Characterization of a cytoplasmic glucosyltransferase that extends the core trisaccharide of the Toxoplasma Skp1 E3 ubiquitin ligase subunit*

Kazi Rahman1,4,8, Msano Mandalasi1,5, Peng Zhao2, M. Osman Sheikh2, Rahil Taujale2,3, Hyun W. Kim1, Hanke van der Wel1, Khushi Matta6, Natarajan Kannan1,3, John N. Glushka2, Lance Wells1,2, Christopher M. West1,5,7

1Department of Biochemistry & Molecular Biology, 2Complex Carbohydrate Research Center, 3Institute of Bioinformatics, University of Georgia, Athens, GA 30602 USA; 4Departments of Microbiology & Immunology, and 5Biochemistry & Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104 USA; 6Chemical and Biological Engineering, State University of New York, Buffalo, NY 14260 USA

Running title: A Novel Cytoplasmic Glucosyltransferase

Keywords: Toxoplasma gondii; Skp1; E3 ubiquitin ligase; SCF-complex; cytoplasmic glycosylation; linkage analysis; glycobiology; parasitology

Skp1 is a subunit of the SCF (Skp1/Cullin-1/F-box protein) class of E3 ubiquitin ligases that are important for eukaryotic protein degradation. Unlike its animal counterparts, Skp1 from Toxoplasma gondii is hydroxylated by an O2-dependent prolyl-4-hydroxylase (PhyA), and the resulting hydroxyproline can subsequently be modified by a five-sugar chain. A similar modification is found in the social amoeba Dictyostelium, where it regulates the social amoeba Dictyostelium, generating GaLa1,3GaLo1,3FucO1,2Galβ1,3GlcNAcO1-. Here, we found that Toxoplasma utilizes a cytoplasmic glucosyltransferase from an ancient clade of CAZy family GT32 to catalyze transfer of the terminal disaccharide in cells, and cytosolic extracts catalyzed transfer of [3H]glucose from UDP-[3H]glucose to the trisaccharide form of Skp1 in a gtl1-dependent fashion. Recombinant Glt1 catalyzed the same reaction, confirming that it directly mediates Skp1 glucosylation, and NMR demonstrated formation of a Glex1,3Fuc linkage. Recombinant Glt1 strongly preferred the full core trisaccharide attached to Skp1, and labeled only Skp1 in gtl1a extracts, suggesting specificity for Skp1. gtl1-knockout parasites exhibited a growth defect not rescued by catalytically inactive Glt1, indicating that the glycan acts in concert with the first enzyme in the pathway, PhyA, in cells. A genomic bioinformatics survey suggested that Glt1 belongs to the ancestral Skp1 glycosylation pathway in protists and evolved separately from related Golgi-resident GT32 glycosyltransferases.

Skp1 is an adaptor subunit of the Skp1/Cullin-1/F-box protein (SCF)9-class of E3 ubiquitin ligases that target proteins for polyubiquitination and degradation via the 26S proteasome (1). In the agent for toxoplasmosis, Toxoplasma gondii, Skp1 is hydroxylated by the cytoplasmic prolyl 4-hydroxylase PhyA at Pro154 (2), and subsequently modified by a linear pentasaccharide (3). In the social amoeba Dictyostelium, Skp1 is also modified by hydroxylation and a pentasaccharide, which represents a novel form of SCF regulation (4). Biochemical and interactome studies indicate that full glycosylation of Skp1 promotes association with three different F-box proteins (FBPs) (5, 6) which, in a developmentally regulated manner, is associated with their reduced steady state levels in cells.10 Many FBPs are substrate receptors for ligands whose polyubiquitination controls their abundance, whereas others are considered to have enzymatic or other functions (2). FBPs possess a 40-amino acid F-box domain that binds to the C-terminal region of Skp1 (7, 8). Recent studies show that glycosylation influences the organization and
range of motions of this region of Skp1, in part by hydrogen bonding along the polypeptide in cis (9). Glycosylation is regulated by PhyA, whose action on Skp1 is rate-limited by O₂-availability in the cell (10, 6). This biochemical mechanism underlies cellular O₂-sensing, which controls the slug-to-fruit switch and sporation during starvation-induced development (4). O₂ signifies positional information in the Dictyostelium’s native soil environment and sensing O₂ is key for the ability of developing cells to navigate to the soil surface for fruiting body formation (11).

PhyA is important for Toxoplasma tachyzoite proliferation on cultured human fibroblasts (12), which involves successive cycles of invasion, intracellular replication, egress, and reinvasion to form cell free plaques. Glycosylation of Toxoplasma Skp1 is important too, because disruption of genes that mediate addition of the first three monosaccharides also result in reduced parasite growth (3). However, the significance of the full pentasaccharide has not been examined because the gene that mediates addition of the final two sugars on the Dictyostelium pentasaccharide, agtA (13), is evidently absent from the Toxoplasma genome (14).

In order to investigate the function of the terminal disaccharide of the Toxoplasma Skp1 glycan, we searched for glycosyltransferase-like genes in the parasite’s genome whose protein products are predicted to reside in the cytoplasm or nucleus, and whose phylogenetic distribution correlates with the presence of the pgtA gene that mediates addition of the second and third monosaccharides (15). This search netted two genes, not found in the amoebozoa including Dictyostelium. Characterization of one of these (TGGT1_205060), which we named glt1, revealed a previously non-annotated gene that encodes a cytoplasmic glycosyltransferase from CAZy glycosyltransferase family 32 that modifies the Fuc terminus of the Skp1 trisaccharide. Gene disruption and complementation studies show that the fourth sugar contributes to Skp1 functionality in parasite growth. Thus, these unrelated protists employ different mechanisms to assemble a related though distinct pentasaccharide to regulate, presumably, the critical process of protein turnover in cells. A deeper understanding of the enzymes involved may offer unique strategies to control toxoplasmosis, which affects a large fraction of the world’s human population and for which control of acute and chronic phases are lacking (16).

Results

Prediction of Candidate TgSkp1 Modifying Glycosyltransferases—To identify candidate proteins for catalyzing addition of the fourth and fifth Skp1 monosaccharides, genome sequence databases were searched for predicted glycosyltransferase domains that reside in the cytoplasm, and exist only in protists whose genomes harbor other predicted Skp1 modification pathway enzymes, phyA, gnt1, and pgtA, but not agtA, as outlined in Fig 1. Briefly, the predicted proteome (8460 proteins) of the T. gondii Type I GT1 strain was searched using i) the SUPERFAMILY server, which assigns protein domains at the SCOP ‘superfamily’ level using hidden Markov models; ii) dbCAN, an automated Carbohydrate-active enzyme Annotation database which utilizes a CAZyme signature domain-based annotation based on a CDD (conserved domain database) search, literature curation, and a hidden Markov model; and iii) the Pfam database. The resulting 45 sequences were scanned by SignalIP 4.1 and TMHMM servers for signal sequences or transmembrane domains, which yielded 10 candidate cytoplasmic glycosyltransferases. Among those, TgGnt1 and TgPgtA were already identified as TgSkp1 modifying glycosyltransferases (3). The remaining 8 sequences were subjected to BlastP analysis against the NCBI non-redundant database to search for their phylogenetic co-distribution with gnt1 and pgtA. This yielded two candidates, one from CAZy family GT32, TGGT1_205060, and another from CAZy family GT8. The GT32 sequence, to be referred to as Glt1, is the subject of this study.

Glt1 Sequence Characteristics and Phylogeny—In the Toxoplasma database ToxoDB V29, TGGT1_205060 (type 1 GT1 strain) is annotated as a 605-amino acid protein encoded by five exons, that is conserved in ME49 (type II) and VEG (type III) strains with 99% nt sequence identity. The protein contains a DxD motif typical of and essential for CAZy GT32 family and superfamily A glycosyltransferases (17). Proteomics analyses of Toxoplasma, as reported on ToxoDB, identified five Glt1 peptides, one of which was phosphorylated at Ser601 (Fig. S1), confirming that the gene is expressed in tachyzoites grown in...
Glt1 is Required for TgSkp1 Glycosylation—To determine whether glt1 is involved in Skp1 glycosylation, the gene was replaced by double-crossover homologous recombination in the type 1 RHΔΔ strain, as illustrated in Fig. 3A. glt1 deletion mutants were confirmed by negative PCR reactions for glt1-coding DNA, and positive PCR products for the insertion of the selection marker hxgprt between glt1 flanking sequences (Fig. 3B). To control for off-target genetic modifications, a complementation construct containing a version of the original genomic DNA was used to replace the hxgprt locus in glt1 disruption clone-1 using counter-selection for loss of hxgprt (Fig. 3A). The same set of PCRs were used to confirm the desired gene restoration in clonal isolates (Fig. 3C). The same clone was also transformed with a mutant version in which three conserved Asp residues, including the DxD motif at positions 363 and 365 and another at codon 348, to generate a potentially inactive mutant Glt1(D363A/D365A/D348N). Strains are listed in Table 1.

Western blot analysis of whole cells suggested that Skp1 from parasites lacking glt1 (Fig. 4A, lane 3) migrates slightly more rapidly than wild-type Skp1 (lane 1 and lane 6) but slower than Skp1 from a phyAΔ strain (lane 2). Complementation with wild-type glt1 (lane 4), but not the triple mutant (lane 5), restored normal mobility. This confirmed the specificity of the genetic disruption to glt1, and further implicated Glt1 as a Skp1 glycosyltransferase. To confirm an effect on Skp1 glycosylation and to pinpoint its location, Skp1 was immunoprecipitated, converted to peptides using trypsin, and analyzed by nanoLC/MS-MS in an Orbitrap Fusion mass spectrometer. As shown in the extracted ion chromatograms from a sample from the glt1-complemented strain (Fig. 4B, left side, dashed box-1), doubly- and triply-charged ions with exact mass matches to the peptide (145IFNIVNDFTPREEAQVR161) that bears the known modification site, Pro154, were readily detected. As expected, the pentasaccharide form was also readily detected at an earlier elution time (Fig. 4E, left side, dashed box-2), as reported previously for the normal parental strain (3) and shown in Table 2. In addition, a very low level of the disaccharide form was apparent in panel C, but no trisaccharide form was detected (panel D). In contrast, Skp1 from a glt1Δ clone yielded no detectable pentasaccharide, but abundant trisaccharide (box-3), as well as the unmodified peptide and a minor level of the disaccharide form. The doubly-charged version of this putative dHex-
Hex-HexNAc-peptide (Fig. 5A) was subjected to further CID analysis, yielding a series of cleavage products confirming the order of sugar residues (Fig. 5B, Table 2), and the sequence of the peptide and its attachment, as expected, at a hydroxylated form of Pro154 (panel C). In a clone in which glt1Δ was complemented with mutated (D363A, D365A, D348N) glt1 genomic DNA, the profiles resembled those of the glt1Δ strain (Table 2), demonstrating the importance of its predicted glycosyltransferase activity. These findings implicate Glt1 as the glycosyltransferase that catalyzes addition of the fourth sugar, and indicates that addition of the fifth sugar depends on the fourth.

**Toxoplasma Expresses a glt1-Dependent Skp1 Glucosyltransferase Activity**—To characterize the implied activity of Glt1, exogenous radioactive sugar nucleotide donors and Skp1 glycoforms were introduced into soluble parasite extracts in an attempt to recapitulate reactions in the cell. Since Toxoplasma and Dictyostelium apparently share the same core trisaccharide on their Skp1’s, and their Skp1 sequences are highly conserved and serve as substrates for each other’s PhyA, Gnt1 and PgtA enzymes (12, 3), we used recombinantly expressed Dictyostelium FGGn-Skp1 (13) as a surrogate acceptor substrate. *In vitro* reactions were performed with cytosolic extracts in the presence of UDP-[3H]Gal, based on the occurrence of Gal in Dictyostelium Skp1 (Fig. 6A). Substantial activity was detected, which was time- and Skp1-dependent, and absent from glt1Δ extracts. To confirm the chemical identity of the transferred 3H, the radioactive Skp1 band was excised after electrophoretic transfer to a PVDF membrane, and subjected to acid hydrolysis and high pH anion exchange chromatography. Surprisingly, radioactivity co-chromatographed with D-Glc (Fig. 6B). New reactions using UDP-[3H]Glc revealed more efficient incorporation (Fig. 6A), whose product also co-eluted with the Glc standard (data not shown). The incorporation of 3H from UDP-[3H]Gal was likely the result of the action of TGGT1_225880, a putative UDP-Glc 4-epimerase in the extract, which would explain the lower incorporation of 3H from this substrate. Therefore, Glt1 is inferred to utilize UDP-Glc as its substrate.

**Activity of Recombinant Glt1**—To test whether Glt1 is capable of directly glucosylating FGGn-Skp1, a codon-optimized Toxoplasma glt1 cDNA (Fig. S1) was expressed in and purified from **E. coli** as an N-terminally His6-tagged protein. After enrichment over a Co\(^{3+}\)-Talon column (Fig. 7A), Western blot analysis of fractions from a Q-anion exchange column revealed an anti-His6-reactive band at the expected \( M_r \) value of 69,500 (Fig. 7C). To simplify detection of enzymatic activity, a synthetic small molecule, FGGn-pNP, was used as a surrogate for FGGn-Skp1, and its validity is described below. A glucosyltransferase activity catalyzing the transfer of \(^3\)H from UDP-[\(^3\)H]Glc to FGGn-pNP coeluted with the 69,500 band (Fig. 7B). Inspection of a parallel gel stained for total protein with Coomassie blue indicated a prominent band that co-migrated with the His6-band (Fig. 7D), suggesting that the Glt1 protein was major component of the preparation. The great majority of Glt1 was full-length (not shown), and densitometry indicated that Glt1 represented 14% of the protein in fraction 21 (not shown). Pilot studies showed that the transferase activity was stable on ice for days and freeze-thawing, most active in the presence of bovine serum albumin, unaffected by NaCl concentration over the range of 10-400 mM, and more active at increasing pH values up to the highest value tested (pH 8.5). Activity was blocked by the addition of EDTA, consistent with the importance of the DxD sequence for divalent cation coordination. A recombinant version of the triple (D348N/D363A/D365A) mutant described above, as well as each of the individual point mutants, expressed well but were inactive (data not shown). MnCl\(_2\), but not MgCl\(_2\), supported activity, and 2 mM was sufficient for maximal activation. These data guided the design of the standard reaction to examine the substrate specificities of the enzyme.

Most glycosyltransferases are capable of specifically hydrolyzing their donor substrates by transferring the sugar to water in the absence of an appropriate acceptor (18). A screen for ability of purified His6-Glt1 to hydrolyze 6 different UDP-sugar donors, based on generation of GDP, revealed strong selectivity for UDP-Glc (Fig. 8A), which was consistent with the transferase activity in cytosolic extracts (Fig. 6). GDP-Man was not a substrate in a glycosyltransferase assay utilizing FG-Bn, which mimics the non-reducing terminal disaccharide of the Skp1 glycan, using a variation of the UDP-Glo assay that measures the generation of GDP or UDP (Fig. 8B). Kinetic analysis of UDP generation yielded a \( K_m \) for UDP-Glc of 7.5 ± 2.3 \( \mu \)M. Transferase activity toward FG-Bn was also
monitored as the incorporation of \(^{3}H\) from UDP-[\(^{3}H\)]Glc into FG-Bn. These data confirmed the expected inverse hyperbolic dependence on UDP-Glc concentration and yielded an apparent \(K_m\) of 6.0 ± 0.77 \(\mu\)M (Fig. 8C), consistent with the value from the other assay. UDP was a more potent inhibitor than GDP of this transferase reaction (data not shown), confirming that GDP-Man is not a substrate.

A comparison of FG-Bn and FGGn-pNP acceptors showed markedly improved time-dependent activity toward the full trisaccharide (Fig. 8D), though an effect of the different aglycon moieties cannot be excluded. Fuc-pNP alone was inactive as an acceptor, as were various other mono- and di-saccharides representing facets of the Skp1 or related glycans (Fig. 8E). Glt1 transferase activity toward varied concentrations of FGGn-pNP yielded a \(K_m\) of 13 mM and a \(V_{max}\) of 330 nmol/h/\(\mu\)g, or about 6 nmol Glc/nmol Glt1/sec (Fig. 8F). To compare with the native substrate, *Toxoplasma* Skp1 was co-expressed in *E. coli* with *Dictyostelium* PhyA and Gnt1, purified to near homogeneity under non-denaturing conditions, and modified to completion *ex vivo* using *Dictyostelium* PgtA. Activity toward *Toxoplasma* FGGn-Skp1 yielded an apparent \(K_m\) of over 3 orders of magnitude lower than toward FGGn-pNP, 4.2 \(\mu\)M, and an apparent \(V_{max}\) that was reduced by about 200-fold (Fig. 8G). Thus at the low \(\mu\)M concentrations expected in the cell, Glt1 is calculated to exhibit a strong preference for the FGGn-trisaccharide when attached to Skp1 compared to unknown potential aglycons.

To question the ability of Glt1 to accumulate in *glt1Δ* parasites and be susceptible to glucosylation in extracts. This was tested by incubation of desalted wild-type or *glt1Δ* parasites extracts in the presence of His\(_6\)Glt1 and UDP-[\(^{3}H\)]Glc, followed by separation of the entire reaction mixture by SDS-PAGE and scintillation counting of gel slices. As shown in Fig. 8H, a high level of incorporation of \(^{3}H\) was observed at the position of Skp1 in the *glt1Δ* extract. Incorporation was dependent on the addition of His\(_6\)Glt1, and was not observed in parental (RH\(\Delta\)) extracts consistent with Skp1 being already modified before cell extraction. Minimal incorporation was detected at higher \(M_t\) positions, but was also observed in the controls.

Thus within the sensitivity of the method and assuming that potential other Glt1 substrates are accessible and not alternatively processed in its absence, Glt1’s main or only substrate is Skp1.

**Linkage position of Glc–** In order to characterize the glycosidic linkage of the Glc residue, FGGn-pNP was used as the acceptor substrate in a scaled up reaction with Glt1, and the glycan product was recovered by solid phase extraction on a C\(_{18}\) Sep-Pak and analyzed by NMR. A 1-D \(^{1}H\)-NMR spectrum revealed that the sample was of high purity, but also contained many overlapping peaks in the proton dimension (Fig. 9A, between ~3.5 and 5.5 ppm). Much of the peak overlap in the proton dimension was resolved in the two-dimensional \(^{1}H\)--\(^{13}C\)-HSQC spectrum, which was provisionally assigned using the CASPER program (19) (Fig. 9B). Assignments were confirmed by analysis of the 2D COSY, TOCSY and HMBC spectra (data not shown). The HMBC spectra of the Glt1 reaction product also revealed through bond connectivities between the anomic carbons to the ring protons (Fig. 9C), and between the ring carbons and anomic protons (Fig. 9D), clearly demonstrating the glycosidic linkage between the terminal αGlc and underlying αFuc as 1→3. Taken together, the NMR analyses are most consistent with the Skp1 glycan structure: Glcα1,3Fuca1,2Galβ1,3GlcNAca1-. Thus, Glt1 is a UDP-Glc: fucoside α1,3-glucosyltransferase.

**Glt1 is Important for Toxoplasma Proliferation**– The ability of the parasite to infect and proliferate on a monolayer culture of fibroblasts, as measured from the area of clearance (plaques) of cells, is a model for potential virulence in animals. Confluent monolayers were infected with parental, mutant, and complemented parasites and plaque areas analyzed 5.5 d later. As shown in Fig. 10, the average plaque area generated by *glt1Δ* clones is reduced compared to that of parental (RH\(\Delta\)) cells, but larger than that of *phyAΔ* strains that lack the entire modification on Skp1. There was no evidence for reduced plating efficiency (data not shown). Genetic complementation with the original *glt1* sequence at the same locus restored normal growth, verifying that the growth defect in the original disruption strain was due to disruption of *glt1*. In contrast, complementation with the enzymatically inactive mutant version of *glt1* failed to rescue the growth defect of *glt1Δ* cells. Therefore, the slow growth of the strain can be
attributed to loss of the enzymatic activity itself, rather than another potential function of Glt1.

Discussion

glt1Δ parasites exhibit a growth defect in fibroblast cultures. The effect was specific for glt1, as complementation with glt1 restored normal growth. Failure to complement with a catalytically inactive sequence demonstrated the importance of the catalytic activity of Glt1. The deficit was not as strong as that of disrupting phyA, an earlier gene in the Skp1 modification pathway. Key to understanding how glt1 contributes to parasite growth, and interpreting the difference between disrupting phyA and glt1, is characterizing the biochemical contributions of the glt1 gene product.

Glt1 is a Novel α3-Glucosyltransferase—Here we find that Glt1, an enzyme from the CAZy GT32 family of retaining glycosyltransferases, mediates addition of the fourth sugar on Toxoplasma Skp1. This enabled us, in turn, to infer the sugar to be the pyranose form of D-Glc in α-linkage to the 3-position of the underlying Fuc. This conclusion is based on i) the mass spectrometric analysis of Skp1 that shows that disruption of glt1 leads to accumulation of the truncated trisaccharide form of Skp1 (FGGn-Skp1) suggesting inability to transfer the fourth sugar, a residue of hexose (Figs. 4, 5; Table 2), ii) the ability of parasite cell extracts to transfer Glc from UDP-α-D-Glc to FGGn-Skp1 by a mechanism that depends upon glt1 (Fig. 6), and iii) the ability of recombinant His6Glt1 (Fig. 7) to directly, specifically and efficiently catalyze the addition of Glc to the Skp1 trisaccharide (FGGn-Skp1) and synthetic glycan models (Fig. 8). UDP-Glc was the only UDP-sugar, of the six tested, to be efficiently hydrolyzed by purified Glt1 (Fig. 8A). GDP-Man was not a substrate (Fig. 8B). Furthermore, UDP-Glc is likely to be the native substrate of Glt1 owing to its ability to efficiently transfer Glc to the synthetic acceptor FGGn-pNP (Fig. 8F), at a calculated turnover number of about 6/sec at maximal velocity, which is rapid for a glycosyltransferase. The Glt1 protein sequence is related to the CAZy GT32 family, and the loss of enzymatic activity of point mutants that inactivate other members of this family supports this association (data not shown). The CAZy GT32 family includes several characterized retaining glycosyltransferases that use either UDP-Gal or GDP-Man, though none are known to utilize UDP-Glc, and catalyze formation of α1,3, α1,4, or α1,6 linkages. The NMR-HMBC analysis confirms the α-linkage, and furthermore establishes glycosidic attachment to the 3-OH of the underlying Fuc (Fig. 9).

Skp1 is likely the Natural Substrate of Glt1—Glt1 and Skp1 both reside in cytoplasmic and possibly nuclear compartments based on the absence of apparent signal peptides or transmembrane domains, consistent with their recovery from cytosolic Toxoplasma extracts. Glt1 has a strong preference for non-reducing terminal Fuc in the context of the native Skp1 trisaccharide, relative to the non-reducing disaccharide, and did not modify GlcNAc or Gal in different contexts (Fig. 8D, E). Furthermore, it did not modify Fuc alone, indicating that it will not target the abundant O-Fuc modifications in the nucleus of Toxoplasma (21). FGGn-Skp1 is also an excellent substrate, with an apparent $K_m$ of 4 μM, dramatically lower than that of FGGn-pNP and consistent with favorable selectivity for the trisaccharide in the context of Skp1. What FGGn-Skp1 gains in $K_m$ is, however, sacrificed in $V_{max}$. Nevertheless, at 1 μM concentrations, FGGn-Skp1 is still calculated to enjoy a 15-fold catalytic advantage over FGGn-pNP. Evidence that Skp1 is a primary target of Glt1 comes from a biochemical complementation experiment, in which Skp1 is the only acceptor substrate detected after incubation of glt1Δ extracts with His6Glt1 and radioactive UDP-Glc (Fig. 8H). Although other substrates could have been missed because they were excluded from or not accessible in the extract, or alternatively modified in the absence of Glt1, the interpretation that Glt1 is dedicated to Skp1 is consistent with evidence that the two earlier glycosyltransferases in the pathway are specific for Skp1 (3), and with more extensive evidence that the Skp1 modification pathway is specific for Skp1 in Dictyostelium (14, 4).

Functional Variations Between Toxoplasma Glt1 and Dictyostelium AgtA—Skp1 from Dictyostelium also possesses a pentasaccharide on the corresponding 4-hydroxyproline residue, and evidence suggests that the core trisaccharide of Toxoplasma Skp1 is identical to that of the recently confirmed (9) core trisaccharide of Dictyostelium Skp1, Fucα1,2Galβ1,3GlcNAcα1-. The evidence for identity is based primarily on sequence homologies of their Gnt1 and PgtA enzymes (3), and confirmation of their donor substrates as UDP-
GlcNAc, and UDP-Gal and GDP-Fuc, respectively. Identity of the trisaccharide cores is also supported by the robust activity of His₆Glt1 toward endogenous Toxoplasma Skp1 in glt1Δ extracts (Fig. 8H). More direct confirmation of the linkages of the core-trisaccharide is confounded by the small amounts of the intracellular pathogen that are available, and difficulties encountered in expressing soluble protein from Toxoplasma pgta cDNA in E. coli or Dictyostelium.¹¹

The fourth sugar in the Dictyostelium Skp1 glycan, an αGlc, differs from the αGlc in Toxoplasma, but conserves the 3-linkage to Fuc. The enzyme catalyzing its addition, AgtA, belongs to CAZy family GT77 and is evolutionarily unrelated to Glt1. AgtA also catalyzes addition of the fifth sugar, an αGal that is 3-linked to the fourth sugar to form a linear chain. In contrast, in Toxoplasma the fifth sugar is added by a separate enzyme.¹² The mechanism of recognition of FGGn-Skp1 is distinct for the two enzymes. Whereas AgtA strongly prefers FG- relative to Fuc- or FGGn- (as conjugates to small aglycons) as an acceptor (22), Glt1 strongly favors the full trisaccharide and is unable to modify Fuc alone (Figs. 8D, E). Genetic and biochemical studies reveal that Dictyostelium AgtA has an independent function in modulating Skp1 activity that likely involves a physical interaction via a separate WD40-repeat domain (13). Unlike AgtA, Glt1 lacks an identifiable domain separate from the catalytic domain, though it does possess short sequences at various positions within the conserved catalytic domain (Fig. S2). However, these insert sequences, variations of which are commonly observed in Toxoplasma proteins (23), are not conserved at the sequence level in another apicomplexan expected to harbor the Skp1 modification pathway, Sarcocystis neurona. Furthermore, the insert sequences tend to be absent from other protists that possess glt1-like genes. Therefore, Glt1 lacks a structural basis for the second function possessed by AgtA.

Structural studies of Dictyostelium Skp1 suggest that the full-length glycan encourages an ensemble of conformations that promote interactions with at least certain FBPs (9). The trisaccharide form of Dictyostelium Skp1 exhibits intermediate interaction with two of the FBPs (6) and, if an analogous mechanism operates in Toxoplasma, this might explain why the glt1 deletion results in a growth phenotype intermediate between complete absence of the glycan and its full assembly. Further studies, to be conducted when the final linkage of the Toxoplasma glycan is ascertained, will address this possible explanation.

Glt1 is Ancestral for Skp1 Glycosylation in Protists—A comprehensive reconstruction of the evolution of CAZy GT32 related sequences indicates that Glt1-like sequences form a distinct clade within the GT32 group and have evolutionarily diverged from other members through variations in the GT domain. This interpretation is supported by conservation of amino acids at ten positions throughout the protein, including three that are exclusive to the Glt1 clade (Fig. S2). Glt1-like sequences are found only in protists (Fig. 2), with one exception in a bacterium which might reflect horizontal transfer. But the broad distribution within this diverse phylogeny suggests that Glt1 was present in the last common protistan ancestor and selectively lost or modified where it did not afford a selective advantage in, e.g., O₂-sensing. Furthermore, the Glt1 clade is populated only by protists that also possess CAZy GT74-like sequences predicted to encode the Skp1 PgtA-like enzyme that assembles the Fuc residue on which Glt1 acts in Toxoplasma. This evidence of co-evolution implicates all members of this clade in the assembly of the Skp1 glycan. However, Glt1-like sequences are absent from the amoebozoa (using BlastP search at E<10⁻¹⁰), which include Dictyostelium, where the CAZy family GT77 member AgtA performs a related biochemical function. The simplest explanation is that Glt1 is the ancestral Skp1 glycosyltransferase whose function was replaced in amoebozoa, which might have occurred to compensate for a loss of Glt1 or the final glycosyltransferase (Gat1),¹² or because of the selective advantage afforded by AgtA’s additional function in Skp1 suppression (13).

Experimental Procedures

Parasite Manipulations—Strain RHΔku80Δhxgprt (RHΔ) of Toxoplasma gondii was cultured on preformed monolayers of human foreskin fibroblasts (ATCC #SCRC-1041), or BJ fibroblasts (ATCC #CRL-2522) if indicated, in the presence of Complete Medium, which consisted of Dulbecco’s Modified Eagle’s Medium (Corning Inc.) supplemented with 10% (v/v) fetal bovine serum, 2 mM L-glutamine and 100 units/ml...
penicillin/streptomycin (Corning Inc.) at 37°C in a humidified CO₂ (5%) incubator. RHΔku80Δglt1 (glt1Δ) strain was selected and maintained in the same medium with 25 µg/ml mycophenolic acid (Sigma) and 25 µg/ml xanthine (Sigma). The medium of complemented glt1Δ strains was supplemented with 250 µg/ml of 6-thioxanthine (Matrix Scientific). Strains were cloned by limiting dilution on BJ fibroblasts in 96-well plates. Plaque assays were conducted in media without selection drug (3). Cytosolic S100 extracts were prepared from tachyzoites isolated from infected fibroblasts and desalted, and soluble S16 urea extracts were prepared, as before (3).

**glt1 Gene Replacement and Complementation—DNA**
For the gene replacement was generated from pminiGFP.hht, in which the hgxprt gene is flanked by multiple cloning sites, as described (3). Briefly, 5’-flanking and 3’-flanking targeting sequences of glt1 from RHΔΔ were PCR amplified with primer pairs a and a’ and pairs b and b’, respectively (Table S1). The 5’-fragment was released by digestion with ApaI and HindIII and cloned into similarly digested pminiGFP.hht. The 3’-fragment was similarly inserted using Xbal and NotI. The resulting vector was linearized with ApaI and electroporated into strain RHΔΔ. Transformants were selected and GFP negative parasites were cloned by limiting dilution. Genomic DNA was prepared and screened by PCR (3).

To complement a glt1Δ clone, the hgxprt cassette of pminiGFP.hht was replaced with a ~5 kb gDNA fragment containing the glt1 coding region from RHΔΔ (2860 nt) plus ~1 kb each of 5’-flanking and 3’-flanking DNA, using the Complementary Annealing Mediated by Exonuclease cloning method (25). Briefly, the vector and insert were PCR amplified separately for 20 cycles by Q5 high fidelity DNA polymerase (Novagen) at 22°C for 2.5 min to inactivate the polymerase, and annealed at 75°C for 20 min in 10 mM dithiothreitol in 50 mM NH₄HCO₃, incubated at 56°C for 1 h, alkylated with 22.5 mM iodoacetamide for 45 min in dark, and digested with trypsin (Promega) at 37°C overnight. The resulting peptides were recovered by addition to a C18 spin column (MicroSpin™ Column, The Nest Group), elution with 0.1% formic acid in 80% acetonitrile, and drying under vacuum. Peptides were reconstituted in 19.5 µl of solvent A (0.1% formic acid) and 0.5 µL of solvent B (0.1% formic acid in 80% acetonitrile), separated on an Acclaim PepMap RSLC C18 column (75 µm × 15 cm), and eluted into the ion source of an Orbitrap Fusion Lumos Trifrib™ mass spectrometer (Thermo Fisher) with a linear gradient consisting of 0.5-100% solvent B over 150 min at a flow rate of 200 nL/min. The spray voltage was set to 2.2 kV and the temperature of the heated capillary was set to 280°C. Full MS scans were acquired from m/z 300 to 2000 at 120k resolution, and MS² scans following collision-induced fragmentation were collected in the ion trap for the most intense ions in the Top-Speed mode within a 3-sec cycle using Fusion instrument software (v2.0, Thermo Fisher Scientific). The acquired raw spectra were analyzed using SEQUEST (Proteome Discoverer 1.4, Thermo Fisher Scientific) with a full MS peptide tolerance of 20 ppm and MS² peptide fragment tolerance of 0.5 Da, and filtered to generate a 1% target decoy peptide-spectrum-match (PSM) false discovery rate for protein assignments. Spectra assigned as glycosylated TgSkp1 peptides were manually validated.

**Preparation of Toxoplasma FGGn-Skp1**
The E. coli dual expression plasmid for Dictyostelium Skp1A (5) and Dictyostelium PhyA, was modified by substitution with synthetic cDNA encoding Toxoplasma Skp1, and introduced into E. coli together with the plasmid encoding Dictyostelium DdDp-Gnt1. Gn-Skp1 was purified as for its Dictyostelium counterpart, and modified in vitro
using *Dictyostelium* FLAG-PgtA. The reaction consisted of 6.25 µM *Toxoplasma* Gn-Skp1, 112 nM *Dictyostelium* FLAG-PgtA, 25 µM UDP-Gal, 50 µM GDP-Fuc, 120 mM NaCl, in 50 mM Tris-HCl (pH 7.5). The reaction was monitored by dot blotting onto nitrocellulose filters and probing with mAb 1C9, which is specific for *Dictyostelium* Gn-Skp1 relative to other modified glycoforms (15), and 1:1000 dilution of pAb UOK104, which is specific for *Dictyostelium* FGGn-Skp1 relative to other glycoforms (26). Alexa680 coupled secondary antibodies were applied and detected in an Odyssey infrared scanner (LiCor). Similar specificity was observed for the *Dictyostelium* and *Toxoplasma* glycoforms, and the glycosylation reaction was taken to completion within the sensitivity of the method (data not shown).

**Expression and Purification of Recombinant His<sub>6</sub>Glt1**- The predicted coding sequence of Glt1 (TGGT1_205060) was codon optimized for expression in *E. coli*, chemically synthesized by GenScript (Piscataway, NJ), and inserted into the pUC57 vector between its NdeI and BamHI sites (Supplemental Fig. S1). After treatment with NdeI and BamHI, the released coding fragment was inserted into similarly digested pET15b (Invitrogen), which resulted in the full-length 605 amino acid coding sequence preceded by an N-terminal His<sub>6</sub>tag and TEV protease cleavage site (MGSSHHHHHHSSGRENLYFQGH-). *E. coli* Gold cells expressing His<sub>6</sub>Glt1 were incubated for 24 hours in 8 × 1 l Terrific Broth medium in the presence of 100 µg/ml of ampicillin, 2 g/l lactose, and 125 µM isopropyl-β-D-galactopyranoside at 20 °C. Pilot studies showed that the auto-induction during bacterial growth (27) was much superior to standard induction at high density for expression of soluble Glt1 (data not shown). After 24 h, cells were collected by centrifugation at 2,000 × g for 10 min and resuspended in 50 mM Na<sup>+</sup>/K<sup>+</sup> phosphate (pH 7.8), 300 mM NaCl, 2 mM benzamidine, 0.5 µg/ml pepstatin A, 5 µg/ml aprotinin, 5 µg/ml leupeptin, and 0.5 mM phenylmethylsulfonyl fluoride at 4 °C; Na<sup>+</sup>/K<sup>+</sup> phosphate buffer was prepared by titrating monosodium phosphate into dipotassium phosphate of equal molarity. Bacteria were lysed using a probe sonicator (model #500, Fisher Scientific). The lysate was immediately centrifuged at 21,000 × g for 30 min at 4 °C, and the supernatant (S21) was applied to a column containing 1.5 ml Co<sup>2+</sup> TALON resin (Clontech Laboratories) pre-equilibrated at 4 °C in the buffer described above. Protein was eluted with 300 mM imidazole in the same buffer. The major A<sub>280</sub> peak was dialyzed against 40 mM Tris-HCl (pH 8.0) overnight, and applied to a 1-ml Hi-Trap Q-Sepharose column (GE-Healthcare) pre-equilibrated at 4 °C in 40 mM Tris-HCl (pH 8.0), 2 mM DTT, 2 mM MgCl<sub>2</sub>, 10% (v/v) glycerol, aprotinin and leupeptin (as above). The column was eluted with a 0-1 M gradient of NaCl in the same buffer. Fractions with highest enzymatic activity were confirmed for the presence of His<sub>6</sub>Glt1 by Western blotting with anti-His<sub>6</sub> monoclonal antibody (Novagen #70796-3), pooled, and frozen as aliquots at -80°C.

**Western blot analysis**— SDS-PAGE was performed on 4-12% preformed gels (NuPAGE Novex, Invitrogen), Western blotted, probed with primary and fluorescent secondary antibodies, and scanned as described previously (3).

**Glt1 Enzyme Activity Assays**— Hydrolysis of UDP-sugars was conducted using the UDP-Glo assay (Promega) as described (18). Briefly, His<sub>6</sub>Glt1 (after Q-column purification) was incubated in the presence of 50 µM sugar nucleotides in 20 µl reactions containing 50 mM HEPES-NaOH (pH 7.4), 2 mM MnCl<sub>2</sub>, 5 mM DTT at 37°C for 16 h, and activity was quantitated based on conversion of the UDP reaction product to ATP.

Glt1 glycosyltransferase activity was assayed as the transfer of <sup>3</sup>H from UDP-[<sup>6</sup>-<sup>3</sup>H]Glc to synthetic acceptors or FGGn-Skp1. Synthetic acceptors were Fucα1-para-nitrophenol (F-pNP) (Sigma), Fucα1,2Galβ1-Bn (FG-Bn), Fucα1,2Galβ1,3GlcNAcβ1-pNP (FGGn-pNP), Galβ1,3GlcNAcβ1-Bn, GlcNAcβ1-Bn, and Galβ1,3Fucα1-Bn (22, 28). The standard reaction consisted of Q-column purified His<sub>6</sub>-Glt1, and 2 mM synthetic acceptor or 1 µM FGGn-TgSkp1, 50 mM HEPES-NaOH (pH 7.4), 2 µM UDP-Glc (Sigma-Aldrich), including 0.5 µCi UDP-[<sup>6</sup>-<sup>3</sup>H]Glc (38.2 µCi/nmol; Perkin Elmer Sciences), 2 mM MnCl<sub>2</sub>, 50 mM NaCl, 5 mM DTT, and 2 mg/ml bovine serum albumin (Sigma), in a final reaction volume of 20 µl that was incubated for 1 h at 37 °C. Concentrations, pH (using a mixture of MES and HEPES), and reaction times were varied as indicated, and confirmed to be linearly dependent on enzyme concentration and reaction time. Reactions using synthetic acceptor substrates were terminated by addition of 1 ml of ice-cold 2 mM sodium EDTA (pH 8.0), and applied to and eluted from a Sep-Pak C<sub>18</sub> cartridge (360 mg, Waters Corp.), and counted for radioactivity as
After incubation for 3 h at 37°C, His6-Glt1, in a final volume of 200 µl containing 863 µl 2 mM Na-EDTA (pH 8.0), 20 µl 20 mg/ml BSA, and 85% TCA to a final concentration of 10%, and incubation on ice. The precipitate was recovered by filtration on a 0.45 µm glass filter, washed with ice-cold 10% TCA and ice-cold acetone, and transferred into Biosafe NA (RPI) scintillation cocktail and counted as above.

Alternatively, Glt1 transferase activity was assayed based on release of GDP or UDP from unlabeled GDP-Man or UDP-Glc, using GDP-Glo or UDP-Glo assays (Promega) according to the manufacturer’s protocol. Conditions were as described above.

Kinetic parameters were determined assuming Michaelis-Menten kinetics using GraphPad Prism software. Acceptor substrate kinetics were determined at 40 µM UDP-Glc, and donor substrate kinetics were determined at 2 mM FG-Bn.

For assay of transferase activity in parasite extracts, the typical 50 µl reaction volume containing 30 µl S100 fraction (1 mg protein/ml), 50 pmol of FGGN-DdSkp1 (5), 1 µCi UDP-[3H]Glc or UDP-[15N]Gal (13), in 50 mM HEPES-NaOH (pH 7.4), 10 mM MgCl2, 2 mM MnCl2, 2 mM DTT, 3 mM NaF, and protease inhibitors, was incubated at 37°C for 1.5 or 3 h. Reactions were stopped by addition of 4×-Laemmli electrophoresis sample buffer and incorporation was after SDS-PAGE as described (3). The chemical form of incorporated radioactivity was determined after electroblot transfer to a polyvinylidene difluoride membrane, acid hydrolysis, and high pH anionic exchange chromatography, using 1.5 nmol each of L-Fuc, D-Gal and D-Man as internal standards (3).

For assay of acceptor substrate activity in extracts, reactions were modified to contain 2 µCi UDP-3H-Glc, 170 µl desalted S100 fraction (1 mg/ml), His6-Glt1, in a final volume of 200 µl. After incubation for 3 h at 37°C, the reaction was concentrated by centrifugal ultrafiltration using a Nanosep 3K concentrator (Pall Corporation). The samples were resolved by SDS-PAGE and each lane was sliced into 36 equal pieces which were counted for radioactivity as above.

Glc-Fuc linkage analysis— To determine the linkage of the Glt1 product, 4 µmol of FGGN-pNP and 8 µmols UDP-Glc were incubated with 34 µg (~0.65 nmol) His6-Glt1 and 40 µg calf intestinal alkaline phosphatase (Promega) in 2 ml reaction buffer (50 mM HEPES-NaOH (pH 8.0), 5 mM DTT, 8 mM MnCl2 and 10 mM MgCl2), at 37°C for 5 h. The reaction was terminated with 8 ml 5 mM Na-EDTA, pH 8.0. The product was recovered using a Sep-Pak C18 cartridge as above, dried under N2, and reconstituted in 50 µl H2O. Quantitative conversion to the tetrasaccharide was confirmed using MALDI-TOF mass spectrometry (ABSciEx 5800, Applied Biosystems), and the concentration was determined spectrophotometrically using an extinction coefficient of 1.15 × 10^4 M⁻¹ cm⁻¹ at 300 nm (22). 3 µmoles of dried GFGGn-Skp1 was resuspended in 99.96% D2O and analyzed on an Agilent 900 MHz DD2 spectrometer equipped with a 5 mm cryogenically cooled probe. NMR experiments were performed at 25°C after stabilization and shimming, and used standard pulse sequences (PRESAT, gCOSY, zTOCSY, HSQCAD, gHMBCAD) from the Agilent library. Two-dimensional data were collected with default values except for increased digital resolution in both dimensions. Spectra were processed with MestReNova software (Mestrelab Research S.L.). Peaks of the 1H–13C-HSQC spectra were provisionally assigned by comparison with predicted chemical shifts calculated by the CASPER program and all signals and residue linkages confirmed from analysis of the 2-D data.

Phylogenetic Analysis— Full length sequences for all the experimentally characterized GT32 family proteins listed in the CAZy database (accessed May 10, 2017, http://www.cazy.org/) (29) were collected. To identify Glt1 related sequences, Toxoplasma gondii (EPR61400.1) and Vitrella brassicaformis (CEM02366.1) Glt1 protein sequences were used as BLAST queries to search against the NCBI nr database and available protist proteome databases [Ensembl protist, JGI (http://protistsensembl.org/index.html, http://genome.jgi.doe.gov/)]. The putative ortholog from the chromerid Vitrella brassicaformis, a close relative of apicomplexa that is predicted to express the Toxoplasma-like Skp1 modification pathway, was included owing to its lack of apicomplexan-specific insertions within its proteins (30). For the GT74 family, the only characterized sequence from Dictyostelium discoideum (AAF82378.1) was used as BLAST query. For both families (GT32 and GT74), the collected BLAST best hit sequences
were then aligned using MAFFT L-INS-I strategy (31) and the alignment was used to generate a Hidden Markov Model (HMM) (32). The built HMM profile was used to identify additional related sequences in NCBI nr from diverse taxonomic groups using an e-value cutoff of 1e-5. A total of 68 GT32 related sequences from bacteria, protists, fungi, chlorophyte and metazoans were collected and used for the phylogenetic tree construction. Hits collected for the GT74 family were used to identify the presence of a GT74 related sequence in the species.

The 68 full length sequences representing the GT32 family members were aligned using MAFFT L-INS-I strategy. The boundaries of the GT32 domain was marked by identifying conserved regions in the alignment. The sequences were initially trimmed to extract only the GT32 domain and re-aligned using the same strategy. This domain alignment was further refined by removing large insert segments, and poorly aligned regions. The trimmed alignment was used to generate the final phylogenetic tree. FastTree (33) was used to build the tree with default parameters using the following options: -wag for the WAG (34) model of amino acid evolution, -gamma for the rescaling of the branch lengths and the computation of Gamma20 based likelihood scores, -pseudo to add pseudo counts for highly gapped segments in the alignment. Local support values for the internal nodes were computed by FastTree using the Shimodaira-Hasegawa test and are displayed in the figure. Trees generated using full length GT domain alignments also resulted in similar tree topologies. Runs using RaXML also generated similar topologies (not shown). Functional annotations were collected from the CAZy database for the characterized sequences. SignalP3.0 (www.cbs.dtu.dk/services/SignalP-3.0/) was used to predict signal peptide sequences, using separate runs for eukaryotes, gram negative and gram-positive bacteria, and TMHMM 2.0 (www.cbs.dtu.dk/services/TMHMM/) was used to predict the transmembrane regions.

Acknowledgments— We thank Dr. Hicham Zegzouti (Promega R&D) for providing the GDP-Glo assay kit, and are grateful to Ira Blader (Buffalo) for reviewing the manuscript.

Conflict of Interest— The authors declare that they have no conflicts of interest with the contents of this article.

Author Contributions— K. R. conducted most of the experiments, analyzed the results, assembled figures, and wrote most of the first draft. P. Z. conducted the mass spectrometry and interpreted the data together with L. W. M. M. and H. vdW. contributed to the enzymatic assays and HPAEC analyses. M. O. S. performed the UDP-sugar hydrolysis assays, and M. O. S. and J. N. G. performed the linkage analyses. H. W. K. prepared the Glt1 enzyme. K. M. synthesized the glycan substrates. R. T. conducted the phylogenetic analyses which were edited by N. K. and C. M. W. C. M. W. conceived most ideas and experimental strategies for the project, and compiled the final draft, which was approved by all authors.

References
1. Willems, A. R., Schwab, M., and Tyers, M. (2004) A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. Biochim. Biophys. Acta 1695, 133–170
2. Skaar, J. R., Pagan, J. K., and Pagano, M. (2013) Mechanisms and function of substrate recruitment by F-box proteins. Nat. Rev. Molec. Cell Biol. 14, 369–381
3. Rahman, K., Zhao, P., Mandalasi, M., van der Wel, H., Wells, L., Blader, I. J., and West, C. M. (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, 4268–4280
4. West, C. M., and Blader, I. J. (2015) Oxygen sensing by protozoans: How they catch their breath. Curr. Opin. Microbiol. 26, 41–47
5. Sheikh, M. O., Schafer, C. M., Powell, J. T., Rodgers, K. K., Mooers, B. H., and West, C. M. (2014) Glycosylation of Skp1 affects its conformation and promotes binding to a model F-Box protein. Biochemistry 53, 1657–1669
6. Sheikh, M. O., Xu, Y., Van der Wel, H., Walden, P., Hartson, S. D., and West, C. M. (2015) Glycosylation of Skp1 promotes formation of SCF complexes in Dictyostelium. Mol. Cell Proteomics
14, 66–80
7. Schulman, B. A., Carrano, A. C., Jeffrey, P. D., Bowen, Z., Kinnucan, E. R. E., Finnin, M. S., Elledge, S. J., Harper, J. W., Pagano, M., and Pavietich, N. P. (2000) Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408**, 381–386

8. Zimmerman, E. S., Schulman, B. A., and Zheng, N. (2010) Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* **20**, 714–721

9. Sheikh, M. O., Thieker, D., Chalmers, G., Schafer, C. M., Ishihara, M., Xu, X., Azadi, P., Woods, R. J., Glushka, J. N., Bendiaik, B., Prestegard, J. H., and West, C. M. (2017) Glycosylation associated with O2-sensing modulates the F-box combining site of *Dictyostelium* Skp1. *J. Biol. Chem.*, in review

10. Xu, Y., Wang, Z.A., Green, R.S., West, C.M. (2012) Role of the Skp1 prolyl-hydroxylation/glycosylation pathway in oxygen dependent submerged development of *Dictyostelium*. *BMC Dev Biol.* Oct 25;12:31

11. Bonner, J. T., and Lamont, D.S. (2005) Behavior of cellular slime molds in the soil. *Mycologia* **97**, 178-184

12. Xu, Y., Brown, K. M., Wang, Z. A., van der Wel, H., Teygong, C., Zhang, D., Blader, I. J., and West, C. M. (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**, 25098–25110

13. Schafer, C. M., Sheikh, M. O., Zhang, D., and West, C. M. (2014) Novel regulation of Skp1 by the *Dictyostelium* AgtA alpha-galactosyltransferase involves the Skp1-binding activity of its WD40 repeat domain. *J. Biol. Chem.* **289**, 9076–9088

14. West, C. M., Wang, Z. A., and van der Wel, H. (2010) A cytoplasmic prolyl hydroxylation and glycosylation pathway modifies Skp1 and regulates O2-dependent development in *Dictyostelium*. *Biochim. Biophys. Acta* **1800**, 160–171

15. Wang, Z. A., van der Wel, H., Vohra, Y., Buskas, T., Boons, G. J., and West, C. M. (2009) Role of a cytoplasmic dual-function glycosyltransferase in O2 regulation of development in *Dictyostelium*. *J. Biol. Chem.* **284**, 28896-28904

16. Blader, I. J., and Koshy, A. A. (2014) *Toxoplasma gondii* development of its replicative niche: in its host cell and beyond. *Eukaryot. Cell* **13**, 965-976

17. Wiggins, C. A., and Munro, S. (1998) Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7945-7950

18. Sheikh, M. O., Halmo, S. M., Patel, S., Middleton, D., Takeuchi, H., Schafer, C. M., West, C. M., Haltiwanger, R. S., Avcı, F. Y., Moremen, K. W., and Wells, L. (2017) Rapid screening of sugar-nucleotide donor specificities of putative glycosyltransferases. *Glycobiology* **27**, 206-212

19. Rönnols, J., Pendrill, R., Fontana, C., Hamark, C., d’Ortoli, T. A., Engström, O., Stähle, J., Zaccheus, M. V., Säwén, E., Hahn, L. E., Iqbal, S., and Widmalm, G. (2013) Complete 1H and 13C NMR chemical shift assignments of mono- to tetrasaccharides as basis for NMR chemical shift predictions of oligosaccharides using the computer program CASPER. *Carbohydr. Res.* **380**, 156–166

20. Varki, A., Cummings, R. D., Aeberi, M., Packer, N. H., Seeberger, P. H., Esko, J. D., Stanley, P., Hart, G., Darvill, A., Kinoshita, T., Prestegard, J. H., Schnaar, R. L., Freeze, H. H., Marth, J. D., Bertozzi, C. R., Etzler, M. E., Frank, M., Vliegenthart, J. F., Lutteke, T., Perez, S., Bolton, E., Rudd, P., Paulson, J., Kanehsia, M., Toukach, P., Aoki-Kinoshita, K. F., Dell, A., Narimatsu, H., York, W., Taniguchi, N., and Kornfeld, S. (2015) Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323-1324

21. Bandini, G., Haserick, J. R., Motari, E., Ouologuem, D. T., Lourido, S., Roos, D. S., Costello, C. E., Robbins, P. W., and Samuelson, J. (2016) O-fucosylated glycoproteins form assemblies in close proximity to the nuclear pore complexes of *Toxoplasma gondii*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11567-11572

22. Ketcham, C., Wang, F., Fisher, S. Z., Erkan, A., van der Wel, H., Locke, R. D., Sirajud-Doula, K., Matta, K. L., and West, C. M. (2004) Specificity of a soluble UDP-galactose:fucoside alpha1,3-galactosyltransferase that modifies the cytoplasmic glycoprotein Skp1 in *Dictyostelium*. *J. Biol. Chem.* **279**, 29050–29059
23. Narasimhan, J., Joyce, B. R., Naguleswaran, A., Smith, A. T., Livingston, M. R., Dixon, S. E., Coppens, I., Wek, R. C., and Sullivan, W. J., Jr. (2008) Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in *Toxoplasma gondii*. *J. Biol. Chem.* **283**, 16591–16601

24. Fox, B. A., Ristuccia, J. G., Gigley, J. P., and Bzik, D. J. (2009) Efficient gene replacements in *Toxoplasma gondii* strains deficient for nonhomologous end joining. *Eukaryot. Cell* **8**, 520-529

25. Sun, S., Huang, H., Qi, Y. B., Qiu, M., and Dai, Z. M. (2015) Complementary annealing mediated by exonuclease: a method for seamless cloning and conditioning site-directed mutagenesis. *Biotechnol. Biotechnol. Equip.* **29**, 105-110

26. Chinoy, Z. S., Schafer, C. M, West, C. M, and Boons, G. J. (2015) Chemical synthesis of a glycopeptide derived from Skp1 for probing protein specific glycosylation. *Chemistry* **21**, 11779-11787

27. Li, Z., Kessler, W., Heuvel, J. V., and Rinas, U. (2011). Simple defined autoinduction medium for high-level recombinant protein production using T7-based Escherichia coli expression systems. *Appl. Microbiol. Biotechnol.* **91**, 1203-1213

28. Rana, S. S., Barlow, J. J., and Matta, K. L. (1981) The chemical synthesis of O-alpha-L-fucopyranosyl-(1 goes to 2)-O-beta-D-galactopyranosyl-(1 goes to 3)-O-[alpha-L-fucopyranosyl-(1 goes to 4)]-2-acetamido-2-deoxy-D-glucopyranose, the Lewis blood-group antigenic determinant. *Carbohydr. Res.* **96**, 231-239

29. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nuc. Ac. Res.* **42** (Database issue), D490-495

30. Fussy, Z., Masarova, P., Krucinska, J., Esson, H. J., and Obornik, M. (2016) Budding of the alveolate alga *Vitrella brassicaformis* resembles sexual and asexual processes in apicomplexan parasites. *Protist* **168**, 80-91

31. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nuc. Ac. Res.* **30**, 3059-3066

32. HMMER: biosequence analysis using profile hidden Markov models [http://hmmer.org/].

33. Price, M. N., Dehal P. S., and Arkin, A. P. (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS One* **5**, e9490

34. Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691-699

* This work was supported in part by Grant 14-140 from the Mizutani Foundation for Glycoscience (to C. M. W.), and National Institutes of Health Grants R01-GM084383 (to C. M. W.), P41-GM103490 (to L. W., senior investigator), and 8P41-GM103390 (Resource for Integrated Glycotechnology). R. T. was supported by NIH training grant T32-GM107004. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains supplemental Table S1 and Figs. S1 and S2.

7 To whom correspondence should be addressed, at the Dept. of Biochemistry & Molecular Biology, University of Georgia, 120 East Green Street, Athens, GA 30602 USA; Tel.: 706-542-4259; E-mail: westcm@uga.edu

8 Current address: HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, Maryland, 21702 USA

9 Abbreviations used: FBP, F-box protein; GGFGn-Skp1, fully glycosylated Skp1 in which G=Gal, F=Fuc, Gn=GlcNAc; HMBC, Heteronuclear Multiple Bond Correlation; HSQC, Heteronuclear Single Quantum Correlation; mAb, monoclonal antibody; pAb, polyclonal antibody; SCF, Skp1/Cullin-1/F-box protein; Ub, ubiquitin;

10 Sheikh, M. O., Nottingham, B., Boland, A., and West, C. M., unpublished data

11 van der Wel, H., and West, C. M., unpublished data

12 Mandalasi, M., Rahman, K., Kim, H., unpublished data
**TABLE 1. Toxoplasma strains used in this study.**

| Strain      | Parental strain | Genotype      | Gene targeted | Selection marker | Selection drug | Ref. |
|-------------|-----------------|---------------|---------------|------------------|----------------|------|
| RHΔΔ        | RH(1)           | Δhxgprt;Δku80 | –             | –                | –              | 24   |
| phyAΔ       | RHΔΔ            | phyAΔ;Δku80   | phyA-exon 1   | Hxgprt           | MPA, xanthine  | 12   |
| glt1Δ       | RHΔΔ            | glt1Δ;Δku80   | glt1-exon 1-5 (all) | Hxgprt | MPA, xanthine | TR  |
| glt1Δglt1Δ  | glt1Δ-clone 1   | Δhxgprt;Δku80 | –             | ΔHxgprt         | 6-thioxanthine | TR  |
| glt1Δ;glt1Δ | glt1Δ-clone 1   | Δhxgprt;Δku80; glt1(D363A/D365A/D348A) | –             | ΔHxgprt         | 6-thioxanthine | TR  |

\[ \text{a TR: this report}\]

\[ \text{b (m): signifies the triple mutant form encoding Glt1(D363AD365A/D348N)}\]

**Table 2. MS detection of Skp1 glycopeptides in strains.** Isoforms of the Skp1 peptide \[^{14S}\text{IFNIVNDFTPPEEAQVR} \] were detected as described in Fig. 4. The distribution of raw ion counts among the detected isoforms are shown for the strains analyzed.

| Strain      | unmodified peptide | Hex-HexNAc-peptide | dHex-Hex-HexNAc-peptide | Hex-Hex-dHex-HexHexNAc-peptide |
|-------------|-------------------|--------------------|-------------------------|--------------------------|
|             | fraction\[^{a}\]  | [M+2H]\(^{\text{²}}\) ΔM\[^{b}\] | [M+3H]\(^{\text{³}}\) ΔM | [M+2H]\(^{\text{²}}\) ΔM |
| RHΔΔ        | 0.68              | 1011.0002          | 0.218                   | n.d.\[^{c}\]             | 0.31            | 1436.6464 |
|             |                   | n.r.\[^{f}\]      |                         |                          |                 | n.r.      |
| glt1Δ       | 0.45              | 1011.0042          | 1.78                    | 2.67                     | 0.03            | 1201.5630 | 2.47 |
|             |                   | 674.3391           |                         |                          |                 | 801.3804 | 0.90 |
| glt1Δ;glt1Δ | 0.27              | 1011.0042          | 1.78                    | 2.67                     | 0.01            | 1201.5657 | 0.22 |
|             |                   | 674.3380           |                         |                          |                 | 801.3818 | 2.65 |
| glt1Δ;glt1Δ | 0.36              | 1011.0020          | 0.40                    | 0.44                     | 0.05            | 1201.5651 | 0.72 |
|             |                   | 674.3370           |                         |                          |                 | 801.3817 | 2.52 |

\[ \text{\[a\] apparent fractional occupancy, averaged from doubly- and triply-charged ion raw spectral counts}\]

\[ \text{\[b\] Δm/z in ppm}\]

\[ \text{\[c\] from ref. 3}\]

\[ \text{\[d\] not reported}\]

\[ \text{\[e\] not detected}\]

\[ \text{\[f\] (m): signifies the triple mutant form encoding Glt1(D363AD365A/D348N)}\]
FIGURE LEGENDS

FIGURE 1. Bioinformatics strategy used to identify potential TgSkp1 modifying glycosyltransferases. A, steps utilized in bioinformatic analysis to identify potential terminal glycosyltransferase candidates. B, the Skp1-glycosylation pathway in Toxoplasma. Arrows from left-to-right show at which step of the pathway putative glycosyltransferases are predicted to function.

FIGURE 2. Phylogenetic distribution of Glt1-like sequences. Phylogenetic tree showing 3 distinct clades for the CAZy GT32 family sequences. Different colored branches and leaves indicate different taxonomic groups as labelled. Local support values for the branch splits are shown. Known enzymatic activity for experimentally characterized sequences are provided in black after the leaf name as per the CAZy database. Teal colored circles indicate the presence of a GT74 fucosyltransferase related protein in the species; empty circles indicate its absence. Orange coloration indicates evidence for a signal peptide or signal anchor sequence expected to direct the protein to the secretory pathway.

FIGURE 3. Disruption and complementation of glt1. A, outline of strategy for double crossover gene replacements, and PCR reactions to monitor outcomes. B, PCR analysis of hxgprt-positive clones that were resistant to mycophenolic acid and xanthine, and GFP-negative. Desired replacement clones were characterized by absence of products of PCR1, and presence of PCR2 and PCR3. C, PCR analysis of gnt1-replacement clone-1 with gDNA encoding wild-type Glt1 or a triple mutant form encoding Glt1(D363AD365A/D348N).

FIGURE 4. glt1Δ parasites accumulate the trisaccharide form of Skp1. A, Western blot analysis of Skp1 from equivalent numbers (3 × 10^6 cells) of parental RHΔΔ, phyAΔ, glt1Δ or complemented cells. Soluble S16 fractions were separated on a 4-12% SDS-PAGE, electroblotted, and probed with anti-TgSkp1 (UOK75) antiserum. B-E, analysis of the Skp1 glycosylation by mass spectrometry, from RHΔΔ (left half) and glt1Δ (right half) cells. Skp1 was enriched by immunoprecipitation and converted to peptides with trypsin. Total peptides were analyzed by nanoLC/MS-MS (*note*: retention times varied between the left and right-hand trials, which were conducted at different times). The vertical dashed boxes show approximate elution intervals for glycoforms of the Pro154-containing peptides from Skp1. Spectral counts corresponding to exact m/z matches that appear as both doubly and triply charged forms are annotated at the side. B, extracted ion scans from full ion mode for the unmodified peptide containing Pro154, showing scans for [M+2H]^2+ and [M+3H]^3+ ions. C, extracted ion scans for corresponding to the HexHexNAc-derivative of the Pro-hydroxylated peptide. D, same for the dHexHexHexNAc- derivative. E, same for the fully glycosylated Hex2dHexHexHexNAc- derivative. No other glycoforms were detected.

FIGURE 5. MS-MS confirmation of the trisaccharide-glycopeptide. A, the doubly charged ion from Fig. 4, whose m/z (1274.5954) matched that expected of IFNIVFTP(HexNAc-dHex-Hex)EEEAQVR. B, CID fragmentation of the doubly-charged precursor ion yields a sequential loss of monosaccharide residues corresponding to dHex, Hex and HexNAc, indicating the presence of a linear trisaccharide. C, the full CID fragmentation spectrum showing b- (blue annotations) and y- (red annotations) ion series that match the predicted peptide sequence, as illustrated in the inset. The glycan is linked via a hydroxylated derivative of Pro154. Peptides with residual sugars are annotated in green.

FIGURE 6. Toxoplasma extracts express gtl1-Dependent Skp1 glucosyltransferase activity. A, desalted cytosolic extracts were assayed for ability to transfer ^3H from UDP-[^3H]Gal or UDP-[^3H]Glc to FGGn-Skp1, based on incorporation into the Skp1 band after SDS-PAGE. Normal (RHΔΔ) and glt1Δ mutant extracts were assayed for 1.5 or 3 h as indicated. Negligible incorporation was detected in the rest of the SDS-PAGE lane (data not shown). These data, which were collected in parallel on the same extracts, are representative of other trials conducted under varied conditions (data not shown). B, after electroblotting from the SDS-PAGE gel to a polyvinylidene difluoride membrane and staining with Ponceau, the Skp1
band and surrounding regions were separately acid hydrolyzed and subjected to high pH anion-exchange chromatography on a Dionex PA-1 column in the presence of internal standards. Monosaccharide elution was monitored by pulsed amperometric detection (nC) and fractions were collected for scintillation counting for $^3\text{H}$. Similar results were obtained for UDP-$[^3\text{H}]\text{Gal}$ (illustrated) and UDP-$[^3\text{H}]\text{Glc}$ reaction products (data not shown).

FIGURE 7. Purification of recombinant Glt1 and Toxoplasma FGGn-Skp1. A, recombinant His$_6$Glt1 was purified from E. coli over Co$^{2+}$-Talon and Q-anion exchange columns. Protein was eluted from the Q-column using a gradient of NaCl and monitored at $A_{280}$. B, fractions were assayed for glucosyltransferase activity using FGGn-pNP as the acceptor substrate. C, fractions were analyzed for Glt1 protein by Western blotting with an anti-His$_6$-tag antibody. The region of the gel showing reactivity with the antibody is shown. $D$, fractions were analyzed for total protein using SDS-PAGE and Coomassie blue staining. Only the region of the gel corresponding to C is shown. Equal volumes of fractions were analyzed in B-D.

FIGURE 8. Substrate specificities of His$_6$Glt1. A, comparison of glycosyl-donor specificities, based on UDP-sugar hydrolysis. His$_6$-Glt1. Recombinant Glt1 was incubated with each UDP-sugar at 50 $\mu$M, and released UDP was detected using the UDP-Glo$^\text{TM}$ assay. Data are representative from three trials, and error bars represent standard deviation of the mean for one reaction conducted in triplicate. $B$, comparison of utilization of UDP-Glc or GDP-Man as donors in a transferase assay toward 1.5 mM FG-Bn, which represents the non-reducing terminal disaccharide of Skp1, using a UDP or GDP (NDP) detection assay. Standard deviations of the mean of three technical replicates are shown. $C$, dependence of transferase activity on UDP-Glc concentration, using 2 mM FG-Bn as the acceptor substrate in reactions monitored by incorporation of $^3\text{H}$ from UDP-$[^3\text{H}]\text{Glc}$. Data represent the mean of two technical replicates ± standard deviation. The $K_m$ calculation is based on the Michaelis-Menten model. $D$, time dependence of reactions comparing FG-pNP with FGGn-Bn, which models the full Skp1 trisaccharide, assayed as in C. The acceptor and donor substrates were 2 mM and 2 $\mu$M, respectively. Error bars indicate standard deviation of the mean between duplicate samples. $E$, comparison of different potential acceptor substrates, as in $D$, in reactions incubated for 2 h. $F$, dependence of transferase activity on FGGn-pNP concentration, in the presence of 40 $\mu$M UDP-Glc. Data are representative of two independent experiments, and the kinetic calculations are based on the Michaelis-Menten model. $G$, dependence on Toxoplasma FGGn-Skp1 concentration, as in $F$. $H$, biochemical complementation of parasite extracts with His$_6$Glt1. Transferase reactions were conducted using UDP-$[^3\text{H}]\text{Glc}$ as the donor, desalted S100 extracts of parental (RHΔΔ) or glt1Δ extracts as the acceptor, in the presence or absence His$_6$Glt1. Samples were separated by SDS-PAGE and, after fixation, gel lanes were sliced from top to bottom. Incorporation was determined by scintillation counting.

FIGURE 9. Linkage analysis of the His$_6$Glt1 reaction product. FGGn-pNP was incubated with UDP-Glc in the presence of His$_6$Glt1. The reaction product was recovered on a C$_{18}$-Sep-Pak and subjected to NMR analysis. $A$, 1D 900 MHz $^1\text{H}$-NMR spectrum of Glc-$\alpha$1,3-Fuc-$\alpha$1,2-Gal-$\beta$1,3-GlcNAc-$\alpha$-pNP. $B$, 2D $^1\text{H}$–$^{13}\text{C}$-HSQC spectrum. All expected peaks are observed and annotated in this region except for Fuc-6 and GlcNAc methyl groups, which are annotated in $A$. Number indicates ring position. $C$, 2D $^1\text{H}$–$^{13}\text{C}$-HMBC spectrum region demonstrating anomeric carbon to ring proton correlations (red dashed lines). $D$, 2-D $^1\text{H}$–$^{13}\text{C}$-HMBC spectrum region demonstrating ring carbon to anomeric proton correlations (red dashed lines) and cartoon diagram of the $T_g$ Skp1 tetrasaccharide attached to pNP with linkages indicated. Glycan symbols used are adopted from Ref. 20. In $C$ and $D$, interglycosidic correlations are annotated in red.

FIGURE 10. Role of Glt1 in parasite growth. Human foreskin fibroblast monolayers were inoculated with the indicated strains in parallel, at an MOI of 0.002. After 5.5 d, monolayers were stained with crystal violet. The images were digitized and areas of manually encircled plaques, representing cleared areas of the monolayer owing to cell lysis, were measured and analyzed using GraphPad Prism. Average plaque areas ± SEM from a representative experiment are shown.
**Figure 1**

A. *Toxoplasma* GT1 proteome (8460 proteins)

Glycosyltransferase prediction using dbCAN, superfamily, pfam and CAZY database search (total 45)

Cytosolic glycosyltransferase prediction using SignalIP, TMHMM (total 10)

BlastP search against NCBI non-redundant databases for phylogenetic co-distribution with Gnt1 and PgtA (total 2)

A. TGGT1_205060 (Glt1)

B. TGGT1_310400 (Gat1)
Figure 3
glt1Δ/glt1+ Skp1 extracted ion scans:  
glt1Δ Skp1 extracted ion scans:

A. Western blot analysis

B. unmodified peptide

Intensity: 7.63E7  
BP: m/z 1011.0042  
theoretical m/z: 1011.0024

Intensity: 1.35E7  
BP: m/z 674.3380  
theoretical m/z: 674.3373

Intensity: 3.58E7  
BP: m/z 1011.0054  
theoretical m/z: 1011.0024

Intensity: 2.68E7  
BP: m/z 674.3391  
theoretical m/z: 674.3373

C. HexHexNAc-peptide

Intensity: 4.39E6  
BP: m/z 1201.5657  
theoretical m/z: 1201.5660

Intensity: 2.92E6  
BP: m/z 801.3818  
theoretical m/z: 801.3797

Intensity: 4.12E6  
BP: m/z 1210.5707  
theoretical m/z: 1210.5660

Intensity: 1.52E7  
BP: m/z 801.3804  
theoretical m/z: 801.3797

D. dHexHexHexNAc-peptide

Intensity: 0  
BP: m/z 1274.5977  
theoretical m/z: 1274.5949

Intensity: 0  
BP: m/z 850.0679  
theoretical m/z: 850.0674

Intensity: 1.28E8  
BP: m/z 850.0674  
theoretical m/z: 850.0674

Intensity: 2.12E8  
BP: m/z 1274.5977  
theoretical m/z: 1274.5949

E. Hex2dHexHexHexNAc-peptide

Intensity: 1.16E8  
BP: m/z 1436.6494  
theoretical m/z: 1436.6477

Intensity: 0  
BP: m/z 958.1023  
theoretical m/z: 958.1008

Intensity: 2.08E8  
BP: m/z 958.1023  
theoretical m/z: 958.1008

Intensity: 0  
BP: m/z 958.1008  
theoretical m/z: 958.1008

116 118 120 122 124 126 128 130 132 134 136 138 140 142 144  
HPLC retention time (min)
Figure 5

A

B

C

HexNAc-Hex-dHex

I F N I V N D F T (HyP) E E E A Q V R

y16 y15 y14 y13 y12 y11 y10 y9 y8 y7 y6 y5 y4 y3

b3 b4 b5 b6 b7 b8 b9 b10 b11 b12 b13 b14 b15 b16

b3 375.33

y3 402.37

b4 473.40

y4 502.53

b5 560.53

y5 602.53

b6 701.61

y6 731.51

b7 816.61

y7 860.65

Figure 5
Figure 6

A

| Time | Strain | FGGn-Skp1 | UDP-[3H]galactose | UDP-[3H]glucose |
|------|--------|-----------|-------------------|------------------|
| 3h   | w/t    | w/t       | -                 | +                |
| 1.5h | +      | +         | +                 | +                |
| 3h   | glt1Δ  | +         | +                 | -                |
| 3h   | w/t    | w/t       | +                 | +                |
| 1.5h | +      | +         | +                 | +                |
| 3h   | glt1Δ  | +         | +                 | +                |

B

UDP-[1H]galactose reaction

Fuc, Gal, Glc, Man

dpm

time, min
Figure 7

A

B

C

D

at National Institutes of Health Library on October 17, 2017
Figure 8

(A) Ratio to UDP detected enzyme activity.

(B) mN NDP/min.

(C) Km = 6.0 ± 0.77 µM

(D) Donor: UDP-[3H]Glc

(E) Donor: UDP-[3H]Glc

(F) Km = 13 ± 3.0 mM

Vmax = 330 ± 47 nmol/h/µg

(G) Km = 4.2 ± 1.5 µM

Vmax = 1.67 ± 0.37 nmol/h/µg

(H) Donor: UDP-[3H]Glc
Figure 10

Strain at National Institutes of Health Library on October 17, 2017
SUPPLEMENT

*Characterization of a cytoplasmic glucosyltransferase that extends the core trisaccharide of the Toxoplasma Skp1 E3 ubiquitin ligase subunit*

Kazi Rahman1,4,8, Msano Mandalasi1,5, Peng Zhao2, M. Osman Sheikh3, Rahil Taujale2,3, Hyun W. Kim1, Hanke van der Wel1, Khushi Matta6, Natarajan Kannan1,3, John N. Glushka2, Lance Wells1,2, Christopher M. West1,5,7

1Department of Biochemistry & Molecular Biology, 2Complex Carbohydrate Research Center, 3Institute of Bioinformatics, University of Georgia, Athens, GA 30602 USA; 4Departments of Microbiology & Immunology, and 5Biochemistry & Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104 USA; 6Chemical and Biological Engineering, State University of New York, Buffalo, NY 14260 USA

Contents:

SUPPLEMENTAL TABLE S1. Oligonucleotide sequences employed.
SUPPLEMENTAL FIGURE S1. Sequence of Glt1 cDNA.
SUPPLEMENTAL FIGURE S2. Alignment of Glt1-like sequences with other CAZy GT32 family proteins.
SUPPLEMENTAL TABLE S1. Oligonucleotide sequences employed.

**Targeting sequence amplification:**

**Tgglt1-disruption**

a) Glt1F1 5’-flank 5’-end 5’-CCGGCCCAAGCAACCTCTCTCTCTAT (ApaI)

a’) Glt1R1 5’-flank 3’-end 5’-GGAAGCTTCAATTCACACTCAAGGAATG (HindIII)

b) Glgt1F2 3’-flank 5’-end 5’-GCTCTAGAGAGATGGCGCAGGATTCACTCTAT (XbaI)

b’) Glt1R2 3’-flank 3’-end 5’-GGCGGCCGCTTGTTAGGCATGCTCTCTT (NotI)

**Tgglt1 complementation** (italics denote complementary regions)

Vector (forward): 5’-GCAGGCCAGATTCTCAAG

Vector (Reverse): 5’-TATTGATGAAACCCCTTGTTCC

Insert (Forward): 5’-ACGGGTTCATCAATAGCTGGCAGATCTTTAAGT

Insert (Reverse): 5’-GAAAATCGTGGCCTCCACATAACTCTTGCTTGACAGAC

**Mutation of Tgglt1 complementation construct:**

\[ \text{363 DVD}^{365} \text{ to } \text{363 AVA}^{365} \]

Forward: 5’-ggtttttatatgccctccagcttccagcagctggctgtatattctgat-3’

Reverse: 5’-attcaggaatatgcgagcttggtggcctgcaggttataaaacc-3’

**D348N:**

Forward: 5’-gcagggcggcaaatattcgatttttcgcagcgcacag-3’

Reverse: 5’-ccgggtgcaaaatcgaatattgcccgcctgc-3’

**PCR confirmation of Tgglt1 disruption and complementation**

**PCR 1**

Forward: 5’-CTGTGACTCAGAACTCCTCAAG

Reverse: 5’-CCTCCGAGAACATCCGATTAGAG

**PCR 2**

Forward: 5’-GCTAGTGTACAGTTCCCACTAAG

Reverse: 5’-AGTCGCGGAACATCTCGTTGAAGT

**PCR 3**

Forward: 5’-ATTTCATCTGAGATAAGGCTCTCGC

Reverse: 5’-TCGAGCATGATTCGCTCTATAG
SUPPLEMENTAL FIGURE S1. Sequence of Glt1 cDNA. Codon optimized synthetic nt sequences of *Toxoplasma* Glt1 cDNA is in black lower case. The translated amino acid sequence is below in black upper case, and matches TGGT1_205060. The N-terminal His$_6$ tag contributed by pET15b(TEV) is in purple. Numbering refers to the sequence of the native cDNA and protein. Peptides from Glt1 detected in strains GT1 and ME49 as recorded at ToxoDB are in red, with the first amino acid underlined. A detected phosphosite is asterisked in red.

**Atgggcaagcacccagcatcactcatatcactcatcagagcacagagaaactttgtatttccagggccat**

M G S S H H H H H H H S G R E N L Y F Q G H pET15b(TEV)

**atgcgtgaagcctgggtgactgtagcgtggattgtctgtcctgcgaaaaacatggt**

M L R E S L G D C S V D C L S S C E K H G 20

**aacatgcgttctccgctgatgcctctggcgtcatctgtaaacctggagctctgaa**

N M R S L D A R G R L A E T T Y R Q N R 40

**atggattttcccagactcccagctcccagcttcagctcccagttaactctggagctctgaa**

I D F P S H S V I Q F N L W S S E 60

**gattaatccgtgctagtctgcctggtggcctgcctgcctgatcgggagctctgaa**

D N T W L S L L R G L H R C N A R R V 80 peptide 1

**ctgcgtgcaagcacagctcagctgctgtagaaaagccagcacatcgctgtgaggtttata**

L L A A S G D P R E P A A L H S E E V T S 100

**aaacgctccgctgtgcgggtgtcgcacagcgagtctgtaactctgtgactgttgaactgtg**

K R H A L P G V A T P L N C D S E L L 120

**aatgcctataaaatggcgttgattctgtaaagggctgtgaccatcactgtccggcaaccggca**

K C Y K N G G D L K G V D H H C P Q P A 140

**ccgcgtgataagcctgtgtgatcgcgtgacgcttgagcagctcattctgcctgcgtgaccgctggtc**

P A D S R G D R D G E A F C A D R A L S 160 peptide 2

**gcagccgctgtaagcctgtgtgatcgcgtgacgcttgagcagctcattctgcctgcgtggtc**

A S L A Q A I R E A K L Y C G H P E R T 180 peptide 3

**gtgcgtgattgcgcagctccttttctgcctgtagttctagaaagcccgtgtcctcagcggctgga**

V V D C R C L E E T G R H T P P E 200

**ctgcgcttttctgcctgtagttctactcatcgcagcagtttaaaaacttcaagctctgcctcttt**

L P L V D G G H S S A G C K N F S S A F 220

**atgaaatccgctgtaatggctcagttgaacaccgctctcatacgccggcagaa**

M K I P S E F V S G K A I P P L L H F V 240

**tggctggtggtgcacccggcgcgccttttttcgcgaaacctttgcggcggcgtcatc**

W L G P P P F P F E T I R Q S W A V H 260

**aatccccgagctctgacgaccgtcactgtagtgagcgcggcgtcagttgcacgctgggat**

N P D L I Q A L W T D A H V E S L L D V 280

**ctgcgacgctaatccgctgagctgctgccggcagaaactccgtgaaacccgctcatcgcagcggct**

L D R K S S R R P K C R K T D H Q D L 300

**ccgccgctttgctggtgtcctcagcggcaagcgcgttggtgaacaccggaatgggaaatgg**

H P L L V D R P T A D A G E T T E W E M 320

**tccgttcagctccgggaaccgcagcagcagctttgtctcagatcgccattaaagccttttcgtaaagga**

S D S V P D S Q T L V H A I K A F R K E 340

**tccgttcagctccgggaaccgcagcagcagctttgtctcagatcgccatcagatgcccgtt**

T C P G A K S D I A R L L I L C H Y G G 360
attacgcggatgtggacatggaagccatcaaaccgctgccgccgtgtctgcgtcaatgc 1140
I Y A D V D M E A I K P L P P C L R Q C 380
accacgtgtttaaatgggcgacgcgcggccgatgcagttgaactggtaacgctctgctcgc 1200
T T V F M G M Q R P D A V E L G N A L I 400
ggtcgcaggtcgcgtcatgaactgattctttcatctctcgacgctgttgggcgcccgtcag 1260
G C S S G H E L I R F I L Q R V G R P Y 420
tcgaatgggtacctgagcagcagatggcagtcgctgctggacattctgaactgac 1320
SQWGTSADQMATVLDILKL440 peptide 4
cacgtgacgctggtctccgagccccgagctcagaactgatgggattcgctgcgtagcaccaac 1380
HVARSPEASELMDSLRSTN460
gacgaagcgcagcagatggcctcagcagcgccggtctgtgctaccggtgaaacgctg 1440
DEATNVIERTGPGLLTRL480 peptide 5
gttggctgcgcgtcagtaactcgtcagctgactgtgcacgctggtcagctgcagcag 1500
AWLDRQLNSSSCARCPRQER500
cgctccatgattacgaacacaaatcaaatctacaacgcggcaacactgcaggtcagctgac 1560
RSHDYEHKSNTEATSVHAD520
gttggctgcgcgtcagcagctgactgtgcacgctggtcagctgcagcagcgcgg 1620
VVGESTGSKEDDSKPEATSG540 acgtgcgttgcattctgaccaacgggtagcaaaagatgactctaaaccggaagcaacgctgac 1680
TCCILATTCCICPPFFYPV560 ccgaatctgctgcaaaagactgctgcgaagcagaaactgacctgcagcagcagctgcaagaaatct 1740
PNHRRELREGKVVQTEHLES580 agttttctcagctcactcagctgctgttcagactgatctttagctcctgcgatctttttctatatcggcagtt 1800 SFSYTVHWRQTWQDSVRQE600 peptide 5 agcga aaatgaaatgtaa 1818 SENEC - 605 *phosphosite 1
SUPPLEMENTAL FIGURE S2. Alignment of Glt1-like sequences with other CAZy GT32 family proteins. Alignment of Glt1-like sequences (third grouping from the top) with representative characterized CAZy GT32 proteins from eukaryotes (top group), from bacteria (second from top grouping), and from non-Glt1-like sequences from other protists (bottom group). Amino acids are color-coded with respect to chemical similarities that were used as the basis for manual alignment giving preference to the registration of hydrophobic residues: green, hydrophobic; blue, acidic, dark red, basic; bright red, structure breaking (P or G). Positions occupied by identical amino acids across all the organisms, or within all of the Glt1-like sequences, are bolded and asterisked. Predicted signal peptides near the N-terminus are italicized. Similar amino acids are highlighted when they have specific representation at a given position. Sequence identification numbers are at the bottom.

| S. 5 |

| Hs1 | 1 | HKSPDPDLLLRLRAGQPVRCTFLIIIGKFKTFFVSIMYNYHVGEPKQELNYMPAEIPCLFTPP |
| Hs2 | 1 | HKRDELQLSTVLLVICGFLYQPTLKLSCLCPLSFKSHQGLLEAL |
| Dm | 1 | MLMLWPIAVARFMIIILVMICYGFLYITSENYHSCHFIEQVLAQATQALDGETNLLDV |
| C1 | 1 | MSSYQLQLTTLTVFSPHRAGFALFILIIISFLVGSATVQGKGDGQTPNQRSLEDFTSPHT |
| A1 | 1 | MRQVIFLIFVNLISFLVRSFVTLLLVLVDRASADAILHEIPSNSLIEQFQ1- |
| M1 | 1 | MSYFQGMFLIHALHIGGTLKLAGAECSVQGQRVCYHTYFVTDEPSGSGIGCDAN1- |
| Cj | 1 | MTEIESFYWPARKQYRIGMIGELTIKSWLDFHYTVDNLKIDFLQGDNLKIDAN1- |
| Sc | 1 | MRSKLILHATIRLKSITJTVTLVLIYLSLTFVSLNKRLLQFYPSKDFQTLFPTTSHQDINLKQTVNKKNQLHNLAGQSFYAPDS |
| Tg | 1 | HRESLGDCSVCLSSCEKHGRSLIDRGLAETTYQRRQIDGPDFSFSVQFLQNSS- |
| Hh | 1 | MLBDARGLIAETTYQRRQIDGPDFSFSVQFLQNSS- |
| Nb | 1 | MCVIPFNLIRFSPNLDLFPRFVRPFSINRNLQPI- |
| Sn | 1 | HNGTERQHTSPQRFFQPFDDRQERGFRQNNNG- |
| Rf | 1 | CQFVNSFKEIILFYLCLNLVYTKSLEENVNLYKSNFDRNLKLKEMTKREATYFGIMKIDAEASER1ICNLDQT |
| Bn | INTEELMNRRPSTQETFQNTGNNKGAKDEGHEINQYLALASSFESMCQMRRKLRLGVSSE- |
| Ot | 1 | HENSSAGQIDTLFSDRTVEKR- |
| Ng | 1 | HTHSTLNTSSQGKLPLDLTNADG- |
| To | 1 | HTRFBFFPFRKLARRKLYDFRESE |
| Ac | 1 | VTSMSANLICYYIRHNLVIGYIPKRRCTCHSNKIPMKVCKNNFDP- |
| Pu | 1 | MDAFNGKVRGWVID- |
| Aa | 1 | MAALLDLEPTWTDFPLHHRGWS- |
| Es | 1 | HIRWNSLPI- |
| Ap1 | 1 | HRRKHAEE- |
| Gt | 1 | HNQMDGCVGKCSSAMGRSRSWDDESSIEFDELRIWREGESERDRLVRGPINS- |
| Eh | 1 | HMKHQLRIFQFRRDFDKGTAVLSFSTYEPKQFDQLYQRMLAAAMLRSRENSQCH- |
| Pp | 1 | MPFPEVQSORSHTRAWAAWSTTSLRCLVFFTLTIILSVLSVVLVSPRIRFAPAKLQSFHI |
| Ap2 | 1 | MMGICYLQSTVAVLAFQLVDFSTPPSLASTDTPFVQVLRTPPHLINSTT |
| Sp | 1 | MHDDVTVYKVARRGCAPFGCVLIALTMLWHLFLLCALHPSTLRHHAAPSTSRRKATPEPAAYQEALSS- |

| Hs1 | -TPSHGPTPNIFFLFLELTST- | -NPFLFMCSVEBAART |
| Hs2 | -SHRRPIFVFLFLELHS- | -EPFLVSCSVEBAKI |
| Dm | LQADPKFSFKNSLFFRDLSCLSKQFLKTLVRQADCAABAMI |
| C1 | FQADPKFSPGNSLFFRDDLSCLSKQFLKTVRQADCAABAMI |
| A1 | -SPBPPGGFKSPGNSLFFRDDLSCLSKQFLKTVRQADCAABAMI |
| M1 | - |
| Cj | - |
| Sc | QAP - |
| Tg | LLAS - | -DPREPAHLSEETSRKRHLFGAVAIPLNCDELLKCYRNGDLKSVDHCPQAAPABRGDRGIEACFADKAL8 |
| Hb | LLAS - | -DPREPAHLSEETSRKRHLFGAVAIPLNCDELLKCYRNGDLKSVDHCPQAAPABRGDRGIEACFADKAL8 |
| Nc | LAAASKAAKSKAAACGDPDHQADLQYAVLNSAAMTKSMRSNIAAPDDTFPCHBQGKGG- |
| Sn | LAR - | -DALLGHVTRAGFVRTTHSELGSAPEG- |
| Rf | RDFQDFQKSVRDDDHHVMNTT1NHNDDDDDGGKSWD- |
| Bn | RVLIEKDQRTMESTKRM1- |
| Ot | DQPREVHVIDLSQSR- |
| Ap1 | APSPKGGPV- |
| Gt | RTOQKNGSCECER |
|     | W318         | D340   | R351     | D363     | D365     |
|-----|-------------|--------|----------|----------|----------|
| Hs1 | PDDRELFQ    | LADNY  | AAVQGKYP | YLVDSDAL | WLVF   |
| Hs2 | LKDRKLLLLD | FSHVY  | NQINASAR  | HSDASRAL | YWYDTS  |
| Dm  | RLESYGALT  | PFCXV  | KDGRLSRKY | SHSDFRL | LLYHYYV |
| C1  | PDLATLFQ    | PLAKY  | SQFEPKYP  | YLSDACR  | IMLDPT  |
| Af  | PEVWRAAQSS  | DFEPL  | TRKSERF    | KTLFDT  | FGCXK   |
|     | ------------ | -------| ----------|----------|----------|
| Sc  | DQVSLSDBG  | IIPFNL  | LYA-FVIP  | IAPFKL  | PGMNLKDF  |
|     | ------------ | -------| ----------|----------|----------|
| Tg  | TVER-SDSVPDQ | TVHAI  | KATFK  | SCGASDAL | LLHICY-YAYDT  |
| Hh  | TVER-DFSA  | TLVSAI  | KATFK  | SCGASDAL | LLHICY-YAYDT  |
| Nc  | 1VER-DFSA  | TLVSAI  | KATFK  | SCGASDAL | LLHICY-YAYDT  |
| Sn  | 1VER-DFSA  | TLVSAI  | KATFK  | SCGASDAL | LLHICY-YAYDT  |
| Rf  | NLEYITYMT  | DSECMN  | MLLAKHGN | SQSVDL | LKFLLRF-YVIDFET |
| Bn  | HYRYIKRDIEE | FLRLH  | CKHLDL  | QSSVEKSD | LVARLYVE-YAYDFTC |
| Ot  | TGEKOSHEL  | LINDRT  | NIVRDARL  | YSEGLTT  |
| Ng  | NTERTADVA  | DLNQ  | HSAMA  | FNTRGKSL | LFLDRLLE-YYDTSE  |
| To  | FPPONRTTAVIMEG=LKSQTSEK-EMHABQ1SNRSHY1-LKCNVNI  |
| Ac  | ONEKIDTOYST | QVEN  | EENL  | -NVRGKDL | YAYDFTC |
| Pu  | NVEKADFQ  | -      | -      | -        | -        |
| Aa  | DVEKARNPAQA  | ELNA  | -      | -        | -        |
| Es  | TVRMTADENV  | SEKRN  | -      | -        | -        |
| Ap1 | DVEKTMAT  | KLEQGO  | -      | -        | -        |
| Ot  | TVEINO-X        | AUKV  | -      | -        | -        |
| Gf  | XAPFDDAKMES  | LLARKHA  | -      | -        | -        |
| Pp  | TVYTVTADNLR  | LEFEL  | -      | -        | -        |
| Ap2 | TVYTVTADN  | LEFEL  | -      | -        | -        |
| Sp  | TVYTVTDDANLA  | LEARH  | -      | -        | -        |

**G385**

|     | E393 | N396 | H405 |
|-----|------|------|------|
| Hs1 |       |      |      |
| Hs2 |       |      |      |
| Dm  |       |      |      |
| C1  |       |      |      |
| Af  |       |      |      |
| Mc  |       |      |      |
| Cj  |       |      |      |
| Sc  |       |      |      |

**Tg**

|     | RQCTTFHN  | QRPAEL  | ALNLCSG  | HELRIFL  | QVRGRQP  |
| Hh  | RQCTTFHN  | QRPAEL  | ALNLCSG  | HELRIFL  | QVRGRQP  |
| Nc  | RQCTTFHN  | QRPAEL  | ALNLCSG  | HELRIFL  | QVRGRQP  |
| Sn  | LDDTIVGAG  | QLDNG  | -      | -        | -        |
| Rf  | QQSLGDCHRSRPQCVLSPMG   | SMTSLFRVNNPHASKK  | HELFKILDNLNLDKFOE |
| Bn  | QQSFYQAL  | SNTUVBBNNIAGK  | HELCAYLNNQTDQFN |
| Ot  | DIPFTDI  | SNTKAPLA  | FLKHCNQ  |
| Ng  | HLDPPQ  | QQRGPQYAP  | SNTKAPLA  |
| To  | SPSDLPLPECNLA  | SNACGYNMHIHAGG  | HELCAYLNNQTDQFN |
| Ac  | QAFQL  | SNTUVBBNNIAGK  | HELFKILDNLNLDKFOE |
| Pu  | KAPFISQA  | ANTVGNS  | SHSVLTH  |
| Aa  | EADFFVQP  | SNVCASNNINLQAK  | HELFKILDNLNLDKFOE |
| Es  | KREQFP  | STNTUVBBNNIAGK  | HELFKILDNLNLDKFOE |
| Ap1 | KETFQYS  | ENTEFYQAL  | SNTSNS     |
| Ot  | RPPDFD  | -      | SNVQIYNNLSCAG  |
| Eh  | RLNNAS  | ILSDQALFK  | IADVRSQTPMTAATQ  |
| Pp  | DFDDD  | RENNFLSSLEPHVNL  | EKSDKALNMAARAR  |
| Ap2 | QMMDL  | DLPFLFSTEPLVHNL  | EGASAPFLYNAASAG  |
| Sp  | SCDL  | EKDFLSTFEPLVHNL  | EKDFLSTFEPLVHNL  |

**S-7**
### Sequence IDs:

**Hs1** Homo sapiens (BAAS9519.1): Hsa4galt, lactosylceramide α4- galactosyltransferase  

**Hs2** Homo sapiens (AAD48406.1): Hsa4gnt, α4-N-acetylglucosaminyltransferase  

**Dm** Drosophila melanogaster (AAF51162.1): Dma4gt1: α4GT1  

**Cl** *Columbia livia* (AADC8388.1): collagen GT, α4-galactosyltransferase  

**Af** *Aspergillus fumigatus* AT293 (EAL85572.1): Mipa, MIPC synthase subunit (SurA), α-mannosyltransferase  

**Mc** *Moraxella catarrhalis* 2951 (AAZ29048.1): Lgt5, α1,4-galactosyltransferase  

**Cj** *Campylobacter jejuni* (AAM90647.1): Cgd1, α1,4-galactosyltransferase  

**Sc** *Saccharomyces cerevisiae* (P13755): Oeh1, α1,6-mannosyltransferase  

**Tg** *Toxoplasma gondii* (EPR61400.1)  

**Hh** *Hammondia hammondi* (XP_008887433.1)  

**Nc** *Neospora caninum* (CEL66241.1)  

**Sn** *Sarcocystis neurona* (EupathDB:SRNC_1642)  

**Rf** *Reticulomyxa filosa* (X6P0U2)  

**Bn** *Bigelowiella natans* (JGI:aug1.98_g20090)  

**Oy** *Oxytricha trifallax* (J9R850)  

**Ng** *Nannochloropsis gaditana* (W7UA89)  

**To** *Thalassiosira oceanica* ([EJK66288.1](https://www.ncbi.nlm.nih.gov/nuccore/EJK66288.1))  

**Ac** *Albugo candida* ([CC139677.1](https://www.ncbi.nlm.nih.gov/nuccore/CC139677.1))  

**Pu** *Pythium ultimum* (K3WJQ7)  

**Aa** *Aureococcus anophageferens*  

**Es** *Ectocarpus siliculosus* (CBJ34172.1)  

**Ap** *Aphanomyces astaci* ([XP_009836009.1](https://www.ncbi.nlm.nih.gov/nuccore/XP_009836009.1))  

**Gt** *Guillardia theta* ([XP_005836918.1](https://www.ncbi.nlm.nih.gov/nuccore/XP_005836918.1))  

**Eh** *Emiliania huxleyi* (R1FXS0)  

**Pp** *Phytophthora parasitica* (V9EJ41)  

**Ap** *Aphanomyces*  

**Sp** *Saprolegnia parasitica* (A0A067C728)
Characterization of a cytoplasmic glucosyltransferase that extends the core trisaccharide of the Toxoplasma Skp1 E3 ubiquitin ligase subunit
Kazi Rahman, Msano Mandalasi, Peng Zhao, M. Osman Sheikh, Rahil Taujale, Hyun W. Kim, Hanke van der Wel, Khushi Matta, Natarajan Kannan, John N. Glushka, Lance Wells and Christopher M. West

J. Biol. Chem. published online September 19, 2017

Access the most updated version of this article at doi: 10.1074/jbc.M117.809301

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2017/09/19/M117.809301.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2017/09/19/jbc.M117.809301.full.html#ref-list-1