Identification of a glycosylphosphatidylinositol anchor-modifying β1-3 N-acetylglucosaminyl transferase in *Trypanosoma brucei*

Luis Izquierdo,1 Masayuki Nakanishi,2 Angela Mehlert,1 Greg Machray,1 Geoffrey J. Barton1 and Michael A. J. Ferguson1*

1Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK, 2College of Pharmaceutical Sciences, Matsuyama University, 4-2 Bunkyo-cho, Matsuyama, Ehime 790-8578, Japan.

Summary

*Trypanosoma brucei* expresses complex glycoproteins throughout its life cycle. A review of its repertoire of glycosidic linkages suggests a minimum of 38 glycosyltransferase activities. Of these, five have been experimentally related to specific genes and a further nine can be associated with candidate genes. The remaining linkages have no obvious candidate glycosyltransferase genes; however, the *T. brucei* genome contains a family of 21 putative UDP sugar-dependent glycosyltransferases of unknown function. One representative, *TbGT8*, was used to establish a functional characterization workflow. Bloodstream and procyclic-form *TbGT8* null mutants were created and both exhibited normal growth. The major surface glycoprotein of the procyclic form, the procyclin, exhibited a marked reduction in molecular weight due to changes in the procyclin glycosylphosphatidylinositol (GPI) anchor side-chains. Structural analysis of the mutant procyclin GPI anchors indicated that *TbGT8* encodes a UDP-GlcNAc: β-Gal-GPI β1-3 GlcNAc transferase. This is only the second GPI-modifying glycosyltransferase to have been identified from any organism. The glycosylation of the major glycoprotein of bloodstream-form *T. brucei*, the variant surface glycoprotein, was unaffected in the *TbGT8* mutant. However, changes in the lectin binding of other glycoproteins suggest that *TbGT8* influences the processing of the poly N-acetyllactosamine-containing asparagine-linked glycans of this life cycle stage.

Introduction

The African trypanosomes are protozoan parasites that cause human African sleeping sickness and Nagana in cattle. The estimated number of cases is currently between 50 000 and 70 000 and, without treatment, sleeping sickness is fatal (WHO, 2006). The parasite undergoes a complex life cycle between the mammalian host and the blood-feeding tsetse fly vector (*Glossina* sp.). Throughout this life cycle, *Trypanosoma brucei* is coated by several million copies of glycosylphosphatidylinositol (GPI)-anchored proteins. The bloodstream form of the parasite in the mammalian host is covered by a coat of 5 × 10⁶ variant surface glycoprotein (VSGs) homodimers and it evades the immune system by periodically replacing the existing VSG coat by a different one, a phenomenon known as antigenic variation (Cross, 1996). When bloodstream-form parasites are ingested by the tsetse fly, they differentiate to the procyclic form in the insect midgut. During this process the VSG coat is replaced by a new coat (Guther et al., 2006) that includes a set of GPI-anchored glycoproteins known as procyclins. These are characterized by internal dipeptide (EP) or pentapeptide (GPEET) repeats which confer a rod-like structure to the protein (Roditi et al., 1989; Treumann et al., 1997). *T. brucei* strain 427, the strain used in this study, contains (per diploid genome) two copies of the GPEET1 gene encoding six GPEET repeats, one copy each of the EP1-1 and EP1-2 genes, encoding EP1 procyclins with 30 and 25 EP repeats, respectively, two copies of the EP2-1 gene, encoding EP2 procyclin with 25 EP repeats and two copies of the EP3-1 gene, encoding EP3 procyclin with 22 EP repeats (Acosta-Serrano et al., 1999). The EP1 and EP3 procyclins contain a single N-glycosylation site, occupied exclusively by a conventional Man₉GlcNAc₂ oligosaccharide, at the N-terminal side of the EP repeat domain (Treumann et al., 1997; Acosta-Serrano et al., 1999). Neither EP2 or GPEET procyclins are N-glycosylated and GPEET1 procyclin is phosphorylated on six of seven Thr residues (Mehlert et al., 1999; Schlaeppi et al., 2003).
Procyclins, both GPEET and EP, contain similar GPI membrane anchors. These are the largest and most complex anchors known and they are characterized by the presence of large poly disperse branched N-acetyllactosamine (Galβ1-4GlcNAc) and lacto-N-biose (Galβ1-3GlcNAC)-containing side-chains (with an average of about 8–12 repeats, depending on the preparation) that can be capped with α2-3-linked sialic acid residues (Ferguson et al., 1993; Treumann et al., 1997). Sialic acid is transferred from serum sialoglycoconjugates to terminal β-galactose residues by the action of a cell surface trans-sialidase enzyme (Engstler et al., 1993; Pontes de Carvalho et al., 1993; Montagna et al., 2002; 2006) and trans-sialylation of surface components plays a role in the successful colonization of the tsetse fly (Nagamune et al., 2004). In addition, it has been postulated that the branched side-chains of the anchor form a dense glycocalyx that contributes to the protective function of the coat against digestive enzymes in the fly midgut (Acosta-Serrano et al., 2001).

The survival strategies of protozoan parasites frequently involve the participation of glycoconjugates. T. brucei presents many surface glycoproteins containing Gal (Mehlert et al., 1998a; Atrih et al., 2005) and a UDP-Glucose 4⁻-epimerase (TbGalE) conditional null mutant showed that this gene, and hence UDP-Gal, is essential for the survival of the parasite in both the bloodstream and procyclic-form life stages (Roper et al., 2002; 2005). From these experiments, it was possible to conclude that one or more of the UDP-Gal-dependent glycosylation pathways are essential to the parasite. This provided the impetus to identify and characterize glycosyltransferase (GT) genes in the parasite. GTs catalyse the transfer of sugar from an activated sugar donor to an appropriate acceptor, typically another sugar, lipid, protein or small molecule. Current estimates suggest that about 1% of the open reading frames (ORFs) of each eukaryotic genome are dedicated to the task of glycosidic bond synthesis (Coutinho et al., 2003; Davies et al., 2005).

In this paper, we mine the T. brucei genome to identify candidate UDP sugar-dependent GTs. One example of these putative GTs (Tb10.389.1450, referred to here as TbGT8) was examined experimentally by gene replacement.

Results

Identification of putative GT genes in the T. brucei genome

A survey of the known T. brucei glycoprotein N-linked glycan structures (Bangs et al., 1988; Zamze et al., 1990; 1991; Strang et al., 1993; Treumann et al., 1997; Mehlert et al., 1998a; Acosta-Serrano et al., 2004; Atrih et al., 2005), GPI anchor glycan structures (Ferguson et al., 1988; Treumann et al., 1997; Mehlert et al., 1998a,b; A. Acosta-Serrano, J.M. Richardson and M.A. Ferguson, unpubl. data) and the other known glycoconjugates, dolichol-phosphate-mannose and β-glucosyl-5-hydroxymethyluracil (base J) (Gommers-Ampt et al., 1993), shows that the parasite makes a minimum of 38 distinct glycosidic linkages (Fig. 1). Three of these, Galβ1-4GlcNAc, GlcNAcβ1-3Gal and GlcNAcβ1-6Gal, occur in poly LacNAC branches in two different contexts, i.e. attached to N-linked glycans and GPI anchors in bloodstream-form and procyclic-form glycoproteins respectively. When the glycosidic linkages and context are not unique to T. brucei (Fig. 1, linkages 1–21 and 34), genome annotations and/or BLAST searches for T. brucei homologues of known eukaryotic GTs identified several plausible T. brucei candidate genes (Table S1). However, T. brucei candidates for linkages 12, 13, 14, 15, 16 and 17 (Fig. 1) were noticeably absent (see Discussion).

One of the putative GTs (Tb10.389.1450) was selected for further study. This gene encodes a 371-amino-acid protein, with a theoretical mass of 43.5 kDa. Semi-quantitative RT-PCR analysis indicated that mRNA for most of these genes was present in both life cycle stages, with some showing higher expression in one or other life cycle stage (Fig. S2).

One of the putative GTs (Tb10.389.1450) was selected for further study. This gene encodes a 371-amino-acid protein, with a theoretical mass of 43.5 kDa. Semi-quantitative RT-PCR analysis indicated that mRNA for most of these genes was present in both life cycle stages, with some showing higher expression in one or other life cycle stage (Fig. S2).

The predicted protein contains a single DXD, in this case DDD, sequence motif that is common to many GTs and likely to be involved in binding a divalent cation (probably Mn²⁺) required for catalytic activity (Wiggins and Munro, 1998). In addition, the protein has a single predicted N-terminal transmembrane domain between residues 5–27 (Krogh et al., 2001) and is likely to be a type II membrane protein, a typical topology for Golgi apparatus GTs (Colley, 1997). The Tb10.389.1450 gene and flanking sequences were amplified with Pfu polymerase from T. brucei strain 427 genomic DNA and three clones from two separate PCR reactions were sequenced in both
directions. The strain 427 (Accession Number AM988671) sequence was very similar to that in the strain 927 genome database. In the ORF there were five apparent nucleotide polymorphisms, only one of which produced an amino acid change (Val254 in place of Ile254). The strain 427 gene and protein product will be referred to here as \( \text{TbGT8} \) and \( \text{TbGT8} \) respectively.

**Creation of bloodstream and procyclic-form \( \text{TbGT8} \) null mutants**

As the \( \text{TbGT8} \) gene is transcribed in both bloodstream- and procyclic-form parasites, it was decided to analyse its function in both life cycle stages by investigating whether a bloodstream-form null mutant could be created and, if so, whether this could be transformed into a procyclic-form null mutant.

Southern blot analysis using a \( \text{TbGT8} \) ORF probe showed that there is a single-copy of the gene per haploid genome (Fig. S3). Both \( \text{T. brucei} \ \text{TbGT8} \) alleles were replaced sequentially in the bloodstream-form parasite with puromycin acetyl transferase (\( \text{PAC} \)) and hygromycin phosphotransferase (\( \text{HYG} \)) drug resistance cassettes by homologous recombination, and selection on the relevant antibiotic(s) (Wirtz et al., 1999), to generate a \( \Delta \text{TbGT8}::\text{PAC/}\Delta \text{TbGT8}::\text{HYG} \) mutant, referred to here as the \( \text{TbGT8} \) null mutant. Southern and Northern blot analysis confirmed that both \( \text{TbGT8} \) gene copies, and their mRNA transcripts, respectively, were absent from the null mutant (Fig. 2A and B). The normal in vitro growth kinetics of the bloodstream-form \( \text{TbGT8} \) null mutant and its ability to infect mice (data not shown) allow us to conclude that \( \text{TbGT8} \) encodes a nonessential gene in bloodstream-form \( \text{T. brucei} \).

Wild type and \( \text{TbGT8} \) null mutant cells were differentiated from bloodstream-form to procyclic-form cells by growing them at 27°C in SDM-79 media supplemented with 3 mM citrate-\( \text{cis} \) aconitate (Brun and Schonenberger, 1981). Both cell lines transformed and grew normally. These data show that the \( \text{TbGT8} \) gene is not
required for the transformation to, or for the normal growth of, procyclic-form *T. brucei* in vitro. The *TbGT8* gene was added back into the procyclic-form null mutant via a constitutively expressed plasmid (Helms et al., 2006). A Southern blot was performed on procyclic-form DNA from wild-type cells, the *TbGT8* null mutant and the mutant harbouring the add-back, confirming the absence of the gene in the null mutant and revealing the ectopic gene copy in the add-back (Fig. 2C).

The procyclic-form *TbGT8* null mutant expresses smaller procyclins than wild-type cells

In order to perform glycosylation phenotyping on procyclic cells, procyclins from wild type and null mutant cells were extracted (Ferguson et al., 1993) and analysed by SDS-PAGE and periodate-Schiff staining for carbohydrate. This revealed a significant reduction in the apparent molecular weights of the *TbGT8* null mutant procyclins compared with wild type (Fig. 3A). Western blots with anti-EP and anti-GPEET procyclin antibodies showed that wild-type null mutant cells were expressing both EP and GPEET types of procyclins and that both underwent a reduction in apparent molecular weight and an apparent increase in quantity upon deletion of the *TbGT8* gene (Fig. S4). The increase in procyclin copy number to compensate for their decreased size is reminiscent of the situation observed under galactose starvation (Roper et al., 2005). SDS-PAGE analysis of procyclins from the *TbGT8* add-back cell line showed recovery of procyclin apparent molecular weight (Fig. 3). This phenotypic reversion shows that the reduction in procyclin apparent molecular weight seen in the null mutant is specifically due to deletion of the *TbGT8* gene. Add-back of *TbGT8* that had been mutated to convert the DDD motif to ADA or AAA failed to rescue the mutant phenotype, strongly suggesting that *TbGT8* encodes the GT activity.

![Fig. 2.](image1.png)

**Fig. 2.** Southern and Northern blot characterization of *TbGT8* null mutants.  
A. Southern blot of genomic DNA from wild type (lane 1) and the *TbGT8*:PAC/*TbGT8*:HYG null mutant (lane 2) digested with Pst I and probed with *TbGT8* ORF (upper panel) and *Tb10.70.3150 ORF* control (lower panel).  
B. Northern blot of total RNA extracted from wild-type cells (lanes 1) and the *TbGT8*:PAC/*TbGT8*:HYG null mutant (lanes 2) probed with the *TbGT8* probe (upper panel) and the β-tubulin control (lower panel).  
C. Southern blot of genomic DNA digested with Pst I and probed with *TbGT8* ORF of wild type (lane 1), *TbGT8* null (lane 2) and *TbGT8* add-back (lane 3) cells.

![Fig. 3.](image2.png)

**Fig. 3.** *TbGT8* null procyclic cells express smaller procyclin molecules than the wild type.  
A. Periodate-Schiff stained SDS-PAGE gel of extracted procyclins from wild type (lane 1) and *TbGT8* mutant cells (lane 2) and *TbGT8* null mutant cells with wild type (lane 3) and mutant forms (lanes 4 and 5) of *TbGT8*. The mutant forms of *TbGT8* contain 254ADA256 and 254AAA256 in place of the wild-type 254DDD256 motif.  
B. To confirm expression of the *TbGT8* constructs total RNA samples of wild type (lane 1), *TbGT8* null mutant (lane 2), *TbGT8* add-back (lane 3), *TbGT8* (254ADA256) (lane 4) and *TbGT8* (254AAA256) (lane 5) were used as templates for RT-PCR with primers for *TbGT8*.  
C. As a control, the same experiment was carried out using primers for a constitutively expressed protein (Dol-P-Man Synthetase).  
D. The same RNA samples were used as template for PCR, to rule out the presence of DNA, using as a positive control *T. brucei* genomic DNA (lane 6).
The reduction in procyclin apparent molecular weight is due to changes in the GPI anchor

To further advance the glycosylation phenotyping, procyclins extracted from wild type and TbGT8 null cells were analysed by MALDI-Tof following aq. HF dephosphorylation and mild acid treatment (Acosta-Serrano et al., 1999; Mehlert et al., 1999). The aq. HF treatment removes the GPI anchor while the mild acid treatment cleaves at the labile Asp-Pro (DP) peptide bonds present on the N-terminal side of the EP-repeats of the EP-procyclins. The procedure yields deglycosylated polypeptide fragments that are detectable by MALDI-Tof and are characteristic of the different procyclin species present (Acosta-Serrano et al., 1999). In this case, both the wild type and TbGT8 null mutant cells can be seen to be expressing GPEET, EP1-1 and EP1-2 procyclins (Fig. S5A). A similar MALDI-Tof analysis was performed on procyclins treated with aq. HF alone. In this case, the EP procyclins retain their N-glycosylated N-terminal domains. The ions detected were those characteristic for EP procyclins containing wild-type Man5GlcNAc2 N-linked glycans (Fig. S5B) (Treumann et al., 1997; Acosta-Serrano et al., 1999). Taken together, these analyses show that the reductions in the apparent molecular weights of the TbGT8 null mutant procyclins are not due to changes in procyclin gene expression or N-glycosylation, but are most likely due to changes in the structure of the GPI anchor side-chains (Fig. 1, linkages 28–36).

Characterization of the changes in the procyclin GPI anchor side-chain in the TbGT8 null mutant

MALDI-Tof analysis of the wild-type intact procyclin preparation revealed two polydisperse components centred around 12.9 and 17.8 kDa (Fig. 4A), corresponding to GPEET and EP procyclins respectively (Treumann et al., 1997). These are replaced by components centred around 9.5 and 14 kDa in the TbGT8 null mutant (Fig. 4B), suggesting an average loss of around 3.6 kDa in both the GPEET and EP procyclin GPI anchor side-chains. To analyse this change further, wild type and TbGT8 null GPI anchor glycans were released from the procyclin peptide and the lyso-phosphatidic acid component of the PI lipid by aq. HF dephosphorylation (a procedure that also removes some sialic acid residues). The released GPI glycans were subsequently permethylated, a procedure that removes the fatty acid from the inositol ring, methylates all free hydroxyl groups and converts the amine group of the glucosamine residue to a positively charged quaternary amine (Baldwin, 2005; Mehlert and Ferguson, 2007). Analysis of the permethylated GPI glycan fractions by positive ion MALDI-Tof produced a relatively complex spectrum for the wild-type sample (Fig. 5A) discussed below, and a relatively simple spectrum for the TbGT8 null sample (Fig. 5B). The [M]+ molecular ions in this spectrum can be accounted for by assuming core structures of Hex6HexNAc5Ino and Hex6HexNAc5Ino substituted with between zero and eight HexHexNAc units with (minor species) and without (major species) sialic acid. Although not a quantitative technique, the distribution of glycan species in the TbGT8 null spectrum suggests that the average number of HexHexNAc units is around three to four while the most abundant species has two. The relative complexity of the MALDI-Tof spectrum of the wild-type permethylated GPI glycan (Fig. 5A), noted above, can be explained in terms of higher degrees of sialylation, due to greater branching. The same samples were also analysed by positive ion ESI-MS and triply charged [M+2Na]3+ ions of the major (HexHexNAc5Hex6HexN(Me3)5Ino species were observed at m/z 888.44. Fragmentation of these ions (MS3) produced an intense doubly charged product ion at m/z 1186.55, arising from elimination of the inositol residue and quaternary amine group (Fig. S6), which on further fragmentation (MS5) produced the product ion spectra shown in Fig. 6. The two spectra show that these isobaric permethylated glycans are different and are consistent with the wild-type (HexHexNAc5Hex6HexN(Me3)5Ino species having a branched side-chain and the TbGT8 null (HexHexNAc5Hex6HexN(Me3)5Ino species having a linear side-chain (see Fig. 6A and Fig. 6B respectively).

The same permethylated GPI glycan fractions were subjected to acid hydrolysis, reduction and acetylation to
provide partially methylated alditol acetate derivatives (PMAAs) that were analysed by GC-MS to provide a methylation linkage analysis (Fig. 7 and Table S2). Relative to the PMAAs derived from the mannose residues of the conserved Man3GlcN-Ino core (Fig. 1), the TbGT8 null GPI glycan is deficient in 3,6-disubstituted Gal residues, terminal Gal residues and 3-substituted Gal residues and relatively rich in 6-substituted Gal residues. Also, whereas the wild-type sample contains both 4- and 3-substituted GlcNAc residues, the TbGT8 null sample contains only 4-substituted GlcNAc. These methylation linkage data, together with the linear (TbGT8 null) versus branched (wild type) side-chain structures of the (HexHexNAc)2 Hex6HexN(Me3)Ino species (Fig. 6), strongly suggest that TbGT8 encodes a β1-3 GlcNac-transferase that acts on the GPI anchor side-chain to produce the first branch point (Fig. 1, linkage 30). Furthermore, the disappearance of 3-substituted GlcNac in the TbGT8 null mutant suggests that the repeats attached through GlcNacβ1-3 linkages are of the Galβ1-3GlcNac lacto-N-biose type while those attached through GlcNacβ1-6 linkages are of the Galβ1-4GlcNac N-acetyllactosamine type. Thus, the β1-4 and β1-3 Gal-transferases appear to be selective for the glycosidic linkage of their acceptor βGlcNac substrate. According to the structure depicted in Fig. 1 for the procyclic-form GPI glycan (A. Acosta-Serrano, J.M. Richardson and M.A. Ferguson, unpubl. data), the loss of linkage 30 would lead to the qualitative changes in methylation linkage analysis described above. However, the significant increase in the yield of 6-substituted Gal residues in the TbGT8 null mutant structure further suggests that deletion of the 3-branch may lead to a partially compensatory increase in the length of the 6-branch.

TbGT8 also influences complex N-glycan processing in bloodstream-form T. brucei

In order to assess the glycosylation phenotype of the bloodstream-form TbGT8 null mutant, we first analysed soluble-form VSG221 (sVSG221) isolated from the TbGT8 null mutant and its parental wild-type cell line. The VSG coat is released in a soluble form upon cell lysis by the action of endogenous GPI-specific phospholipase C (Cardoso de Almeida and Turner, 1983; Ferguson et al., 1985). The mature wild-type sVSG221 glycoprotein contains two N-glycosylation sites, one at Asn296 occupied by small complex structures and another at Asn428 occupied by oligomannose structures (Zamze et al., 1991), and a highly galactosylated GPI anchor glycan (Mehlert et al., 1998b). Thus, we use the glycosylation status of sVSG221 as a convenient reporter to assess any effects that a mutation has on the formation of glycosidic linkages 1–14 and 18–27 (Fig. 1) (Jones et al., 2005; Urbaniak et al., 2006). Analysis of

Fig. 5. MALDI-Tof mass spectra of permethylated GPI glycans isolated from wild type and TbGT8 null mutant trypanosomes. Permethylated glycans of the GPI anchors of wild type (A) and TbGT8 null mutant (B) were analysed by positive MALDI-Tof MS after co-crystallization with a 2,5-dihydroxybenzoic acid matrix. A series of GPI glycans with several Hex-HexNAc repeats are indicated. Open circles show a similar series minus a terminal Hex residue, presumably attached to the mannose core (solid arrows indicate 204 Da mass differences, equivalent to one methylated hexose residue). Asterisks indicate molecules with sialic acid residues (up to three sialic acid residues can be detected in the wild-type GPI anchors; dashed arrows indicate 361 Da mass differences, equivalent to methylated sialic acid residues).
the glycoform pattern of sVSG221 from wild type and TbGT8 null cells by electrospray-mass spectrometry (ES-MS) did not reveal any alterations in its glycosylation pattern (Fig. S7 and Table S3). As the high-molecular-weight branched poly LacNAc-containing N-glycans of bloodstream-form \textit{T. brucei} are attached to glycoproteins other than VSG (Atrih \textit{et al.}, 2005), the next analysis performed was SDS/urea extraction of total glycoproteins from trypanosome ghosts (that are depleted of VSG) and Western blotting with lectins. Mutations that affect the formation of complex and poly LacNac structures (Fig. 1, linkages 12–17) would be expected to alter the intensity and/or apparent molecular weight pattern of one or more of the blots.

A TbGT8 conditional null mutant (cKO) was prepared by introduction of a tetracycline-inducible ectopic copy of TbGT8 into the \textit{rDNA} locus using pLew100 (Wirtz \textit{et al.}, 1999). SDS/urea extracts of trypanosome ghosts from wild type and cKO cells before and after tetracycline induction were analysed by tomato lectin (TL), ricin and wheat germ agglutinin (WGA) blotting (Fig. 8). The wild type and induced cKO patterns were similar to each other and to previous TL and ricin blots (Urbaniak \textit{et al.}, 2006; Manthri \textit{et al.}, 2008; Stokes \textit{et al.}, 2008). However, in the absence of tetracycline, and thus TbGT8 activity, the binding of all three lectins was significantly reduced. The primary ligands for TL are structures containing \(\beta1-3\)-linked linear poly LacNac repeats (Merkle and Cummings, 1987) whereas for ricin they are structures bearing terminal Gal residues (Baenziger and Fiete, 1979). We interpret these results in terms of a reduction in UDP-GlcNac : \(\beta1\)Gal GlcNac transferase activity required for \(\beta1-3\)-linked poly LacNac synthesis in bloodstream-form \textit{T. brucei} (i.e. linkage 16 in Fig. 1). This would also reduce the branching of the large poly LacNac chains, consistent with the reduction in ricin binding. However, rationalizing the effect of the mutation on WGA binding is more difficult as the binding properties of this lectin are extremely complex (Muraki \textit{et al.}, 2002) and this will have to await detailed structural analysis of the high-molecular weight N-glycans from the bloodstream-form mutant.

Fig. 6. MS3 product ion spectra of permethylated GPI-glycans show the presence of branched and a linear side-chain in the wild type and TbGT8 null mutant GPI anchors respectively. Permethylated glycans of the GPI anchors of wild type (A) and TbGT8 null mutant (B) were analysed by positive ESI-MS. Triply charged \([M + 2Na]^{3+}\) ions observed at \(m/z\) 888.44 were fragmented (MS3) to generate the product ion spectra shown in the figure. Assignments of the major product ions are indicated on the inset diagrams. Some product ions appear as both singly and doubly charged ions, as indicated. The product ions in (B) at \(m/z\) 1664 and 708 and at \(m/z\) 1437 and 935 are diagnostic of the linear side-chain arrangement shown. The key is the same as for Fig. 8.
Discussion

As described earlier, a survey of the known glycoprotein N-linked and GPI glycan structures shows that the parasite makes a minimum of 38 distinct glycosidic bonds (Fig. 1 and Table S1). Further assuming that: (i) it is likely that there will be some redundancy in GT function and (ii) additional glycoconjugate structures remain to be discovered [e.g. O- and/or phosphate-linked glycans, as found in other trypanosomatids (Fergusson, 1999; Guha-Niyogi et al., 2001; Mendonca-Prevìato et al., 2005], the final figure is likely to be substantially higher. Indeed, it has been noted that typically about 1% of the protein coding genes in a eukaryotic genome are dedicated to the task of glycosidic bond synthesis (Coutinho et al., 2003). The 26-megabase T. brucei genome contains 9068 predicted genes, including approximately 900 pseudogenes (Berri man et al., 2005). According to this analysis, the number of GTs in T. brucei genome is likely to be around 80. Whatever the final count turns out to be, to date only five GTs have been assigned functions experimentally in T. brucei and only a further nine have been assigned putative functions based on BLAST searches (Table S1). The iscanps analysis described here, initiated with the human β3GnT-5 amino acid sequence, provided a further 21 candidate UDP sugar-dependent GT genes, of which one, TbGT8, has now been assigned the function as a GPI side-chain UDP-GlcNAc : βGal β1-3 GlcNAc-transferase that can also affect, in an as yet undefined way, N-linked poly LacNAc chain synthesis in the bloodstream form of the parasite. This apparent dual specificity of TbGT8 is unexpected. Clarification must await detailed analysis of the change(s) in N-linked poly LacNAc chain structure in the bloodstream-form mutant. Methodology to address such changes is currently under development, but is hampered by the inherent resistance of the predominant Galβ1-6GlcNAc inter-LacNAc repeat linkage in this material to endo-β-galactosidase (Atri h et al., 2005).

Another way to address this issue, and indeed to provide unambiguous evidence that TbGT8 is a GT and not simply required for the activity of another gene product, would be to express TbGT8 as a recombinant protein and demonstrate enzymatic activity. We tried to express the full-length enzyme, as well as truncated forms without the transmembrane domain, with hexahistidine or glutathione S-transferase tags at the N- and C-termini in both E. coli and Leishmania tarentole. However, apart from one N-terminal GST-tagged TbGT8 (minus transmembrane domain) construct, which produced low levels of soluble protein complexed to chaperones, we were unable to produce recombinant protein. We tried to assay the chaperone-complexed protein using UDP-[^3H]GlcNAc

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd, Molecular Microbiology, 71, 478–491
lacto-N-neotetraose oligosaccharides as acceptors, using Dionex high-pH anion exchange chromatography with pulsed amperometric detection to look for products. 
Unfortunately, no activity was detected in any assay. However, the presence of the chaperones suggested that the recombinant TbGT8 may have been at least partially denatured in this preparation. Despite this failure, we were able to provide supporting data that TbGT8 encodes an active GT. Thus, mutation of the typical GT DXD motif (in this case DDD) to ADA or AAA abrogated reversion of phenotype in add-back experiments.

Despite the fact that many GPI anchors carry side-chains (Ferguson et al., 1999), this is only the second GPI-modifying GT to have been identified. The other is the Smp3 gene that adds the fourth αMan residue to the conserved GPI core in the ER of yeast and mammals (Grimme et al., 2001; Taron et al., 2004). All the 21 candidate GT genes from the iscanps analysis appear to be similar to the mammalian β3-GT group of genes (Narimatsu, 2006). That group is characterized by an N-terminal transmembrane domain followed by three conserved motifs: (I/1)RXWXG (F/Y)(V/L/M)XXDD and (E/D)D(A/V)(Y/F)XGX(C/S). The comparable motifs in the T. brucei genes are slightly different: WG, Y(I,V,F)XXDDD and ED(A/V/I/L/M)(M I-1)X(G/A) (Fig. S1). It is conceivable that some of these genes might encode the β1-4 and β1-6 GTs needed to synthesize the Galβ1-4GlcNAc lactosamine units and GlcNAcβ1-6Gal interlactosamine linkages that are abundant in both life cycle stages (Fig. 1). BLAST searches with mammalian β4Gal-T sequences and β6GnT sequences fail to return any convincing candidates. Similarly, β3Gal-T and α3Gal-T sequences fail to return any convincing candidates. Thus, there are either additional classes of currently unidentifiable UDP-Gal/GlcNAc-dependent GT genes to be found in the genome and/or this organism has adapted the β1-3 GT family to catalyse a variety of different linkages. Additional gene replacement and glycosylation phenotyping, like those described here, should clarify this issue.

**Experimental procedures**

*Cultivation of trypanosomes*

Bloodstream-form *T. brucei* genetically modified to express T7 polymerase and the tetracycline repressor protein were cultivated in HMI-9 medium containing 2.5 μg ml⁻¹ G418 at 37°C in a 5% CO₂ incubator as described in Wirtz et al. (1999). Procyclic-form *T. brucei* cells were grown in SDM-79 medium (Brun and Schonenberger, 1979). Differentiation to procyclic form was initiated by incubating 1.5 × 10⁷ log-phase bloodstream-form parasites at 27°C in 5 ml of SDM-79 media supplemented with 15% fetal calf serum and 3 mM cis-aconitate (Brun and Schonenberger, 1981), after washing in 5 ml of SDM-79 medium. The cell density of the differentiated cells growth in absence of tetracycline have altered poly LacNAc N-glycosylation.

as a donor and both procyclin purified from the mutant (before and after neuraminidase treatment) and asialofetuin as acceptors, as well as using unlabelled UDP-GlcNAc as a donor and synthetic lacto-N-tetraose and

**Fig. 8.** Bloodstream-form *T. brucei* TbGT8 conditional null mutant cells growth in absence of tetracycline have altered poly LacNAc N-glycosylation.

A. Blot of total glycoprotein extracts of wild type (lane 1), TbGT8 conditional null mutant minus Tet (lane 2) and TbGT8 conditional null mutant plus Tet (lane 3), incubated with TL without (upper panel) and with inhibitors (specificity control, middle panel), and stained with Ponceau red (loading control, lower panel).

B. Blot of total glycoprotein extracts of wild type (lane 1), TbGT8 conditional null mutant minus Tet (lane 2) and TbGT8 conditional null mutant plus Tet (lane 3), incubated with ricin without (upper panel) and with inhibitors (specificity control, middle panel), and stained with Ponceau red (loading control, lower panel).

C. Blot of total glycoprotein extracts of wild type (lane 1), TbGT8 conditional null mutant minus Tet (lane 2) and TbGT8 conditional null mutant plus Tet (lane 3), incubated with WGA without (upper panel) and with inhibitors (specificity control, middle panel), and stained with Ponceau red (loading control, lower panel).
procyclic-form parasites was maintained, by serial subculture, between 1 and $6 \times 10^6$ cells ml$^{-1}$.

**DNA isolation and manipulation**

Plasmid DNA was purified from *Escherichia coli* (DH5α) using Qiagen Miniprep or Maxiprep kits, as appropriate. Gel extraction was performed using Qiagith quick kits. Custom oligonucleotides were obtained from Thermo Hybaid or the Dundee University oligonucleotide facility. *T. brucei* genomic DNA was isolated from $-2 \times 10^8$ bloodstream-form cells or from $1 \times 10^9$ procyclic cells using DNAzol (Helena Biosciences).

**Generation of constructs**

The 394 bp 5' and 398 bp 3'-UTR sequences next to the Tb10.389.1450 ORF were PCR-amplified from genomic DNA using *Pfu* with 5'-atgGGGCGGgcgaagagtgcgcattcgctg-'3' and 5'-ggaagtgtgcaagagggatgaggagac-3' and 5'-gccg atcggggagatgagcgcgctg-3' and 5'-gatGCGGCCGCGaga cgacaaggtgcgcgccaaac-3' as forward and reverse primers respectively. The two PCR products were used together in a further PCR reaction to yield a product containing the 5'-UTR linked to the 3'-UTR by a short HindIII and BamHI cloning site (underlined) and NotI restriction sites at each end (capital letters). The PCR product was cloned into the NotI site of the pGEM-5Zf(+) vector (Promega) and the HYG and PAC drug-resistance genes were introduced into the targeting vector using the HindIII and BamHI cloning sites. For re-expression of Tb10.389.1450 the ORF was PCR-amplified from genomic DNA using *Pfu* with 5'-ggaagcttgcaagagggatgaggagac-3' and 5'-GCCGGCGGCGaga cgacaaggtgcgcgccaaac-3' and cloned into pGL927 (Helms *et al.*, 2006) (a gift from Jeremy Mottram, WCMP, University of Glasgow, UK), cut with EcoRV, to generate pGLGT8. This plasmid was then digested with NotI, and purified for transfection.

**Transformation of bloodstream-form *T. brucei***

Constructs for gene replacement and ectopic expression were purified using the Qiagen Maxiprep kit, digested with NotI to linearize, precipitated and washed twice with 70% ethanol and re-dissolved in sterile water. The linearized DNA was electroporated into *T. brucei* bloodstream cells (strain 427, variant 221) that were stably transformed to express T7 RNA polymerase and the tetracycline repressor protein under the HindIII and BamHI cloning sites. For re-expression of Tb10.389.1450 the ORF was PCR-amplified from genomic DNA using *Pfu* with 5'-GCCGGCGGCGaga cgacaaggtgcgcgccaaac-3' and cloned into pGL927 (Helms *et al.*, 2006) (a gift from Jeremy Mottram, WCMP, University of Glasgow, UK), cut with EcoRV, to generate pGLGT8. This plasmid was then digested with NotI, and purified for transfection.

**Semi-quantitative RT-PCR and quantitative RT-PCR**

In order to assess the amount of Tb10.389.1450 mRNA, RT-PCR reactions were performed using AccessQuick RT-PCR System (Promega). As template 4, 16 and 64 ng of total RNA from bloodstream and procyclic forms of the parasite were used. The reaction was carried out in a GeneAmp PCR System 2700 from Applied Biosystems. Tb10.389.1450 mRNA 366 bp fragment was amplified with primers 5'-CGCAATCGGGGTCTTTCACCGCGTTCT-3' and 5'-GTC CTCTCAACCTTCTGAAGCGCGAAACCC-3'. As a control of a similar mRNA level in both life stages of the parasite, primers 5'-AATGGATGCGGACCTTCCAGCACCACCC-3' and 5'-TAGACCGTGAGCCCGTGCTCATAC-3' amplifying a 449 bp product of dolichyl-phosphate-mannose synthetase (Tb10.70.2610) were used. As a procylic-form marker control, a 433 bp product of CAP5.5 protein (Tb04.1D20.740) (Hertz-Fowler *et al.*, 2001) was amplified with 5'-AAGACGCTGATGCGGAACCAACCC-3' and 5'-CCACCTGGAACCTC TGCCAATTCATCC-3' primers. Finally, as a bloodstream-form mRNA marker, the primers 5'-GAAAAACAGCTCGGCCCTTGACCCTGCAC-3' and 5'-GCTGAACGTACCGGAGGCCAC ATC-3' were used to amplify a 451 bp product from AcP115 mRNA (Tb05.30F7.630) (Bakalara *et al.*, 2000). The cycling parameters used were 48°C for 45 min, 94°C for 2 min, 25 cycles of 94°C for 40 s, 60°C for 1 min 15 s, and 68°C for 1 min followed by a final 7 min extension time at 68°C.

Quantitative RT-PCR to quantify the mRNA expression level of TbGT8 in both life cycle stages was done using Sybgreen master mix (Bio-Rad) and, as template, cDNA generated from procyclic-form and bloodstream-form RNA. The reaction was performed using the iCycler Thermal Cycler from Bio-Rad, using the following programme: 95°C for 3 min, 40 cycles of 95°C for 20 s and 57°C for 45 s, 95°C for 1 min, 55°C for 1 min, followed by 80 cycles of 55°C for 10 s. The values were normalized to the expression of dolichyl-phosphate-mannose synthetase (Tb10.70.2610).

**Purification of procyclins**

Procyclins (both GPEET and EP forms) were purified from 10⁶ freeze dried trypanosomes by organic solvent extraction followed by octyl-Sepharose chromatography (Amersham Pharmacia Biotech) (Ferguson *et al.*, 1993; Treumann *et al.*, 1997; Mehler *et al.*, 1999). For some specific procedures, such as Western Blot, procyclins were extracted from batches of $-1 \times 10^8$ cells and the blot was performed by octyl-Sepharose chromatography.

© 2008 The Authors
**SDS-PAGE and Western blotting of procyclins**

Procyclin extracts, equivalent to $2 \times 10^7$ cells, were subjected to SDS-PAGE on 4–12% NuPAGE (Invitrogen) gels and transferred to polyvinylidene difluoride Hybond-P membranes (Amersham Biosciences) in a semi-dry transfer apparatus at 40 mA for 1 h. After blocking for 1 h with 3% BSA in PBS, the membranes were washed three times with PBS, incubated for 1 h with anti-EP-procyclin mouse monoclonal antibody 247 (Richardson et al., 1988) or antiphosphorylated-GPEET-procyclin mouse monoclonal antibody 5H3 (Richardson et al., 1986) diluted 1:2500 in PBS, washed three times with PBS, and incubated 1 h with anti-mouse conjugated to horseradish peroxidase diluted 1:5000 with the same buffer. The membranes were washed three times with PBS, and developed with ECL (Amersham Biosciences) according to the manufacturer’s instructions.

**Analysis by MALDI-ToF MS**

The 250 pmols of octyl-Sepharose-purified procyclins was dried and treated with 50 µl of ice-cold 50% aqueous hydro- gen fluoride for 24 h at 0°C to cleave the GPI anchor ethanalamine-phosphate bond. Some preparations were further treated with 50 µl of 40 mM trifluoroacetic acid, 100°C for 20 min to cleave Asp-Pro bonds and remove N-glycosylated N termini (Acosta-Serrano et al., 1999). The samples were dried and re-dissolved in 5 µl of 0.1% trifluoroacetic acid. Aliquots (1 µl) of each sample were mixed with 1 µl of 10 mg ml$^{-1}$ sinapinic acid in 50% acetonitrile, 0.1% trifluoroacetic acid and analysed by positive ion MALDI-ToF. Data collection was in linear mode on a Voyager-DE STR instrument. The accelerating voltage was 25 000 V, and grid voltage was set at 94% with an extraction time delay of 700 ns. Data were collected manually at 500 shots/spectrum with the laser intensity set at 2500. The analysis of the native procyclins (500 pmol per sample) was performed in negative mode using sinapinic acid as the matrix. For the analysis of the permethylated GPI glycans, the instrument was set in reflectron mode, positive ion and a matrix of 2,5-Dihydroxybenzoic acid was used.

**Permethylation and ES-MS of GPI glycans**

Samples were dried and permethylated by the sodium hydroxide method, as described in Ferguson (1994). The permethylated glycan samples were dissolved in 50 µl of 80% acetonitrile, and aliquots (5 µl) were dried and recovered in 80% acetonitrile, 0.5 mM sodium acetate before loading into nanotips (Micromass type F) for positive ion ES-MS, ES-MS$^2$ and ES-MS$^3$ on a LTQ Orbitrap XL mass spectrometer (Thermo Scientific). Source and capillary voltages were 0.63 kV and 48 V respectively, and the collision energy was 17–19%.

**Methylation linkage analysis by GC-MS**

The remainder of the permethylated glycan samples were subjected to acid hydrolysis, Na$^+$BH$_4$ reduction and acetylation (to yield PMAAs), and analysed by GC-MS as described in Ferguson (1994). The PMAAs were analysed on an Agilent 6890 N GC-MS system fitted with an HP-5 column.

**Small-scale sVSG isolation**

Soluble-form VSG was isolated from 100 ml cultures containing $2 \times 10^8$ bloodstream-form $T. brucei$. The cultures were chilled in ice water and centrifuged at 2500 g for 10 min. The pellet was washed twice in trypanosome dilution buffer (Cross, 1975) and transferred to a 1.5 ml Eppendorf tube. The pellet was re-suspended in 300 µl of lysis buffer (10 mM NaH$_2$PO$_4$ buffer, pH 8.0, containing 0.1 mM 1-chloro-3-tosylamido-7-amino-2-heptanone, 1 µg ml$^{-1}$ leupeptin and 1 µg ml$^{-1}$ aprotinin) pre-warmed to 37°C and incubated for 5 min at the same temperature. The sample was centrifuged at 14 000 g for 5 min, and the supernatant was applied to a 200 µl DE52 anion exchange column pre-equilibrated in lysis buffer. Fresh lysis buffer (800 µl without protease inhibitors) was applied in four stages, and the pooled column eluate was concentrated and dialyzed with water on a YM-10 spin concentrator (Microcon). The final sample of 50–100 µg sVSG221 was recovered in a volume of 100 µl of water.

**ES-MS analysis of intact VSG**

Samples of the sVSG preparations were diluted to $-0.07$ µg µl$^{-1}$ in 50% acetonitrile, 1% formic acid, loaded into nanotips (Micromass type F) and analysed by positive ion ES-MS on a QToF2 instrument (Micromass) with tip and cone voltages of 1 kV and 30 V respectively. Data were collected and processed using the maximum entropy algorithm of the MassLynx software.

**Lectin blotting of $T. brucei$ cell extracts**

Bloodstream-form $T. brucei$ cells were washed twice in TDB, solubilized with 2% SDS and 4 M urea, subjected to SDS-PAGE on 4–12% NuPAGE (Invitrogen) gels and transferred to nitrocellulose membranes (Amersham Biosciences) in a semi-dry transfer apparatus at 45 mA for 1 h. After blocking for 1 h with 3% BSA in PBS, the membranes were incubated for 1 h with biotin-conjugated ricin (Vector Laboratories) diluted 1/3000 in PBS, with or without inhibitors [10 mg ml$^{-1}$ D(+)-galactose and 10 mg ml$^{-1}$ α-lactose], washed three times with PBS, incubated for 1 h with Extravidin-HRP diluted 1/10 000 in PBS, washed three times with PBS and developed with ECL (Amersham Biosciences) according to the manufacturer’s instructions. In the case of WGA, blocked membranes were incubated for 1 h with WGA (Vector Laboratories) diluted 1/2000 in WGA buffer (50 mM Tris, 0.15 M NaCl, 1 mM EDTA, 0.05% NP40, 0.25% BSA, pH 7.4), with or without inhibitors (1/10 Chitin hydrolysate), washed three times with WGA buffer, incubated for 1 h with α-biotin antibody (diluted 1/10 000 in WGA buffer), washed three times with WGA buffer, incubated for 1 h with Extravidin-HRP diluted 1/10 000 in WGA buffer, washed three times with WGA buffer and developed with ECL (Amersham Biosciences) according to the manufacturer’s instructions. Finally, in the case of TL, blocked membranes were incubated for 1 h with biotin-conjugated TL (Vector Laboratories).

© 2008 The Authors

diluted 1/10 000 in TL buffer (50 mM Tris, 0.5 M NaCl, 0.05% NP40, 0.25% BSA, pH 7.4), with or without inhibitors (1/10 Chitin hydrolyosate), washed three times with TL buffer, incubated for 1 h with Extravidin-HRP diluted 1/10 000 in TL buffer, washed three times with TL buffer and developed with ECL (Amersham Biosciences) according to the manufacturer’s instructions.

Acknowledgements

This work was supported in part by a programme grant from the Medical Research Council to M.A.J.F. L.I. was supported in part by a Marie Curie Fellowship. We would like to thank Alvaro Acosta-Serrano, Frauke Graalfs, Joao Rodrigues and Isabelle Nett for helpful discussions, Douglas Lamont for assistance with the MS3 experiments and Frauke Graalfs for performing the quantitative RT-PCR experiment.

References


Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.