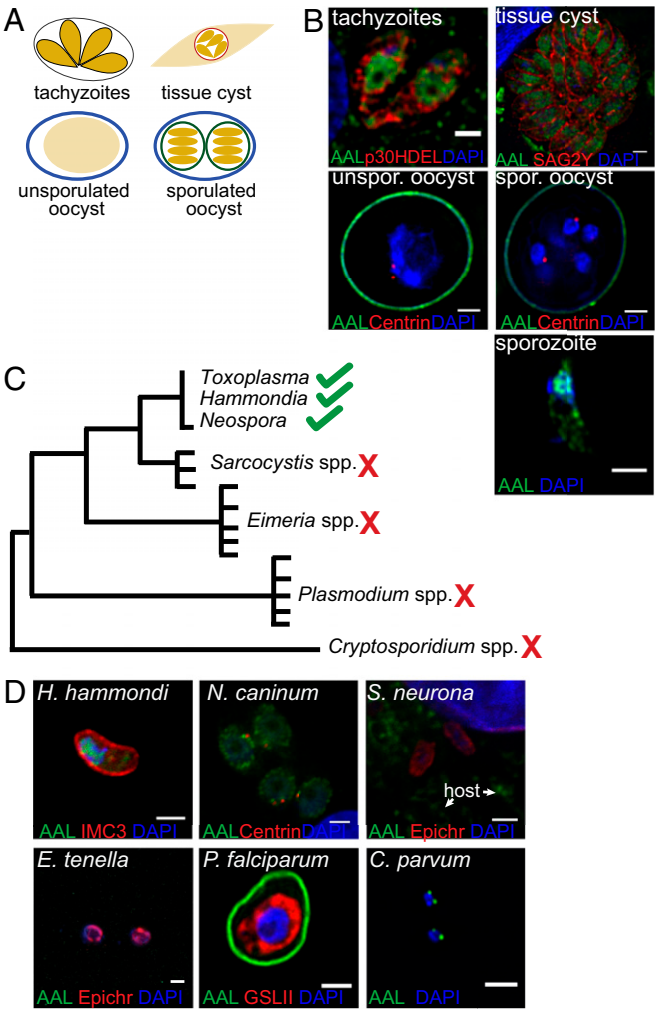

Fig. 1. AAL binding to nuclei of *T. gondii* is stage- and species-specific. (*A*) Schematic of the *T. gondii* life cycle. (*B*) AAL binds to nuclei of tachyzoites, bradyzoites, and sporozoites but not oocysts. (*C*) The Apicomplexa phylogenetic tree. (*D*) AAL binds to closely related organisms (*H. hammondi* and *N. caninum*) but not to more distantly related apicomplexans (*S. neurona*, *E. tenella*, and *P. falciparum*). Binding to sporozoites is shown for all species, with the exception of *P. falciparum* for which a trophozoite is shown. IMC3 marks the inner membrane complex, centrin marks the centrosome, and epichromatin marks the nuclear periphery. *Griffonia simplicifolia* lectin II (GSLII) labels the short *P. falciparum* N-glycans. All images were taken by deconvolution fluorescence microscopy. (Scale bars: 2 μm.)

closely related species, but not to the nuclei of the more distantly related *Sarcocystis neurona*, *Eimeria tenella*, or *Plasmodium falciparum* (Fig. 1*D*). Although AAL labels numerous loci within the nucleus of *T. gondii*, the lectin labels only one or two spots in the perinuclear region of *Cryptosporidium parvum* sporozoites. These results show that strong labeling of nuclei by AAL is not a common property of apicomplexans but instead appears to have developed after the divergence of *T. gondii* and *S. neurona*.

AAL binding to extracellular tachyzoites can be inhibited by preincubation with methyl-α-fucopyranoside (αMeFuc), as shown by flow cytometry and lectin blotting (Fig. 2 *A* and *B*), indicating the specificity of the carbohydrate–lectin interaction. Proteases and β-elimination treatments were also effective in blocking AAL binding to cell lysate, whereas peptide-N-glycosidase F (PNGase F) had no effect (Fig. 2 *B* and *C* and Fig. S2*C*), indicating that AAL binds to *O*-glycosylated proteins in *T. gondii*. In contrast, AAL binds to the N-glycans in the secretory pathway of human foreskin fibroblasts (HFFs), as shown by loss of binding on lectin blots only upon PNGase F treatment and by colocalization by immunofluorescence analysis with RCA120, a galactose-binding lectin (Fig. S2 *A* and *B*).

***T. gondii* Nuclear Fucosylation Is Dependent on the de Novo GDP–Fucose Biosynthetic Pathway.** Because *T. gondii* and other apicomplexans have no salvage pathway for fucose (25, 26), GDP-fucose is synthesized from GDP-mannose in a three-step process catalyzed by GDP-mannose dehydratase (GMD, TGGT1_238940) and GDP-fucose synthase (Fig. S2*D*) (21). Expression of Cas9 and a guide RNA (gRNA) targeting the first exon of the *T. gondii gmd* gene results in the loss of AAL binding to tachyzoite nuclei at 48 h post electroporation (Fig. 2*D*). Immunostaining against inner membrane complex protein 3 (IMC3) indicates that the subpopulation of AAL-negative cells still maintains the correct cell morphology (27). No loss of AAL binding was observed when cells were electroporated with Cas9 but without gRNA. So far we have been unable to isolate the *gmd*-depleted cells by cloning by limiting dilution with or without a selection agent (28), suggesting either that *gmd* is essential or that its deletion causes a strong growth phenotype in *T. gondii* tachyzoites.

The only characterized downstream user of GDP-fucose (GDP-Fuc) in *T. gondii* is the Skp1 glycosylation pathway. *T. gondii* is one of a few organisms that add a pentasaccharide containing GlcNAc, Gal, and fucose to a hydroxyl-proline on Skp1, a E3 ubiquitin ligase adaptor (22, 29). Knockout of Skp1 proline hydroxylase (*phyA*), GlcNAc transferase (*gnt1*), and bifunctional galactose and fucose transferase (*pgtA*) did not affect AAL binding (Fig. S3), indicating that biosynthesis of the AAL-positive glycan is independent of the Skp1 glycosylation pathway.

AAL Binds to *O*-Fucose in *T. gondii*. LC-MS/MS of tryptic peptides of AAL-enriched proteins from extracellular tachyzoites identified 69 unique glycopeptides containing one to six deoxyhexose(s) (dHexs), each linked to Ser or Thr (Dataset S1). These glycopeptides correspond to 50 different peptide sequences, and in many cases different numbers of dHexs were observed on the same peptide. All glycopeptides were manually reviewed and

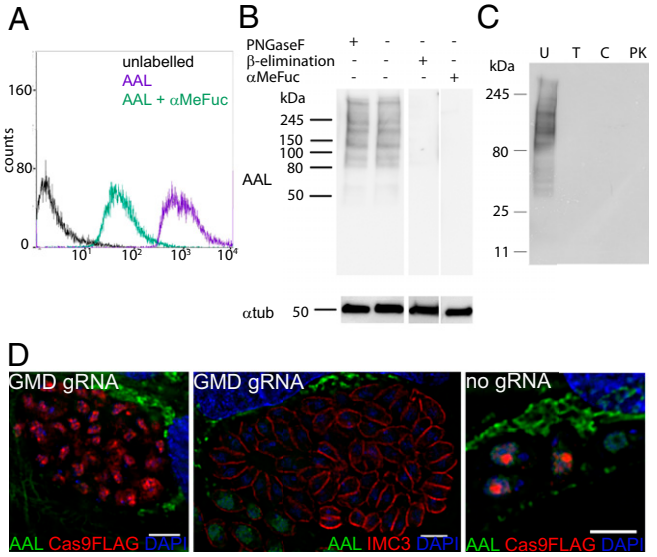

Fig. 2. AAL specifically binds to *O*-glycans on *T. gondii* glycoproteins. (*A*) AAL binding to extracellular tachyzoites by flow cytometry shows inhibition by preincubation with αMeFuc. (*B*) β-Elimination, but not PNGase F treatment, inhibits AAL binding to tachyzoite cell lysates. Tubulin is shown as a loading control. (*C*) Protease treatments show that AAL binds exclusively to glycoproteins. C, chymotrypsin; PK, proteinase K; T, trypsin; U, untreated. (*D*) Disruption of GDP-fucose synthesis using CRISPR/Cas9 eliminates AAL binding by deconvolution fluorescence microscopy. (Scale bars: 4 μm.)

confirmed by either prompt neutral loss of at least one dHex, as indicated by the presence of peaks in the MS that correspond to different glycoforms detected at the same retention time, and/or by peaks in the MS/MS spectra that can be confidently assigned to peptide fragments containing dHex (Dataset S1). No sugars other than dHex were observed, indicating that AAL binds to *O*-fucose (*O*-Fuc). In a few cases, the same peptide was observed modified with a dHex and, at a later retention time, in its unglycosylated form, suggesting that the addition of *O*-Fuc may be a probabilistic event, i.e., that distinct pools of modified and unmodified protein may be present at that site (Dataset S1). The amino acid modified with *O*-Fuc could not be determined for most glycopeptides because of the labile nature of the *O*-Fuc and the low complexity of the amino acid sequence of most of the modified peptides (Fig. S4*A*). Also, we cannot exclude the possibility of heterogeneity in the glycosylation sites. Combining high-energy collision dissociation (HCD) and electron-transfer dissociation (ETD) MS/MS data, we often were able to narrow the modification sites to two to six likely Ser/Thr residues (Fig. S4*B* and Dataset S1). Furthermore, we were able to specify the modification site for one glycopeptide, T610 on TGGT1_203780 (Fig. 3*B*), a putative FG-Nup and one of the most abundant proteins identified in the AAL-enriched fraction (Dataset S2). GC-MS monosaccharide composition analysis of the sugars released by reductive β-elimination identified fucose as the only dHex present in the AAL-enriched fraction (Fig. S4 *C* and *D*).

AAL Recognizes Proteins Involved in Gene Regulation. In most cases, *O*-Fuc is added to low-complexity Ser-rich domains (SRDs) or to sequences adjacent to these domains (Fig. 4 *A* and *F*), and peptides with long SRDs are likely to have more than one dHex (Fig. 3*A* and Dataset S1). We define SRDs as sequences with five or more Ser residues in tandem. More than 70% of the peptides identified in this study came from proteins that had one or more SRDs and/or contained more than 10% Ser (Fig. 3*A*, Fig. S5 *A* and *B*, and Dataset S2). Consistent with this observation, a higher number of proteins with Ser-15-mer are found in the apicomplexans that show nuclear AAL binding than in those that do not (Fig. 1*D* and Fig. S5*C*). Comparison of data from five biological repeats resulted in a set of 69 Ser-rich proteins reproducibly pulled down by AAL and/or for which we observed glycopeptides. We identified glycopeptides for 33 of the 69 AAL-enriched proteins (Table 1 and Dataset S2), and it is likely that numerous *O*-fucosylated peptides on long SRDs were not detected because of the paucity of flanking trypsin cleavage sites (Fig. 4*A*). In contrast, proteins that were present in AAL-enriched pull-downs but that are likely contaminants because of their high abundance (e.g., ribosomal proteins, cytoskeletal components, chaperones, and so forth) do not contain SRDs and have an average Ser content of 6%. A control set comprising *T. gondii* proteins associated with nuclear and cytoplasmic Gene Ontology (GO) terms was shown to have an average Ser content of 9%, compared with the 15% Ser present in the *O*-fucosylated proteins or the SRD set (Fig. S5*B*). About 40% of the 69 proteins with one or more SRDs contain a canonical NLS (cNLS), as identified by cNLS mapping (Table 1, Fig. S5*A*, and Dataset S2) (30). Analysis of PTMs showed that the majority of AAL-enriched proteins exhibit some degree of phosphorylation, but none is modified by ubiquitin. This result is in agreement with the datasets in refs. 31 and 32. AAL enriches five of seven predicted *T. gondii* nucleoporins (four FG-Nups and a Nup54 ortholog), each with one or more SRDs (Table 1 and Dataset S2). Also present in the AAL-enriched fraction are proteins predicted to be involved in mRNA processing, protein–protein interactions, ubiquitination, and enzymes that catalyze the addition/removal of phosphate groups from proteins and polyphosphate phosphatidylinositol (33, 34). Transcription regulators and proteins with nucleotide-binding and chromosome-binding domains are also present (Table 1 and Dataset S2). Numerous hypothetical proteins are found, and these often are conserved in *T. gondii*, *H. hammondi*, and *N. caninum* but are absent in the apicomplexans that do not bind AAL (Fig. 1).

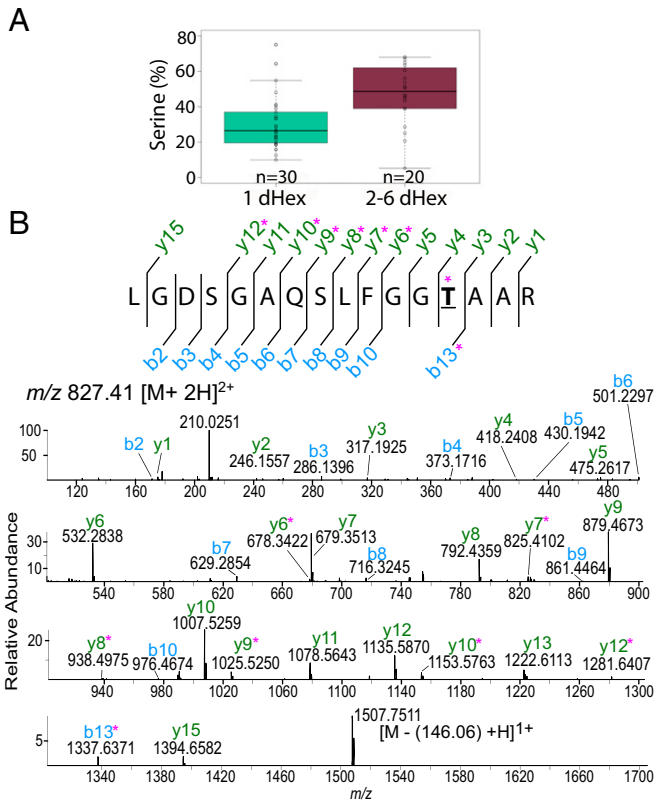

Fig. 3. AAL-bound proteins are modified by *O*-Fuc on Ser/Thr residues. (*A*) Peptides with a higher Ser percentage are more likely to be modified with more than one *O*-Fuc (dHex) residue. (*B*) HCD MS/MS spectrum of a peptide from a FG-Nup (TGGT1_203780), *m/z* 827.41 $[M+2H]^{2+}$. Ions b13+ dHex and y6+dHex were observed, indicating that T610 (bold underlined) is the modified residue. Ions plus dHex are marked by a pink asterisk.

***O*-Fucosylation Directs Glycoproteins to AAL-Labeled Assemblies Associated with the Nuclear Membrane in Close Proximity to the NPC.** Two AAL-enriched proteins were chosen for further studies: a putative Gly-Pro-Asn-loop GTPase (hereafter "GPN") and the transcriptional activator SWI/SNF2 (hereafter "SNF2"). GPN has a long N-terminal SRD but lacks an NLS (Fig. 4*A*). The protein was tagged with 3×MYC at its C terminus (Fig. S6*A*) and shown to be concentrated, like AAL, at the nuclear periphery but also was present in the cytosol, where no AAL labeling is observed (Fig. 4*B*). The tagged protein was also enriched by AAL (Fig. 4*C*). Addition of *O*-Fuc as a probabilistic event might explain why fucosylated GPN is present in the nucleus but unmodified GPN remains in the cytosol.

The SRD from GPN was fused to the N terminus of YFP (SRD-YFP) and expressed in tachyzoites (Fig. S6*B*). SRD-YFP localizes to the nucleus and to a lesser extent to the cytosol (Fig. 4*D*). Both SRD-YFP and AAL are excluded from the nucleolus, as shown by costaining with anti-fibrillarin antibodies (Fig. S7 *A* and *B*). A band corresponding to the observed molecular weight of SRD-YFP was labeled by AAL on Western blots, and AAL reactivity was eliminated upon β-elimination (Fig. 4*E*), indicating that SRD-YFP is modified with *O*-Fuc. In contrast, when the NLS from TgGNC5 (17) was added to the N terminus of YFP, it targeted the fluorescent protein to the nucleoplasm, including the nucleolus (Fig. 4*D* and Fig. S7*A*). Western blots showed that NLS-YFP, despite its relative abundance as compared with SRD-YFP, does not bind AAL, indicating that it is not *O*-fucosylated (Fig. 4*E*).

SNF2 has an SRD at amino acids 1871–1935 (Fig. 4*F*) and three predicted NLS (amino acids 1005–1016, 2402–2410, and 2507–2517). SNF2, which was endogenously tagged with a 3×HA

MICROBIOLOGY

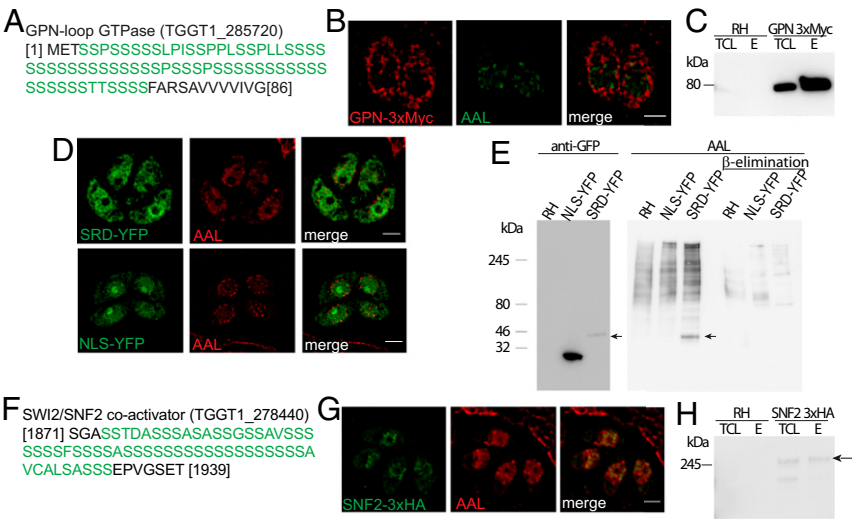

Fig. 4. *O*-fucosylation directs proteins to the nuclear periphery. (*A*) The N-terminal SRD of GPN. (*B*) An ectopic copy of the protein localizes to the cytoplasm and to the nuclear periphery with AAL by ELYRA SIM. (*C*) AAL pull-down followed by anti–c-MYC Western blotting shows GPN enrichment. (*D*) SIM shows that the GPN SRD fused to YFP (SRD-YFP) partially colocalizes with AAL, but NLS-YFP does not. (*E*) AAL recognizes an additional band only in the SRD-YFP cell lysate, consistent with the molecular weight of SRD-YFP as defined by anti-GFP blot (black arrows). (*F*) The SRD of the SNF2 transcriptional coactivator. (*G*) SNF2 colocalizes with AAL at the nuclear periphery. (*H*) AAL pull-down followed by anti-HA blotting shows that SNF2 is present in the AAL-bound fraction (black arrow). E, elution; RH, wild-type cell lysate; TCL, total cell lysate. (Scale bars: 2 μm.)

tag, localizes to the nucleus with AAL (Fig. 4*G*). AAL pull-down confirmed that SNF2 is *O*-fucosylated (Fig. 4*H*).

Superresolution microscopy shows that AAL binds to punctate structures that are adjacent to epichromatin (35), a conformational epitope of DNA in complex with histones H2A and H2B that marks the nuclear periphery (Fig. 5*C*). AAL binding is also found in proximity to euchromatin (Fig. S7*C*), as defined by antibodies against two histone modifications associated with active transcription, H3K9Ac and H3K4Me3 (9). Last, two predicted FG-Nups—Nup68, which is enriched by AAL, and Nup67, which is not—were expressed with a C-terminal YFP under their endogenous promoters (Fig. S6 *C* and *D*). Both fusion proteins partially colocalize with AAL at the nuclear membrane (Fig. 5 *A* and *B*), and in both cases, a subset of the AAL-labeled assemblies partially overlaps with the NPCs (see *Insets* in Fig. 5 *A* and *B*). The glycopeptides observed in Nup68 and other two FG-Nups (TGGT1_203780 and TGGT1_313430) indicate that *O*-Fuc modifies the FG repeats region (Dataset S1).

Discussion

Almost all eukaryotes present a glycosylation pathway dedicated to the modification of cytosolic and nuclear proteins in which *N*-acetylglucosamine is transferred to Ser/Thr in disordered domains by *O*-GlcNAc transferase (OGT) (36). Yeast, one of the few organisms lacking an OGT, recently has been shown to use *O*-mannose (*O*-Man) instead of *O*-GlcNAc to modify its nucleocytosolic proteins (37). In contrast, *T. gondii* has three cytosolic glycosylation pathways: an OGT (38), the hydroxylase and glycosyltransferases (GTs) that modify Skp1 with a pentasaccharide (22), and the nucleocytosolic *O*-fucosylation system described here. In the host and presumably in the parasite, the donor for the OGT reaction, UDP-GlcNAc, is sensitive to the metabolic state of the cell, and modification by *O*-GlcNAc affects protein activity (36). Similarly, glycosylation of Skp1 is required for normal growth of tachyzoites in culture, and proline hydroxylation on Skp1 is sensitive to the redox status of *T. gondii* (22, 29). In contrast, this study suggests that the addition of *O*-Fuc targets *T. gondii* proteins to assemblies closely associated with the nuclear membrane and that targeting of endogenous and exogenous proteins to the AAL-labeled assembly may occur in the absence of an NLS. SRD-YFP and two endogenous proteins, one containing a predicted NLS, are modified with *O*-Fuc and localize to the nuclear periphery, whereas the addition of an NLS to either Cas9 or YFP targets proteins to the nucleoplasm, including the nucleolus. Whether the addition of *O*-Fuc affects the activity of *T. gondii* proteins or protein–protein interactions was not determined.

Although *O*-fucosylation of Ser/Thr residues in secreted proteins of eukaryotic cells has been described previously, here we identify this modification on nuclear proteins. AAL staining suggests this pathway is conserved only in *T. gondii*, *H. hammondi*, and *N. caninum*. Both human fibroblasts (Fig. S2*B*) and bovine turbinate cells (Fig. 1*D*, *S. neurona*) did not show nuclear staining by AAL. Further studies should be performed in different taxonomic groups, but the limited data so far suggest that nuclear *O*-fucosylation may be restricted to these three species. Recognizing the limitations we have encountered thus far in defining the modification site(s) precisely, it appears that the unidentified *T. gondii O*-fucosyltransferase (OFucT) differs from host protein *O*-fucosyltransferases (POFUT1 and POFUT2) in its location and acceptor specificity. First, POFUTs are glycosyltransferases resident in the endoplasmic reticulum (ER) (21), whereas we predict that the *T. gondii* OFucT is either cytosolic or nuclear. Second, the host enzymes transfer *O*-Fuc to epidermal growth factor-like (POFUT1) or thrombospondin type I (POFUT2) repeats, both of which are characterized by conserved disulfide bonds (39, 40), whereas disulfides do not form in cytosolic and nuclear proteins.

Our working model of protein *O*-fucosylation in *T. gondii* is shown in Fig. 5*D*. AAL enrichment and the identified glycopeptides suggest

Table 1. Proteins containing Ser-rich domains identified by AAL pull-down and grouped by function (putative or annotated)

Protein family	Nuclear localization signal +	Nuclear localization signal −	Unique *O*-fuc peptides (proteins*)
Autophagy	1	0	1 (1)
DNA-binding proteins	3	2	3 (3)
Hypothetical proteins	9	15	29 (13)
Kinases/phosphatases	2	4	2 (1)
mRNA processing	5	2	2 (2)
NPC/nuclear transport	2	4	14 (4)
Nucleotide-binding domain	0	2	
Protein–protein interaction	5	5	16 (7)
Transcription regulators	2	1	1 (1)
tRNA synthesis	2	0	
Ubiquitination related	1	2	1 (1)
	32	37	69 (33)
Total	69		

*Number of proteins from the corresponding family for which glycopeptides were observed.

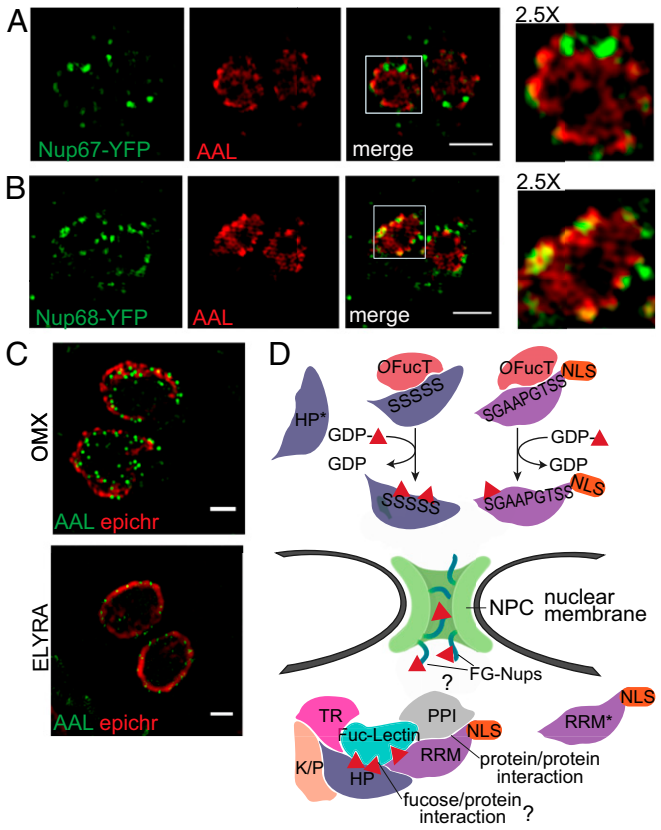

Fig. 5. *O*-fucosylated proteins localize in assemblies at the nuclear membrane near the NPCs. (*A* and *B*) ELYRA SIM shows partial association between AAL-labeled assemblies and Nup67-YFP (*A*) and Nup68-YFP (*B*), one of the FG-Nups bound by AAL. For both fusion proteins, YFP was detected using anti-GFP. Boxed areas are shown at higher magnification. (Scale bars: 2 μm.) (*C*) SIM superresolution microscopy using either OMX or ELYRA shows that AAL labels punctate assemblies at the nuclear periphery. (Scale bars: 1 μm.) (*D*) Model for *O*-fucosylation of nuclear proteins in *T. gondii*: An as yet unidentified OFucT transfers fucose (red triangle) to Ser/Thr residues on SRD-containing proteins. *O*-fucosylated proteins then are shuttled to the nuclear periphery where they associate in assemblies, via protein–protein interactions and, possibly, an endogenous fucose-binding lectin (question mark). How the assembly of *O*-fucosylated proteins interacts with the NPC remains to be determined (question mark). According to the model, nonfucosylated proteins (asterisks) will stay in the cytosol or go to nucleus depending on the presence of an NLS. HP, hypothetical protein; K/P, kinase or phosphorylase domain; PPI, protein–protein interaction domain; RRM, RNA-recognition motif; TR, transcriptional regulator.

that for the most part the acceptors of the putative OFucT are SRDs or proteins containing such domains. However, it seems likely that not all nucleocytosolic proteins with SRDs were identified in the AAL enrichment. In our model, we speculate that an as-yet-unidentified fucose-binding lectin would recognize *O*-fucosylated proteins and participate in their accumulation in assemblies closely associated with the nuclear membrane. This idea would be similar to the host cell secretory pathway in which lectins bind glucosylated and mannosylated N-glycans (41). Many hypothetical proteins, with no homology to any known conserved domains and specific to *T. gondii*, have been identified in the AAL pull-downs, and they may be important in forming the assemblies, i.e., the hypothesized fucose-binding lectin. The AAL-labeled assembly is also likely to contain proteins that are not *O*-fucosylated; in our model, these proteins would associate via protein–protein interactions. Protein–protein interactions, disrupted during lysis, reform during lectin enrichment, and we therefore cannot exclude the possibility that some of the proteins isolated in the pull-down that were categorized as contaminants are actually nonfucosylated members of the assemblies. Isolation of the intact assemblies will be required to discriminate between nonfucosylated proteins that are members of the assemblies and true contaminants. This distinction might be complicated further by the possibility that the protein composition of the assembly is itself heterogeneous. Furthermore, we observed the addition of *O*-Fuc to be a probabilistic event for ∼20% of the identified glycopeptides. According to our model, the fucosylated form of the protein might be present in the AAL-labeled assembly, and the nonfucosylated form might diffuse into the nucleus if an NLS is present or remain in the cytosol if an NLS is absent, as suggested by the localization of GPN-3×MYC (Fig. 4*D*). Whether addition of fucose as a probabilistic event is due to the OFucT mechanism and/or kinetics or is observed because nuclear *O*-Fuc, like *O*-GlcNAc, is a reversible modification cannot be clearly stated at this point. No putative *O*-fucosidase could be identified so far in *T. gondii*.

Multiple pieces of evidence point to the potential importance of the AAL-labeled assemblies in the nucleus of *T. gondii*. First, proteins in AAL pull-downs include numerous putative nucleoporins, mRNA processing enzymes, transcription regulators, and signaling proteins. Second, disruption of GDP-Fuc biosynthesis, which eliminates the binding of AAL to nuclei, appears to affect growth severely, as suggested by our inability to clone the AAL-negative cells. The absence of AAL binding to oocyst nuclei suggests that these proteins are either absent or not glycosylated in these life stages, because electron microscopy showed that the nuclear membrane of sporulating *T. gondii* remains intact (42). Last, five of seven predicted nucleoporins are present in the AAL enrichments, and tagged versions of Nup68 and Nup67 partially colocalize with AAL-labeled assemblies. Furthermore, MS showed that *O*-Fuc is found on three FG-Nups: Nup68, TGGT1_203780 [the second most abundant protein in the pull-down (Dataset S2)], and TGGT1_313430 [the *T. gondii* ortholog of yeast Nup98/96 (43)]. In all three instances, the sugar modifies the FG repeats region. FG regions are the disordered sequences that characterize the NPC channel and interact with karyopherins to mediate nuclear transport (19). In higher eukaryotes and yeast, FG regions are highly decorated with *O*-GlcNAc and *O*-Man, respectively, and evidence suggests that this PTM could affect cargo selectivity (36, 37).

O-fucosylation thus appears to be a mechanism by which *T. gondii* proteins involved in gene expression and mRNA processing (gating hypothesis) are gathered at the nuclear membrane, often in close proximity to the NPCs (14), and components of the NPC itself are *O*-fucosylated.

Methods

Parasite Cell Culture and Manipulation. *T. gondii* type I RH tachyzoites culture, manipulation, and in vitro differentiation to bradyzoites were performed as previously described (7, 44). The RH Δ*ku80* (44) transgenic cell line has been described previously. RH Δ*ku80* Δ*phyA*, Δ*gnt1*, and Δ*pgtA* (22) were kind gifts from Christopher West, University of Georgia, Athens, GA and Ira Blader, SUNY at Buffalo School of Medicine, Buffalo, NY. RH SNF2-3xHA was a kind gift from William Sullivan, Jr, Indiana University School of Medicine, Indianapolis. The generation of transgenic cells expressing GPN-3xMYC, SRD-YFP, NLS-YFP, Nup68-YFP, and Nup67-YFP is described in *SI Methods*. All animal work was approved by the Institutional Animal Care and Use Committee at Boston University. *C. parvum* (Iowa strain) oocysts were purchased from Bunch Grass Farm. *S. neurona* SN3 was a kind gift from Michael Grigg, NIH, Bethesda, and *P. falciparum* was a kind gift from Jeffrey Dvorin, Harvard Medical School, Boston. *N. caninum* Nc-1 and *H. hammondi* were kind gifts of Jon Boyle, University of Pittsburgh, Pittsburgh. For oocyst excystation and fixation protocols, see *SI Methods*.

DNA Manipulation. All primers used in this study are listed in Table S1. Cloning strategies and plasmids used are described in *SI Methods*.

Immunofluorescence Microscopy. Fixation, permeabilization, blocking, and labeling conditions are described in *SI Methods*. Deconvolution microscopy was performed on a Olympus XI70 inverted microscope and DeltaVision SoftWoRx as previously described (45). Superresolution microscopy was

performed on an OMX-3D, V3 type (Applied Precision) or on an ELYRA (Zeiss) microscope. In both cases images were acquired with a 100×/1.4 oil immersion objective, 0.125-μm *z* sections at room temperature. For the OMX microscope, images were processed with OMX SoftWoRx software, and projections of 10× *z* stacks are shown. For the ELYRA microscope, images were processed for structured illumination (SIM) using ZEN software, and single optical sections were selected using Fiji (46).

Lectin Pull-Down. Proteins were pulled down from total cell lysate by incubation with biotinylated AAL (Vector Labs) and Dynabeads MyOne Streptavidin T1 (Thermo Fisher) and were eluted by incubation with αMeFuc (Carbosynth). See *SI Methods* for details (47).

MS. Proteins in the elution were reduced, alkylated, and digested with trypsin either in solution or in gel. Peptides and glycopeptides were analyzed by LC-MS/MS using a ChIP or an ultra-performance LC (UPLC) C18 column on a 6550-QTOF (Agilent Technologies), an LTQ-Orbitrap-XL-ETD, or a Q Exactive Plus Quadrupole hybrid Orbitrap (Thermo Scientific) MS system. The PTM search function within the analysis software PEAKS (Bioinformatic Solutions) (48) was used to produce a list of putative *O*-fucosylated peptides that was manually verified. Scaffold (Proteome Software) (49) was used to compare the five biological repeats, and further analyses were performed using RStudio. Monosaccharide composition analysis was performed on a Bruker Scion-SQ GC-MS (Bruker). See *SI Methods* for details.

ACKNOWLEDGMENTS. We thank Dr. Aparajita Chatterjee and Dr. Carolina Agop-Nersesian for invaluable assistance with oocyst excystations; Dr. Christopher West for helpful discussions; Dr. Eliza Vasile at the Swanson Biotechnology Center Microscopy Core Facility for technical assistance with OMX microscopy; and the Harvard Centre for Biological Imaging (HCBI) for infrastructure and support. The ELYRA microscope was acquired through NIH Shared Instrumentation Grant Award S10 RR27990 (to the HCBI). Support for this study came from NIH Grants R01 AI110638 (to J.S.), R01 GM031318 (to P.W.R.), and P41 GM104603 (to C.E.C.), NIH-National Heart, Lung, and Blood Institute Contract HHSN268201000031C (to C.E.C.), and a grant from the Mizutani Foundation of Glycoscience (to J.S.).

1. Torgerson PR, Mastroiacovo P (2013) The global burden of congenital toxoplasmosis: A systematic review. *Bull World Health Organ* 91(7):501–508.
2. Torrey EF, Yolken RH (2013) *Toxoplasma* oocysts as a public health problem. *Trends Parasitol* 29(8):380–384.
3. Lorenzi H, et al. (2016) Local admixture of amplified and diversified secreted pathogenesis determinants shapes mosaic *Toxoplasma gondii* genomes. *Nat Commun* 7:10147.
4. Behnke MS, Zhang TP, Dubey JP, Sibley LD (2014) *Toxoplasma gondii* merozoite gene expression analysis with comparison to the life cycle discloses a unique expression state during enteric development. *BMC Genomics* 15(1):350.
5. Radke JR, et al. (2005) The transcriptome of *Toxoplasma gondii*. *BMC Biol* 3:26.
6. Balaji S, Babu MM, Iyer LM, Aravind L (2005) Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* 33(13):3994–4006.
7. Walker R, et al. (2013) The *Toxoplasma* nuclear factor TgAP2XI-4 controls bradyzoite gene expression and cyst formation. *Mol Microbiol* 87(3):641–655.
8. Walker R, et al. (2013) *Toxoplasma* transcription factor TgAP2XI-5 regulates the expression of genes involved in parasite virulence and host invasion. *J Biol Chem* 288(43):31127–31138.
9. Nardelli SC, et al. (2013) The histone code of *Toxoplasma gondii* comprises conserved and unique posttranslational modifications. *MBio* 4(6):e00922–e13.
10. Bhatti MM, Livingston M, Mullapudi N, Sullivan WJ, Jr (2006) Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell* 5(1):62–76.
11. Saksouk N, et al. (2005) Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. *Mol Cell Biol* 25(23):10301–10314.
12. Blobel G (1985) Gene gating: A hypothesis. *Proc Natl Acad Sci USA* 82(24):8527–8529.
13. Gordon MR, Pope BD, Sima J, Gilbert DM (2015) Many paths lead chromatin to the nuclear periphery. *BioEssays* 37(8):862–866.
14. Burns LT, Wente SR (2014) From hypothesis to mechanism: Uncovering nuclear pore complex links to gene expression. *Mol Cell Biol* 34(12):2114–2120.
15. Gissot M, et al. (2012) *Toxoplasma gondii* chromodomain protein 1 binds to heterochromatin and colocalises with centromeres and telomeres at the nuclear periphery. *PLoS One* 7(3):e32671.
16. Brooks CF, et al. (2011) *Toxoplasma gondii* sequesters centromeres to a specific nuclear region throughout the cell cycle. *Proc Natl Acad Sci USA* 108(9):3767–3772.
17. Bhatti MM, Sullivan WJ, Jr (2005) Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite T*oxoplasma gond*ii. *J Biol Chem* 280(7):5902–5908.
18. Dixon SE, Bhatti MM, Uversky VN, Dunker AK, Sullivan WJ, Jr (2011) Regions of intrinsic disorder help identify a novel nuclear localization signal in Toxoplasma gondii histone acetyltransferase TgGCN5-B. *Mol Biochem Parasitol* 175(2):192–195.
19. Frankel MB, Knoll LJ (2009) The ins and outs of nuclear trafficking: Unusual aspects in apicomplexan parasites. *DNA Cell Biol* 28(6):277–284.
20. Field MC, Koreny L, Rout MP (2014) Enriching the pore: Splendid complexity from humble origins. *Traffic* 15(2):141–156.
21. Ma B, Simala-Grant JL, Taylor DE (2006) Fucosylation in prokaryotes and eukaryotes. *Glycobiology* 16(12):158R–184R.
22. Rahman K, et al. (2016) The e3 ubiquitin ligase adaptor protein SKP1 is glycosylated by an evolutionarily conserved pathway that regulates protist growth and development. *J Biol Chem* 291(9):4268–4280.
23. Chen C-T, Gubbels M-J (2013) The *Toxoplasma gondii* centrosome is the platform for internal daughter budding as revealed by a Nek1 kinase mutant. *J Cell Sci* 126(Pt 15):3344–3355.
24. Ouologuem DT, Roos DS (2014) Dynamics of the *Toxoplasma gondii* inner membrane complex. *J Cell Sci* 127(Pt 15):3320–3330.
25. Sanz S, et al. (2013) Biosynthesis of GDP-fucose and other sugar nucleotides in the blood stages of *Plasmodium falciparum*. *J Biol Chem* 288(23):16506–16517.
26. Gajria B, et al. (2008) ToxoDB: An integrated *Toxoplasma gondii* database resource. *Nucleic Acids Res* 36(Database issue):D553–D556.
27. Anderson-White B, et al. (2012) Cytoskeleton assembly in *Toxoplasma gondii* cell division. *Int Rev Cell Mol Biol* 298:1–31.
28. Shen B, Brown KM, Lee TD, Sibley LD (2014) Efficient gene disruption in diverse strains of Toxoplasma gondii using CRISPR/CAS9. *MBio* 5(3):e01114–e14.
29. Xu Y, et al. (2012) The Skp1 protein from *Toxoplasma* is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba *Dictyostelium*. *J Biol Chem* 287(30):25098–25110.
30. Kosugi S, et al. (2009) Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *J Biol Chem* 284(1):478–485.
31. Silmon de Monerri NC, et al. (2015) The ubiquitin proteome of *Toxoplasma gondii* reveals roles for protein ubiquitination in cell-cycle Transitions. *Cell Host Microbe* 18(5):621–633.
32. Treeck M, Sanders JL, Elias JE, Boothroyd JC (2011) The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe* 10(4):410–419.
33. Ashburner M, et al.; The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25(1):25–29.
34. Marchler-Bauer A, et al. (2013) CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Res* 41(Database issue):D348–D352.
35. Vanagas L, et al. (2013) Epichromatin is conserved in *Toxoplasma gondii* and labels the exterior parasite chromatin throughout the cell cycle. *Parasitology* 140(9):1104–1110.
36. Bond MR, Hanover JA (2015) A little sugar goes a long way: The cell biology of *O*-GlcNAc. *J Cell Biol* 208(7):869–880.
37. Halim A, et al. (2015) Discovery of a nucleocytoplasmic *O*-mannose glycoproteome in yeast. *Proc Natl Acad Sci USA* 112(51):15648–15653.
38. Banerjee S, Robbins PW, Samuelson J (2009) Molecular characterization of nucleocytosolic *O*-GlcNAc transferases of *Giardia lamblia* and *Cryptosporidium parvum*. *Glycobiology* 19(4):331–336.
39. Valero-González J, et al. (2016) A proactive role of water molecules in acceptor recognition by protein *O*-fucosyltransferase 2. *Nat Chem Biol* 12(4):240–246.
40. Al-Shareffi E, et al. (2013) 6-alkynyl fucose is a bioorthogonal analog for *O*-fucosylation of epidermal growth factor-like repeats and thrombospondin type-1 repeats by protein *O*-fucosyltransferases 1 and 2. *Glycobiology* 23(2):188–198.
41. Helenius A, Aebi M (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 73:1019–1049.
42. Ferguson DJ, Birch-Andersen A, Siim JC, Hutchison WM (1979) An ultrastructural study on the excystation of the sporozoites of *Toxoplasma gondii*. *Acta Pathol Microbiol Scand [B]* 87(5):277–283.
43. Iwamoto M, Asakawa H, Hiraoka Y, Haraguchi T (2010) Nucleoporin Nup98: A gatekeeper in the eukaryotic kingdoms. *Genes Cells* 15(7):661–669.
44. Huynh M-H, Carruthers VB (2009) Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. *Eukaryot Cell* 8(4):530–539.
45. Bushkin GG, et al. (2012) β-1,3-glucan, which can be targeted by drugs, forms a trabecular scaffold in the oocyst walls of *Toxoplasma* and *Eimeria*. *MBio* 3(5):e00258–e12.
46. Schindelin J, et al. (2012) Fiji: An open-source platform for biological-image analysis. *Nat Methods* 9(7):676–682.
47. Güther MLS, et al. (2009) Fate of glycosylphosphatidylinositol (GPI)-less procyclin and characterization of sialylated non-GPI-anchored surface coat molecules of procyclic-form *Trypanosoma brucei*. *Eukaryot Cell* 8(9):1407–1417.
48. Zhang J, et al. (2012) PEAKS DB: De novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* 11(4):M111.010587–M111.010587.
49. Searle BC (2010) Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* 10(6):1265–1269.
50. Walzer KA, et al. (2013) *Hammondia hammondi*, an avirulent relative of *Toxoplasma gondii*, has functional orthologs of known *T. gondii* virulence genes. *Proc Natl Acad Sci USA* 110(18):7446–7451.
51. Chatterjee A, et al. (2010) Evidence for mucin-like glycoproteins that tether sporozoites of *Cryptosporidium parvum* to the inner surface of the oocyst wall. *Eukaryot Cell* 9(1):84–96.
52. Sidik SM, Hackett CG, Tran F, Westwood NJ, Lourido S (2014) Efficient genome engineering of *Toxoplasma gondii* using CRISPR/Cas9. *PLoS One* 9(6):e100450.