Deletion of the \textit{TbALG3} gene demonstrates site-specific N-glycosylation and N-glycan processing in \textit{Trypanosoma brucei}

Sujatha Manthri$^2$, M Lucia S Guthér$^2$, Luis Izquierdo$^2$, Alvaro Acosta-Serrano$^3$, and Michael A J Ferguson$^{1,2}$

$^2$The Division of Biological Chemistry and Drug Discovery, The Wellcome Trust Biocentre, College of Life Sciences, University of Dundee, Dundee DD1 5EH; and $^3$The Wellcome Centre for Molecular Parasitology, Glasgow Biomedical Research Centre, University of Glasgow, Glasgow G12 8TA, Scotland, UK

Received on January 5, 2008; accepted on February 3, 2008

We recently suggested a novel site-specific N-glycosylation mechanism in \textit{Trypanosoma brucei} whereby some protein N-glycosylation sites selectively receive Man$_9$GlcNAc$_2$ from Man$_5$GlcNAc$_2$-PP-Dol while others receive Man$_9$GlcNAc$_2$ from Man$_5$GlcNAc$_2$-PP-Dol. In this paper, we test this model by creating procyclic and bloodstream form null mutants of \textit{TbALG3}, the gene that encodes the \(\alpha\)-mannosyltransferase that converts Man$_9$GlcNAc$_2$-PP-Dol to Man$_9$GlcNAc$_2$-PP-Dol. The procyclic and bloodstream form \textit{TbALG3} null mutants grow with normal kinetics, remain infectious to mice and tsetse flies, respectively, and have normal morphology. However, both forms display aberrant N-glycosylation of their major surface glycoproteins, procyclin, and variant surface glycoprotein, respectively. Specifically, procyclin and variant surface glycoprotein N-glycosylation sites that are modified with Man$_9$GlcNAc$_2$ and processed no further than Man$_5$GlcNAc$_2$ in the wild type are glycosylated less efficiently but processed to complex structures in the mutant. These data confirm our model and refine it by demonstrating that the biantennary glycan transferred from Man$_5$GlcNAc$_2$-PP-Dol is the only route to complex N-glycans in \textit{T. brucei} and that Man$_5$GlcNAc$_2$-PP-Dol is strictly a precursor for oligomannose structures. The origins of site-specific Man$_5$GlcNAc$_2$ or Man$_9$GlcNAc$_2$ transfer are discussed and an updated model of N-glycosylation in \textit{T. brucei} is presented.

\textbf{Keywords:} ALG3/mannosyltransferase/N-glycosylation/\textit{Trypanosoma brucei}/variant surface glycoprotein

\section*{Introduction}

The trypanosomatid \textit{Trypanosoma brucei} is a parasitic protozoan organism that causes nagana in cattle and human African sleeping sickness. The organism undergoes a complex life cycle, involving major biochemical and morphological changes, between its mammalian host and tsetse fly vector. These changes include complete remodeling of the major cell surface coat molecules.

The tsetse midgut-dwelling procyclic form of \textit{T. brucei} has a surface coat of \(3 \times 10^6\) polyanionic, rod-like, procyclin glycoproteins (Mowatt and Clayton 1987; Roditi et al. 1987, 1989; Richardson et al. 1988; Treumann et al. 1997) as well as other unidentified glycoproteins (Guthér et al. 2006). In \textit{T. brucei} strain 427, used in this study, the parasites contain (per diploid genome) two copies of the \textit{GPEET1} gene, encoding a procyclin with 6 Gly-Pro-Glu-Thr repeats; one copy each of the \textit{EP1-1} and \textit{EP1-2} genes, encoding EP1 procyclins with 30 and 25 Glu-Pro repeats, respectively; two copies of the \textit{EP2-1} gene, encoding EP2 procyclin with 25 Glu-Pro repeats; and two copies of the \textit{EP3-1} gene, encoding EP3 procyclin with 22 Glu-Pro repeats (Acosta-Serrano et al. 1999; Roditi and Clayton 1999). The EP1 and EP3 procyclins contain a single N-glycosylation site, at the N-terminal side of the Glu-Pro repeat domain, occupied exclusively by a conventional Endo-H-sensitive triantennary Man$_9$GlcNAc$_2$ oligosaccharide (Treumann et al. 1997). Neither EP2 nor GPEET procyclin is N-glycosylated but GPEET1 procyclin is phosphorylated on six out of seven Thr residues (Butikofer et al. 1999; Mehliert et al. 1999). GPEET and EP procyclins contain similar glycosylphosphatidylinositol (GPI) membrane anchors, based on the ubiquitous ethanolamine-P-$\alpha$Man$_1$-2Man$_1$-$\alpha$Man$_1$-4GlcN$_1$-6PI core (Ferguson 1999), where in this case, the phosphatidylinositol (PI) lipid is a 2-O-acetyl-$\alpha$-phosphatidylglycerol (Zamze et al. 1999). The bloodstream form of \textit{T. brucei} has a surface coat of \(5 \times 10^6\) variant surface glycoprotein (VSG) homodimers (Mehliert, Richardson, et al. 1998). The VSG coat serves as a physical barrier to components of the host complement system and undergoes antigenic variation (Pays et al. 2004; Taylor and Rudenko 2006). There are many VSG genes and each encodes a GPI-anchored glycoprotein with one to three N-glycosylation sites (Mehliert, Richardson, et al. 1998). The cell line used in this study expresses VSG variant 221 (also known as MiTat1.2). VSG221 has a GPI anchor with the same core as the procyclin, with six terminal sialic acid residues (Mehliert et al. 1999). The bloodstream form of \textit{T. brucei} has a surface coat of \(5 \times 10^6\) variant surface glycoprotein (VSG) homodimers (Mehliert, Richardson, et al. 1998). The VSG coat serves as a physical barrier to components of the host complement system and undergoes antigenic variation (Pays et al. 2004; Taylor and Rudenko 2006). There are many VSG genes and each encodes a GPI-anchored glycoprotein with one to three N-glycosylation sites (Mehliert, Richardson, et al. 1998). The cell line used in this study expresses VSG variant 221 (also known as MiTat1.2). VSG221 has a GPI anchor with the same core as the procyclin but with a dimyristoyl-PI component and carbohydrate side chains of between two and six Gal residues (Mehliert, Zitzmann, et al. 1998). VSG221 has two N-glycosylation sites: the Asn428 site, five residues from the GPI attachment site, is occupied mostly by Endo-H-sensitive oligomannose structures (Man$_5$GlcNAc$_2$), whilst the Asn263 site is occupied by small Endo-H-resistant biantennary structures ranging from Man$_3$GlcNAc$_2$ to GalGlcNacMan$_3$GlcNAc$_2$ (Zamze et al. 1991).
Protein N-glycosylation in eukaryotes serves a wide variety of functions including signaling through interaction with lectins, protein stabilization, protease resistance, endocytotic sorting functions, and protein folding (Varki 1993; Rudd and Dwek 1997; Helenius and Aebi 2004). In eukaryotes, a precursor for N-glycosylation is built up in the endoplasmic reticulum (ER) on the lipid carrier dolichol pyrophosphate (Dol-PP) (Burda and Parodi 1997; Helenius and Aebi 2004). Processing of the precursor structure by glycosidase and glycosyltransferase enzymes within the ER and Golgi apparatus generates the final set of mature structures (Kornfeld R and Kornfeld S 1985; Schachter 2000).

In most eukaryotes, the mature precursor used by OST is Glc3Man9GlcNAc2-PP-Dol. However, genomic and experimental comparisons have shown that some lower eukaryotes do not possess all the ALG genes needed to make Glc3Man9GlcNAc2-PP-Dol and that they transfer smaller glycans to protein (Parodi 1993; Samuelson et al. 2005). Differences in the compositions and donor specificities of eukaryotic OST complexes, which usually contain eight different subunits, have also been noted (Kelleher and Gilmore 2006; Kelleher et al. 2007).

Seminal work by Parodi and colleagues on several trypanosomatid parasites (excluding T. brucei) showed that protein N-glycosylation in these organisms is aberrant (reviewed in Parodi 1993). None of these organisms make Dol-P-Glc and so fail to make glycosylated Dol-P-oligosaccharide precursors. The mature Dol-PP-oligosaccharide species used for transfer to protein vary according to species. For example, Trypanosoma conhorini, Trypanosoma diontisi, Leptomonas samueli, Herpetomonas manuelpsessoui, and Herpetomonas muscarum utilize triantennary Man5GlcNAc2-PP-Dol; Crithidia fasciculata, Crithidia Harmosa, and Leishmania enriettii utilize biantennary Man9GlcNAc2-PP-Dol; Leishmania mexicana, Leishmania adleri, and Blastocystis cuthicus utilize biantennary Man9GlcNAc2-PP-Dol (Parodi et al. 1981; Parodi and Quesada-Allue 1982; Prevato et al. 1986; de la Canal and Parodi 1987); and Trypanosoma cruzi, the causative agent of Chagas’ disease in the Americas, utilizes Man9GlcNAc2-PP-Dol during most of its life cycle but uses both Man9GlcNAc2-PP-Dol and Man7GlcNAc2-PP-Dol in its bloodstream trypomastigote stage (Doyle et al. 1986).

In a recent study, we showed that wild-type bloodstream form T. brucei utilizes both Man9GlcNAc2-PP-Dol and Man7GlcNAc2-PP-Dol to glycosylate the major surface coat glycoprotein VSG221 and, uniquely, does so in a site-specific manner (Jones et al. 2005). Thus, whereas Man9GlcNAc2-PP-Dol (Parodi et al. 1981; Parodi and Quesada-Allue 1982; Prevato et al. 1986; de la Canal and Parodi 1987); and Trypanosoma cruzi, the causative agent of Chagas’ disease in the Americas, utilizes Man9GlcNAc2-PP-Dol during most of its life cycle but uses both Man9GlcNAc2-PP-Dol and Man7GlcNAc2-PP-Dol in its bloodstream trypomastigote stage (Doyle et al. 1986).

In a recent study, we showed that wild-type bloodstream form T. brucei utilizes both Man9GlcNAc2-PP-Dol and Man7GlcNAc2-PP-Dol to glycosylate the major surface coat glycoprotein VSG221 and, uniquely, does so in a site-specific manner (Jones et al. 2005). Thus, whereas Man9GlcNAc2-PP-Dol is used to glycosylate Asn428, Man5GlcNAc2-PP-Dol is used to glycosylate Asn263. In addition, characterization of VSG221 made by an α-glucosidase II null mutant showed that the Asn263-linked Man9GlcNAc2 glycan serves exclusively on VSG221 as the acceptor for the parasite’s unfolded glycoprotein glucosyltransferase (UGGT) (Jones et al. 2005). We further speculated that Man9GlcNAc2 is the obligate precursor of all Endo-H-resistant (i.e., small biantennary oligomannose and complex) structures whereas Man7GlcNAc2 is the obligate precursor of all conventional Endo-H-sensitive (i.e., Man5–9GlcNAc2) oligomannose structures.

To test and refine our model of site-specific N-glycosylation and N-glycan processing, we constructed bloodstream and procyclic form T. brucei null mutants of the TbALG3 gene and observed the effects of limiting the parasite to solely a Man9GlcNAc2-PP-Dol lipid-linked precursor.

**Results**

**Identification and cloning of the TbALG3 gene**

A BLASTp (Altschul et al. 1990) search of the geneDB T. brucei predicted proteins database with the predicted amino acid sequence of the yeast ALG3 gene (accession no. NP 009471) yielded a putative T. brucei Dol-P-Man-dependent α-1,3-mannosyltransferase sequence encoded by gene Tb10.70.0260, referred to from now on as TbALG3. The genome strain of T. brucei (strain 927) is different from that used in this study (strain 427). The TbALG3 gene was, therefore, cloned and sequenced from strain 427. The consensus sequence (accession no. AM850907) has four single base polymorphisms compared to the strain 927 database sequence, resulting in only a single amino acid polymorphism (Y in place of C at position 278). Southern blot analysis using a TbALG3 ORF probe revealed a single fragment after restriction enzyme digestion by HindIII, SacII, Msci, PsiI, and Xhol whereas BsrGI yielded two fragments as it cuts the gene in the center of the gene (supplementary Figure S1). These data indicate that the TbALG3 gene is present as a single copy per haploid genome.

The predicted 46.4 kDa TbALG3 multispanning membrane protein has 26.2%, 33.4%, 33.7%, and 34.7% identity with the Saccharomyces cerevisiae, Mus musculus, Drosophila melanogaster, and Arabidopsis thaliana ALG3 sequences, respectively. The TbALG3 protein sequence conforms to that of a member of the α,3-mannosyltransferase superfAMILY composed of dolichol cycle ALG3 α3-mannosyltransferases and GPI biosynthesis PIG-M α4-mannosyltransferase sequences (Oriol et al. 2002). Based on conserved amino acid motifs, TbALG3 can further be distinguished as an ALG3-type α3-mannosyltransferase (Oriol et al. 2002) (Figure 1). This confirms that the sequence patterns described in Oriol et al. (2002) for discriminating ALG3 genes from related sequences can be extended to this ancient and highly divergent eukaryote. The TbALG3 sequence contains, like those of S. cerevisiae, D. melanogaster, and M. musculus, a putative C-terminal ER retention signal.

Genes encoding proteins similar to TbALG3 were also found in the T. cruzi (gene Tc00.1047053510187.404) and Leishmania major (gene LmjF36.2040) databases with 48.5% and 44.5% predicted amino acid identity to TbALG3, respectively. Interestingly, the L. major sequence has 8- and 16- amino-acid inserts that are absent from the T. brucei and T. cruzi sequences.

**Creation of TbALG3 null mutants**

The two TbALG3 alleles were sequentially replaced by homologous recombination (Wirtz et al. 1999) with genes encoding puromycin acetyltransferase (PAC) followed by hygromycin phosphotransferase (HYG), for bloodstream form T. brucei, or PAC followed by blasticidin deaminase (BSD), for procyclic form T. brucei. Drug resistant clones were selected and characterized by Southern blotting of HindIII- and SacII-digested genomic DNA with a TbALG3 ORF probe followed by Southern blotting of HindIII- and SacII-digested genomic DNA with a TbALG3 ORF probe followed
Novel aspects of N-glycosylation in Trypanosoma brucei

Fig. 1. Predicted amino acid sequence of TbALG3 and alignment with related sequences. Predicted amino acid sequences of ALG3 genes from Trypanosoma brucei (T. b), Arabidopsis thaliana (A. t), Saccharomyces cerevisiae (S. c), Drosophila melanogasta (D. m), and Mus musculus (M. m) were aligned using ClustalW. Residues in black boxes are conserved between all sequences while those in gray boxes indicate conservative changes between the sequences. Residues indicated by # and * are those that, according to Oriol et al. (2002), are common to the α3,4-mannosyltransferase superfamily and that define the ALG3 α3-mannosyltransferase family, respectively. The C-terminal sequences in bold italics are putative ER-retention signals by a β-tubulin probe, to ensure equal loading. A representative blot for one of the bloodstream form TbALG3 null mutants (∆TbALG3::PAC/∆TbALG3::HYG) is shown in Figure 2.

The TbALG3 gene is nonessential to bloodstream and procyclic form T. brucei.

The in vitro growth rates of the TbALG3 null mutants of both life-cycle stages were indistinguishable from their parental cell lines (supplementary Figure S2). Similarly, scanning electron micrographs of the TbALG3 null mutants of both life-cycle stages were indistinguishable from their parental cell lines (supplementary Figure S3). Furthermore, bloodstream form TbALG3 null mutants were infective to mice and procyclic form TbALG3 null mutants were infective to tsetse flies, although the infectivity may have been marginally reduced (Figure 3). From these data we conclude that TbALG3 is a nonessential gene to the procyclic form and the disease-causing bloodstream form of T. brucei.

TbALG3 encodes a dolichol cycle α3-mannosyltransferase

Previous studies on dolichol-linked oligosaccharides and protein N-glycosylation in T. brucei have noted that although bloodstream form T. brucei make and utilize Man9GlcNAc2-PP-Dol (Zamze et al. 1991; Jones et al. 2005), the steady-state levels of species larger than Man5GlcNAc2-PP-Dol are extremely low in this life-cycle stage (Low et al. 1991). On the other hand, the Man6GlcNAc2-PP-Dol to Man9GlcNAc2-PP-Dol species are easier to observe in GDP-[3H]Man labeling experiments using the procyclic form cell-free system (Low et al. 1991; Leal et al. 2004). For this reason, we utilized procyclic form cell-free systems to assess the metabolic lesion caused by deletion of both TbALG3 alleles.
Washed membranes (cell-free systems) prepared from wild-type and TbALG3 null mutant procyclic form trypanosomes were incubated UDP-GlcNAc and GDP-[3H]Man, which labels both GPI and dolichol-PP-linked oligosaccharide precursors and their biosynthetic intermediates (Acosta-Serrano et al. 2004). To selectively analyze the dolichol-PP-linked oligosaccharide species, the labeled glycolipids were treated with mild acid and partitioned between butan-1-ol and water. The mild acid-labile pyrophosphate bonds of the Dol-PP-oligosaccharides result in the recovery of their radiolabeled free oligosaccharides in the aqueous phase, whereas the mild acid stable radiolabeled GPI species partition into the butan-1-ol phase. Following reduction of the radiolabeled free oligosaccharides with sodium borohydride and desalting, the resulting oligosaccharide alditols were analyzed by high-performance thin layer chromatography (HPTLC) and fluorography alongside authentic sodium borotritiide-reduced oligosaccharide alditol standards (Figure 4A). The fluorograph clearly shows that, as expected, the TbALG3 null mutant cell-free system is unable to produce oligosaccharides larger than Man5GlcNAc2. The radiolabeled free oligosaccharide alditols were further analyzed before and after digestion with Aspergillus saitoi Manα1–2Man-specific α-mannosidase (ASAM) (Figure 4B). The digestion patterns are consistent with the Man5GlcNAc2 structure being Manα1–6(Manα1–2Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc, as summarized in Figure 4C.

Together, these data provide clear evidence that the largest Dol-PP-linked oligosaccharide made by the TbALG3 null mutant is a biantennary Manα1–3mannosyltransferase responsible for the conversion of Manα1–6(Manα1–2Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc-PP-Dol (Manα5GlcNAc2-PP-Dol) to Manα1–3Manα1–6(Manα1–2Manα1–2Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc-PP-Dol (Manα5GlcNAc2-PP-Dol).

Procyclin N-glycosylation is affected in TbALG3 null mutant procyclic form T. brucei

Procyclins were extracted from wild-type and TbALG3 null trypanosomes by differential solvent extraction and octyl-Sepharose hydrophobic interaction chromatography. The proteins were first analyzed by SDS–PAGE and periodate-Schiff staining for carbohydrate. There was a small but perceptible downward shift in the apparent molecular weight range of the intact procyclins in the TbALG3 null mutants compared to those in the wild type (Figure 5, lanes 1 and 3), whereas the two samples appeared similar after peptide N-glycosidase-F (PNGaseF) treatment (Figure 5, lanes 2 and 4).

To gain a more detailed insight, procyclins from wild-type and TbALG3 null mutant cells were treated in three different ways, as described in Acosta-Serrano et al. (1999) and Leal et al. (2004), and analyzed by negative-ion mode MALDI-TOF. The three treatments were as follows: (i) dephosphorylation with cold aqueous HF alone (aq. HF), to remove their heterogeneous GPI anchors and reveal the masses of their full amino acid sequence plus N-glycans; (ii) dephosphorylation with cold aqueous HF followed by de-N-glycosylation with PNGaseF (aq. HF + PNGaseF), to reveal the masses of their full amino acid sequences; (iii) dephosphorylation with cold aqueous HF followed by mild acid hydrolysis of Asp-Pro peptide bonds with trifluoroacetic acid (aq. HF + TFA), to reveal the masses...
Fig. 4. Comparison of mild acid-released glycans from the Dol-PP-oligosaccharides of wild-type and TbALG3 null mutant cells. Panel A: radiolabeled oligosaccharides were released from GDP-[3H]mannose labeled cell free systems of wild-type (WT) and TbALG3null mutant (TbALG3−/−) cells, reduced with NaBH4, and analyzed by HPTLC and fluorography. The positions of authentic NaB[3H]4-reduced oligosaccharide standards (names in bold italics) are shown on the left together with presumed descriptors for the remaining bands. The upper band is [3H]mannitol, derived from the release of [3H]Man from Dol-P-[3H]Man and its subsequent reduction to [3H]mannitol. Panel B: the NaB[3H]4-reduced Man5GlcNAc2 standard (lanes 1 and 2), and the wild type (lanes 3 and 4) and TbALG3 null mutant (lanes 5 and 6) reduced oligosaccharides were treated with (+) and without (−) A. saitoi α1-2-specific α-mannosidase (ASAM), as indicated, and analyzed by HPTLC and fluorography. The conversion of the Man5GlcNAc2 standard to Man4GlcNAc2 is a positive control for the ASAM activity. The uppermost band running ahead of [3H]mannitol in lanes 4 and 6 is free [3H]Man released from the oligosaccharides by ASAM. Panel C: schematic representation of the ASAM-induced interconversions of the radiolabeled oligosaccharides seen in panel B.

Fig. 5. SDS–PAGE and peridate-Schiff staining of wild-type and TbALG3 null mutant procyclins before and after PNGaseF digestion. Lanes 1 and 2 contain wild-type procyclins before and after PNGaseF treatment, respectively. Lanes 3 and 4 contain the TbALG3 null mutant procyclins before and after PNGaseF treatment, respectively. The positions of molecular weight markers are shown on the left.

of their diagnostic C-terminal peptide domains. The aforesaid treatments and their products are summarized at the top of Figure 6 and Table I and the MALDI-TOF spectra are shown in Figure 6A–F. The aq. HF/TFA data are the simplest to interpret, since they show the diagnostic C-terminal domain peptide fragments that define which of the procyclin variants are present in the sample (Acosta-Serrano et al. 1999). From these data, we can see that the wild-type cells are expressing a mixture of EP3, EP1-2 and EP1-1 procyclins, whereas the TbALG3 null mutant is expressing only EP3 and EP1-2 (Figure 6A and B). Such changes in procyclin expression are known in procyclic form T. brucei (Treumann et al. 1997; Morris et al. 2002; Vassella et al. 2004). The aq. HF/PNGaseF data gave the expected result for the wild-type procyclins, i.e., the EP3, EP1-2, and EP1-1 full-length peptides were observed (Figure 6C). However, the TbALG3 null mutant data were unusual in that, although the EP3 full-length peptide is clearly seen (at m/z 8489), the majority of the EP1-2 procyclin was observed as homodimers at m/z 18496 or as EP1-2/EP3 heterodimers at m/z 17777 (Figure 6D). A trace of EP3 homodimer is also seen at m/z 17704. Why the procyclins from the TbALG3 null mutant, and in particular EP1-2, should have this propensity to appear as dimers by MALDI-TOF is unclear. None of the procyclins have cysteine residues, ruling out disulfide formation. The aq. HF only data gave the expected result for the wild-type procyclins, i.e., the EP3, EP1-2, and EP1-1 full-length peptides plus conventional triantennary Man5GlcNAc2 N-linked glycans were observed at m/z 9707, 10417, and 11519, respectively (Figure 6E). On the other hand, the aq. HF only data for the TbALG3 null mutant procyclins were strikingly different. In this case, a major ion at m/z 8492, corresponding to non-N-glycosylated EP3 procyclin was seen together with a cluster of three ions at m/z 9384, 9748, and 10119 that would correspond to EP3 procyclin containing Hex3HexNAc2, Hex3HexNAc3, and
Fig. 6. The MALDI-TOF mass spectra of treated wild-type and TbALG3 null mutant procyclins. Procyclins from wild-type cells (panels A, C, and E) and TbALG3 null mutant cells (panels D, E, and F) were analyzed by negative-ion MALDI-TOF mass spectrometry after treatment with aq. HF and mild acid (panels A and B), aq. HF and PNGaseF (panels C and D), and aq. HF alone (panels E and F). A schematic representation of the products of these treatments for wild-type EP1 procyclin is shown above the spectra.
Table I. Procyclin species observed in Figure 6 by negative-ion MALDI-TOF

<table>
<thead>
<tr>
<th>Procyclin</th>
<th>Mass</th>
<th>Assignment</th>
<th>Wild type</th>
<th>TbALG3 null</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP3</td>
<td>5185</td>
<td>P(EP)3+G</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP3</td>
<td>5397</td>
<td>PDP(EP)3+G</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-2</td>
<td>5864</td>
<td>P(EP)2+G</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-2</td>
<td>6076</td>
<td>PDP(EP)2+G</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-1</td>
<td>6994</td>
<td>P(EP)1+G</td>
<td>+</td>
<td>−</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-1</td>
<td>7206</td>
<td>PDP(EP)1+G</td>
<td>+</td>
<td>−</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-2 dimer</td>
<td>11723</td>
<td>P(EP)3+G × 2</td>
<td>trace</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP1-2 dimer</td>
<td>12152</td>
<td>PDP(EP)2+G × 2</td>
<td>trace</td>
<td>+</td>
<td>Aq, HF + TFA</td>
</tr>
<tr>
<td>EP3</td>
<td>8496</td>
<td>Full-length unglycosylated</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP1-2</td>
<td>9223</td>
<td>Full-length unglycosylated</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP1-1</td>
<td>10304</td>
<td>Full-length unglycosylated</td>
<td>+</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP3 dimer</td>
<td>17041</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP3 + EP1-2 dimer</td>
<td>17777</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP1-2 dimer</td>
<td>18496</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF + PNGaseF</td>
</tr>
<tr>
<td>EP3</td>
<td>9707</td>
<td>Full length + Hex3HexNAc2</td>
<td>+</td>
<td>−</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP1-2</td>
<td>10417</td>
<td>Full length + Hex3HexNAc2</td>
<td>+</td>
<td>−</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP1-1</td>
<td>11519</td>
<td>Full length + Hex3HexNAc2</td>
<td>+</td>
<td>−</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3</td>
<td>8492</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3</td>
<td>9384</td>
<td>Full length + Hex3HexNAc2</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3</td>
<td>9748</td>
<td>Full length + Hex4HexNAc3</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3</td>
<td>10119</td>
<td>Full length + Hex4HexNAc3</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3 + EP1-2 dimer</td>
<td>17723</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP1-2 dimer</td>
<td>18428</td>
<td>Full-length unglycosylated</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3 + EP1-2 dimer</td>
<td>18615</td>
<td>Full-length 1 × unglycosylated + 1 × Hex3HexNAc2</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3 + EP1-2 dimer</td>
<td>19333</td>
<td>Full-length 1 × unglycosylated + 1 × Hex3HexNAc2</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
<tr>
<td>EP3 + EP1-2 dimer</td>
<td>20668</td>
<td>Full-length 1 × unglycosylated + 1 × Hex3HexNAc2</td>
<td>−</td>
<td>+</td>
<td>Aq, HF</td>
</tr>
</tbody>
</table>

“Full length” refers to the entire protein and an ethanolamine linked to the C-terminal glycine.

Hex3HexNAc2 N-linked glycans (Figure 6F). Again, in this spectrum, the EP1-2 procyclin appears predominantly as dimers in nonglycosylated and glycosylated form, containing the same Hex3HexNAc2, Hex4HexNAc3, and Hex5HexNAc4 structures. The combinations of possible dimer glycoforms give rise to the relatively complex peak pattern between m/z 17764 (the unglycosylated EP3/EP1-2 heterodimer) and 20106 (the EP1-2 homodimer with one Hex3HexNAc2 N-linked glycan) (Table I).

The identities of the N-linked glycans in the TbALG3 null mutant procyclins were analyzed by releasing them with PNGaseF followed by permethylation and positive-ion MALDI-TOF mass spectrometry. Ions corresponding to the [M + Na]+ ions of three major species of composition Hex3HexNAc2, Hex4HexNAc3, and Hex5HexNAc4 (together with trace amounts of Hex3HexNAc3, Hex4HexNAc4, and Hex5HexNAc5) were observed (Figure 7A), consistent with the assignments in Table I and Figure 6. The same sample of permethylated glycans was analyzed by electrospray tandem mass spectrometry (ES-MS/MS). Ions at 821.9 and 1047.0, corresponding to [Hex3HexNAc3 + 2Na]+ and [Hex4HexNAc4 + 2Na]+, respectively, were selected and subjected to collision-induced dissociation; the resulting production spectra contained intense m/z 486 ions, characteristic of N-acetylglactosamine (LacNAc) containing species (Figure 7B, C). The lack of product ion at m/z 935 in Figure 7B, corresponding to a LacNAc2 fragment ion, suggests that the Hex3HexNAc3 species is a conventional Galβ1-4GlcNAcβ1-2Manα1-6Galβ1-4GlcNAcβ1-2Manα1-3Manβ1-4GlcNAcβ1-4GlcNAc (NA2-type) complex structure with one LacNAc unit on each branch.

In summary, these data show that whereas wild-type procyclin glycoproteins and process these to conventional triantennary Man9GlcNAc2 oligomannose structures (Treumann et al. 1997; Acosta-Serrano et al. 1999), there is significant under-N-glycosylation and aberrant processing to biantennary complex structures when biantennary Man5GlcNAc2-PP-Dol is the only donor available to the OST(s) present in that life-cycle stage.

VSG N-glycosylation is affected in TbALG3 null mutant bloodstream form T. brucei

VSG glycoprotein was purified (Cross 1984) in its GPI-specific phospholipase C (GPI-PLC) cleaved soluble form (sVSG) (Ferguson et al. 1985) from wild-type and TbALG3 null mutant cells. Analysis by SDS–PAGE and Coomassie blue staining showed that whereas sVSG prepared from the wild-type cells appeared as a single ~54 kDa band, sVSG prepared from the TbALG3 null mutant appeared as a doublet (Figure 8). To assess whether this change in the VSG pattern was due to changes in glycosylation, the wild-type and TbALG3 null mutant sVSGs were analyzed by mass spectrometry.

First, the upper and lower bands of the TbALG3 null mutant sVSG were excised separately from a gel, digested with trypsin, and the resulting peptides were analyzed by MALDI-TOF/TOF. Both bands were clearly identified as VSG221 (data not shown), the same VSG variant as the wild-type cells, suggesting that the differences in the SDS–PAGE pattern are indeed due to the difference in posttranslational modifications of the same sVSG variant.

Next, the positive-ion ES-MS spectra of the intact glycoproteins were collected and deconvolved using a Bayesian
Fig. 7. Mass spectrometric analyses of PNGaseF released and permethylated glycans from \(TbALG3\) null mutant procyclins. Panel A: positive-ion MALDI-TOF of the released and permethylated glycans, revealing a set of \([M+Na]^+\) ions. Panel B: ES-MS/MS product ion spectrum of an \(m/z\) 821.9 \([M+2Na]^{2+}\) ion (corresponding to the \(m/z\) 1620.6 \([M+Na]^+\) ion in panel A). Panel C: The ES-MS/MS product ion spectra of an \(m/z\) 1047.0 \([M+2Na]^{2+}\) ion (corresponding to the \(m/z\) 2069.9 \([M+Na]^+\) ion in panel A). The proposed structures and fragment patterns of the glycans are shown in the panel insets.

Fig. 8. SDS–PAGE of sVSG221 from bloodstream form wild-type and \(TbALG3\) null mutant parasites. Soluble form sVSG221 was purified from wild-type cell cells (lane 1) and \(TbALG3\) null mutant cells (lane 2), subjected to SDS–PAGE, and stained with Coomassie blue. The positions of molecular weight standards are indicated on the left.

Protein reconstruction method. The wild-type sVSG spectrum (Figure 9A) was the same as that described previously for this sVSG variant (Jones et al. 2005; Urbaniak et al. 2006) and the compositions of the isobaric glycoforms, resulting from structural heterogeneity at the two N-glycosylation sites (Zamze et al. 1991) and in the GPI anchor (Mehlert, Richardson, et al. 1998), are described in Table II. The \(TbALG\) null mutant sVSG spectrum (Figure 9B) showed two groups of sVSG glycoforms, consistent with the SDS–PAGE result. The lower molecular weight group of glycoforms has masses consistent with the C-terminal Asn428 N-glycosylation site being unoccupied (Table II). The higher molecular weight group of glycoforms, with both N-glycosylation sites occupied, has significantly higher N-acetylhexosamine to hexose ratios than the wild-type glycoforms, suggesting that the N-glycans contain LacNAc structures at one or both sites (Table II). These data provided the first direct evidence that deletion of the \(TbALG3\) gene affects VSG glycosylation.

Wild-type and \(TbALG3\) null mutant sVSG were also prepared from trypanosomes that had been grown in the presence of a cocktail of \(\alpha\)-mannosidase inhibitors (kifunensine, swainsone, and deoxymannojirimycin that, between them, can inhibit all known ER and Golgi \(\alpha\)-mannosidase activities) and these
Novel aspects of N-glycosylation in Trypanosoma brucei

Fig. 9. sVSG221 glycoform analysis by ES-MS. Samples of sVSG221 from wild-type cells (panels A and C) and TbALG3 null mutant cells (panels B and D) grown in the absence (panels A and B) or presence (panels C and D) of α-mannosidase inhibitors were analyzed by the positive-ion ES-MS and deconvolved spectra of the various isobaric glycoforms were generated. The compositions of these glycoforms are given in Table II.

To ascertain more precisely the changes induced by the deletion of TbALG3 gene (and the effects of MI) on sVSG glycosylation, the N-glycans present at each of the two N-glycosylation sites and the GPI-anchor structures attached to the C-terminal serine residue were determined by ES-MS and ES-MS/MS analysis of Pronase digests. The Pronase glycopeptides were also reanalyzed after digestion with Manα1-2Man-specific Aspergillus saitoi α-mannosidase (ASAM).

From these data, summarized in Figure 10, we can conclude that (i) glycosylation of the internal Asn263 residue is largely unchanged by the deletion of the TbALG3 gene. (ii) The dramatic increase in the Manα-containing glycans attached to the Asn263 site when wild-type and TbALG3 null VSG is made in the presence of the α-mannosidase inhibitor cocktail confirms that biantennary Manα5GlcNAc2 is the precursor of all of the structures found at this glycosylation site. (The presence of some smaller structures suggests that the α-mannosidase block was slightly incomplete in these experiments.) (iii) Although ASAM digestion converts most of the biantennary Manα5GlcNAc2 and Manα5GlcNAc2 structures to Manα3GlcNAc2, some of the structures at Asn263 contain six Hex residues that are not digested by ASAM. This suggests that the HexαHexNAc2 glycan attached to Asn263 is Manα1-6(Glcα1-3Manα1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc. This glycan species that occupies about 10% of the Asn263 site in the wild-type VSG has not previously been noted in mature VSG molecules. On the other hand, similar residual αGlc is quite abundant in the N-linked glycans of L. major and L. mexicana promastigote surface protease glycoprotein, also known as gp63, (Olafson et al. 1990; Ilg et al. 1994). (iv) The most striking change in VSG glycosylation upon TbALG3 deletion is seen at the C-terminal N-glycosylation site where the mixture of conventional triantennary oligomannose Manα5−9GlcNAc2 glycans is replaced by complex biantennary structures. (v) Some
Fig. 10. Representations of wild-type and *TbALG3* null mutant VSG glycoforms. The principle glycans at each site of wild-type (WT) and *TbALG3* null (ALG3 null) VSGs from untreated cells (panel A) and α-mannosidase inhibitor treated cells (panel B) are indicated, together with their approximate relative percentages. The glycans indicated above and marked (+ASAM) are those found after *A. saitoi* Manα1-2Man-specific α-mannosidase treatment.
of the VSG molecules in the \textit{TbALG3} null mutant fail to be N-glycosylated at the C-terminal Asn428 site. (vi) When grown in the presence of \(\alpha\)-mannosidase inhibitors, the complex structures at Asn428 in the \textit{TbALG3} null mutant are replaced by atypical hybrid structures, suggesting that the usual processing of biantennary Man\(_5\)GlcNAc\(_2\) after transfer to Asn428 is the addition of a LacNAc unit to the \(\alpha\)-6 arm of Man\(_5\)GlcNAc\(_2\), either after or during the removal of the two \(\alpha\)-1-2-linked Man residues from the \(\alpha\)-1-3 arm, and the subsequent addition of up to a further 3 LacNAc units. If the two outer \(\alpha\)-1-2-linked Man residues on the \(\alpha\)-1-3 arm are retained (as in the VSG from \(\alpha\)-mannosidase inhibitor treated cells), elaboration appears to be limited to the addition of a single LacNAc unit to the \(\alpha\)-1-6 arm. (vii) There are no qualitative and only minor quantitative changes to the VSG GPI anchor glycan side chains in the \textit{TbALG3} null mutant.

\textbf{Analysis of other glycoproteins by lectin blotting}

As described above, sVSG analysis by mass spectrometry showed that the glycan at the C-terminal (Asn428) N-glycosylation site were changed from oligomannose type to complex type by deletion of the \textit{TbALG3} gene, resulting in an increase in the number of terminal galactose residues on the glycopeptides. We wanted to investigate other glycoproteins to see if this shift from oligomannose-type to complex-type glycans was a general effect of \textit{TbALG3} deletion.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
Protein & I-cP & GlcN & EtNP & HexNAc & Hex & \(M_r\) & Wild type & \textit{TbALG3} null & Wild-type +MI & \textit{TbALG3} null +MI \\
\hline
1 & 1 & 1 & 1 & 2 & 11 & 48,987 & + & − & − & − \\
1 & 1 & 1 & 1 & 2 & 12 & 49,146 & + & − & − & − \\
1 & 1 & 1 & 1 & 3 & 11 & 49,188 & + & − & − & − \\
1 & 1 & 1 & 1 & 2 & 13 & 49,312 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 14 & 49,465 & + & − & − & − \\
1 & 1 & 1 & 1 & 3 & 13 & 49,513 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 14 & 49,676 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 17 & 50,363 & + & − & − & − \\
1 & 1 & 1 & 1 & 5 & 16 & 50,401 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 15 & 50,446 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 18 & 50,525 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 16 & 50,610 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 19 & 50,687 & + & − & − & − \\
1 & 1 & 1 & 1 & 5 & 18 & 50,726 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 17 & 50,773 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 16 & 50,812 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 20 & 50,849 & + & − & − & − \\
1 & 1 & 1 & 1 & 5 & 19 & 50,889 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 17 & 50,975 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 21 & 51,012 & + & − & − & − \\
1 & 1 & 1 & 1 & 5 & 20 & 51,053 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 19 & 51,088 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 18 & 51,137 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 22 & 51,174 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 18 & 51,178 & + & − & − & − \\
1 & 1 & 1 & 1 & 5 & 21 & 51,211 & + & − & − & − \\
1 & 1 & 1 & 1 & 6 & 20 & 51,250 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 19 & 51,291 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 23 & 51,327 & + & − & − & − \\
1 & 1 & 1 & 1 & 8 & 18 & 51,340 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 20 & 51,453 & + & − & − & − \\
1 & 1 & 1 & 1 & 4 & 24 & 51,500 & Trace & + & + & + \\
1 & 1 & 1 & 1 & 8 & 19 & 51,502 & + & − & − & − \\
1 & 1 & 1 & 1 & 7 & 21 & 51,617 & + & − & − & − \\
1 & 1 & 1 & 1 & 9 & 19 & 51,704 & + & − & − & − \\
1 & 1 & 1 & 1 & 9 & 20 & 51,869 & + & − & − & − \\
1 & 1 & 1 & 1 & 10 & 22 & 52,397 & + & − & − & − \\
\hline
\end{tabular}
\caption{Isobaric glycoforms of sVSG221 detected by ES-MS}
\end{table}

*Protein \(M_r\) is based on the amino acid sequence of the VSG221 precursor (accession no. P26332) minus residues 1–27 (signal peptide) and 460–476 (GPI attachment signal peptide) and allows for four disulfide bonds (\(M_r = 46,284\)).

*Components specific to the GPI anchor and common to all glycoforms; I-cP myo-inositol-1,2 cyclic phosphate; EtNP, ethanolamine phosphate.

*The most abundant glycoform of wild-type sVSG221 is expected to contain a GPI anchor of composition of Man3Gal5, a C-terminal attachment signal peptide) and allows for four disulfide bonds (\(M_r = 46,284\)).

The molecular weights of different glycoforms of sVSG221 were calculated according to the indicated compositions. The +, −, trace, +++, +++, ++++ scores indicate the relative abundances of those glycoforms observed in sVSG preparations of wild-type and \textit{TbALG3} null mutants with and without mannosidase inhibitors (MI).
and Bhattacharyya 1986; Naismith and Field 1996). The reduction in ConA binding to the \( TbaL3 \) null mutant glycoproteins is consistent with the expected general ablation of oligosaccharide (i.e., triantennary \( \text{Man}_3\text{GlcNAc}_2 \)) structures. The same lysates were subjected to blotting with ricin, a lectin that binds terminal Gal residues, except this time approximately three times more wild type than \( TbaL3 \) null mutant cell lysate was loaded and HRP-conjugated ricin and a galactose/lactose mixture were used as the lectin probe and lectin inhibitor, respectively.

Taken together, these data are consistent with a general up-regulation of Gal-terminating complex glycans at the expense of conventional oligomannose structures in the \( TbaL3 \) null bloodstream form trypanosomes.

**Discussion**

Deletion of the \( TbaL3 \) gene in bloodstream and procyclic form \( T. brucei \) has no significant effect on their growth, morphology or the infectivity of the parasite to mice and tsetse flies, respectively. We may, therefore, conclude that the Dol-P-Man-dependent \( \alpha_1-3 \) mannosyltransferase encoded by \( TbaL3 \) is not useful drug target against African trypanosomiasis. However, the retention of \( ALG3, ALG12, \) and \( ALG9 \) genes in \( T. brucei \) indicates that the ability to make \( \text{Man}_9\text{GlcNAc}_2 \)-PP-Dol presumably has some selective advantage for the organism and we cannot rule out a role for \( TbaL3 \) in other life-cycle stages or in parasite differentiation processes.

The glycosylation phenotypes of the null mutant parasites supports the model of site-selective N-glycosylation, originally proposed by Bangs et al. (1988) and elaborated in Jones et al. (2005). That model predicts that preventing the synthesis of dolichol cycle intermediates beyond \( \text{Man}_5\text{GlcNAc}_2 \)-PP-Dol should have little or no effect on the glycosylation and subsequent processing of the Asn263 site of VSG221, since this site normally receives \( \text{Man}_9\text{GlcNAc}_2 \) from \( \text{Man}_5\text{GlcNAc}_2 \)-PP-Dol in wild-type cells. This is indeed the case, apart from an increase (from approximately 10 to 19 mol%) in \( \text{Glc}_1\text{Man}_5\text{GlcNAc}_2 \) at this site. The latter effect may be due to the aberrant and underglycosylation that is observed at the C-terminal Asn263 site, perhaps causing greater activity of UGGT in the UGGT/calreticulin/\( \alpha \)-glucosidase II-mediated refolding and quality control cycle.

The underglycosylation of the Asn263 site of VSG221 in the \( TbaL3 \) null mutant is significant because it provides evidence that the OST(s) that normally recognize this sequon have a preference for \( \text{Man}_9\text{GlcNAc}_2 \)-PP-Dol and that, when this donor is absent, they utilize \( \text{Man}_5\text{GlcNAc}_2 \)-PP-Dol relatively poorly. Exactly the same phenomenon is seen in the procyelic form of the parasite, where the procyelins that normally receive \( \text{Man}_9\text{GlcNAc}_2 \) are similarly underglycosylated. These observations support a refined model (Figure 12) whereby two different classes of OST activity exist in \( T. brucei \): one, that we will refer to as TbOST-1, that recognizes sites like Asn263 in VSG221 and preferentially utilizes \( \text{Man}_9\text{GlcNAc}_2 \)-PP-Dol and another, that we will refer to as TbOST-2, that recognizes sites like Asn428 in VSG221 and those of EP3 and EP1-2 procyclin and preferentially utilizes \( \text{Man}_9\text{GlcNAc}_2 \)-PP-Dol. The genome sequencing supports this model in so far as \( T. brucei \) has three OST catalytic subunit \( STT3 \) genes (although, intriguingly, it lacks candidate genes for all other known OST subunits [Kelleher and Gilmore 2006]). The three \( STT3 \) genes are found in a tandem array at the end of chromosome 5 and two of them encode almost identical proteins. We speculate that the unique \( STT3 \) utilizes exclusively either \( \text{Man}_9\text{GlcNAc}_2 \)-PP-Dol or \( \text{Man}_5\text{GlcNAc}_2 \)-PP-Dol and that one or both of the two similar \( STT3 \)s utilize the other donor. In this model, the different VSG glycosylation sites are predicted to recruit the appropriate TbOST, and thus dictate the glycan transferred to that site. The fact that posttranslational N-glycosylation of Endo-H-sensitive (but not Endo-H-resistant) sites has been observed for two different VSG variants (Ferguson et al. 1986; Bangs et al. 1988) also provides support for this model and further suggests that TbOST-1 is associated with co-translational glycosylation whereas TbOST-2 is associated with posttranslational glycosylation. Testing of this revised model awaits the construction and characterization of \( STT3 \) mutants and a (glyco)proteome-wide analysis of the occupancy of \( T. brucei \) N-glycosylation sites with Endo-H-sensitive and resistant glycans to identify the flanking peptidic features that are recognized by TbOST-1 and TbOST-2. This work is currently in progress in our laboratory.

The other key result for this study is the processing of \( \text{Man}_9\text{GlcNAc}_2 \) (i.e., aberrantly transferred to “Man5\text{GlcNAc}_2” sites in the \( TbaL3 \) null mutants) to complex biantennary structures on both Asn428 of VSG221 and the N-glycosylation sites of EP3 and EP1-2 procyclin. This supports the notion (Jones et al. 2005) that \( \text{Man}_9\text{GlcNAc}_2 \) is destined for processing into complex structures and, importantly, suggests that there is nothing inherent in these sites that prevent their processing to complex structures.

This, in turn, begs the question as to why sites that normally receive \( \text{Man}_9\text{GlcNAc}_2 \) can only be processed to other Endo-H-sensitive oligomannose structures (i.e., triantennary \( \text{Man}_3\text{GlcNAc}_2 \)). The answer may lie partly in the absence of a putative \( T. brucei \) Golgi \( \alpha \)-mannosidase II gene. Golgi \( \alpha \)-mannosidase II normally removes the \( \alpha_1-3 \) and \( \alpha_1-6 \)-linked Man residues from the 6-arm of the trimannosyl core to provide a substrate for elaboration by GnT-II into complex...
Novel aspects of N-glycosylation in *Trypanosoma brucei*

**Fig. 12.** Model of the site-specific protein N-glycosylation and N-glycan processing in *T. brucei*. Panel A illustrates the proposed existence of two distinct oligosaccharyltransferase activities, one (TbOST-1) that cotranslationally transfers Man$_5$GlcNAc$_2$ to specific glycosylation sites (e.g., Asn263 of VSG221) and the other (TbOST-2) that preferentially transfers Man$_9$GlcNAc$_2$ posttranslationally to any remaining sites (e.g., Asn428 of VSG221 and procyclin sites). Panel B illustrates the processing of Man$_9$GlcNAc$_2$ and Man$_5$GlcNAc$_2$ in VSG221 and procyclin in wild-type trypanosomes. The structures in boxes are those that have been experimentally observed at the respective glycosylation sites in the mature glycoproteins. The inability of *T. brucei* to process Man$_9$GlcNAc$_2$ beyond triantennary Man$_5$GlcNAc$_2$ is thought to be due to unusual specificity of TbGnT-I and the absence of Golgi α-mannosidase II. Panel C illustrates the processing of Man$_5$GlcNAc$_2$ in VSG221 and procyclin in *TbALG3* null mutant trypanosomes. Structures in boxes are those that have been experimentally observed at the respective glycosylation sites in the mature glycoproteins. The +/- symbol indicates inefficient oligosaccharide transfer of Man$_5$GlcNAc$_2$ by TbOST-2 that normally transfers Man$_9$GlcNAc$_2$. The processing of Man$_5$GlcNAc$_2$ to complex biantennary structures at Asn428 of VSG221 supports the model that sites that receive Man$_5$GlcNAc$_2$ are destined to be processed to complex structures in *T. brucei*.
structures (Schachter 2000). However, this alone does not explain why Gnt-I activity could not convert triantennary Man9GlcNAc2 into conventional hybrid structures. However, whereas eukaryotic Gnt-I enzymes typically work on triantennary Man9GlcNAc2, analysis of ConA-resistant procyclin-processed glycan structures (Leal et al. 2004). In the null mutant (and in ConA-resistant mutants), the procyclin-processed glycan structures. Combining these observations leads to the model shown in Figure 12 that satisfies most or all of the current experimental data. For example, it explains why the effects of TbALG3 and TbALG12 deletion are similar with respect to procyclin underglycosylation (because TbOST2 prefers Man9GlcNAc2 to procyclin underglycosylation (because TbOST2 prefers Man9GlcNAc2, analysis of ConA-resistant procyclin underglycosylation (because TbOST2 prefers Man9GlcNAc2, analysis of ConA-resistant procyclin-processed glycan structures (Leal et al. 2004). In the TbALG12 null mutant (and in ConA-resistant mutants), the procyclin-linked Man9GlcNAc2 structure can be processed by ER α1–2 mannosidases (Man9GlcNAc2-α1–2GlcNAc (Man9GlcNAc2) and further elaborated on the 3- arm of the trimannosyl core by a LacNac unit (Acosta-Serrano et al. 2000; Hwa and Kho 2000; Leal et al. 2004) whereas in the TbALG3 mutant, the Manα1–6(Manα1–2Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc (Man9GlcNAc2) structure attached to procyclin is processed by ER α1–2 mannosidases (Manα1–6(Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc (Man9GlcNAc2) that can receive LacNac units on both arms of the trimannosyl core.

In summary, the revised model presented here serves to highlight fundamental differences in the mechanisms of protein N-glycosylation and glycan processing between T. brucei and its mammalian hosts. Further study may reveal theoretically exploitable differences.

Materials and methods

Cultivation of trypanosomes

Bloodstream form T. brucei strain 427 that has genetically been modified to express T7 polymerase and the tetracycline repres sor protein, referred to as wild type for convenience, were cultured in the HMI-9 medium (Hirumi H and Hirumi K 1994) with the addition of 2.5 µg/mL G418 at 37°C in a 5% CO2 incubator. For some experiments the trypanosomes were grown for 48 h in the presence of a cocktail of α-mannosidase inhibitors as previously described (Jones et al. 2005). Procyclin T. brucei strain 427 clone 29.13, referred to as wild-type culture, were cultured in SDM-79 (Brun and Schonenberger 1979) in the presence of 15% fetal bovine serum (FBS) and 7.5 mg/L haemin (Wirtz et al. 1999) together with 15 µg/mL G418 and 50 µg/mL hygromycin in a 28°C incubator. The concentrations of drugs used for the selection of mutants were 4 µg/mL for hygromycin and 0.1 µg/ mL for puromycin for bloodstream form cells and 10 µg/mL for puromycin for procyclic form cells.

DNA manipulation

Plasmid DNA was purified using Miniprep or Maxiprep kits (Qiagen, Crawley, UK). Gel extraction and cleanup were performed with a Qiaquick kit (Qiagen). Genomic DNA was isolated from ~5 × 10^8 bloodstream form and procyclic form cells using DNAzol (Helena Biosciences, Gateshead, UK).

Cloning and sequencing of the TbALG3 ORF

The 1214 bp ORF was amplified by PCR from genomic DNA using Pfu polymerase with 5'-CCCAAGCTTATGGGTATT-CGTTGGCTG-3' forward and 5'-GGGATCCATTTTCCCTTTTACAG-3' reverse primers using the cycling parameters 94°C for 5 min, 30 cycles of 94°C for 30 s, 52°C for 2 min, and 68°C for 30 s, followed by 10 min extension at 68°C. Four separate PCR products were purified, cloned into pCR-Blunt II TOPO (Invitrogen, Paisley, UK), and sequenced twice in both directions.

Gene replacement constructs

To generate the T. brucei gene replacement cassettes, 500 bp 5'-UTR and 500 bp 3'-UTR immediately adjacent to the TbALG3 ORF were amplified by PCR from T. brucei genomic DNA using Pfu polymerase with primers 5'-ATAAGGA-TGGCGCGCCTCATTTTTTTGATTTGGTACCTC-3' and 5'-GGTAAAACTTACGGACCCTCAAGTCTTTTACTGTT-GGGGAC-3'. For the 5'-UTR, and primers 5'-GACGTT-CCTAATTACGATCCACGAAACGTCGATGG-GGGGAC-3' for the 5'-UTR, and primers 5'-GACGTT-CCTAATTACGATCCACGAAACGTCGATGG-GGGGAC-3'. The products from the above PCR reactions were joined together in a further PCR via a short BamHI–Pmel–HindIII linker region contained within the primers and a NotI site (underlined) at each end. This PCR product was ligated into pGEM-SZII (+) (Promega) via the NotI sites. Antibiotic resistance markers were ligated between the BamHI and HindIII restriction sites, to produce constructs containing puromycin acetyltransferase (PAC), hygromycin phosphotransferase (HYG) and bactericidal deaminase (BSD) resistance genes. After sequencing, plasmids were purified and digested with NotI. The digestion products were precipitated and washed in 70% ethanol, dissolved in sterile distilled water, and used for electroporation to sequentially replace TbALG3 alleles by homologous recombination (Wirtz et al. 1999).

Southern blotting

T. brucei genomic DNA, prepared from 100 mL cultures of bloodstream form and 10 mL cultures of procyclic form cells, was digested with the various restriction enzymes overnight and resolved on a 0.8% agarose gel and transferred to a Hybond-N positively charged membrane (GE Healthcare, Amersham, UK). The membrane was hybridized overnight in ULTRA-HYB (Ambion) at 42°C with the fluorescein-labeled TbALG3 ORF probe (CDP-Star random prime labeling kit, GE Healthcare). After two 20-min washes (once with 2 × SSC, 0.1% SDS and once with 0.2 × SSC, 0.1% SDS) the probe was detected by using the CDP-Star detection system.

Tsetse fly infections

Pupae of Glossina morsitans were obtained from Institute of Zoology, Slovak Academy of Science (Bratislava, Slovakia). Newly hatched (teneral) flies were fed with an infected bloodmeal, which consisted of 10^7 parasites mixed with washed defribinated horse blood (containing 10% FBS). Infected flies were fed with bloodmeals every 2–3 days. After 2 weeks or 3 weeks, midguts were isolated from infected flies and disrupted by
mechanical force in cold SDM-79 containing 10% FBS. Isolated parasites from individual midguts were kept on ice until counted on a haemocytometer.

**Preparation and radiolabeling of trypanosome cell free systems**

Trypanosome membranes were prepared from wild-type cells and ThALG3 null mutant cells as described by Masterson et al. (1989). Cell lysates were thawed on ice and washed twice with a HKML buffer (50 mM HEPES, pH 7.4, 25 mM KCl, 5 mM MgCl₂, 1 μg of leupeptin/mL, 0.1 mM TLCK) and resuspended at a concentration of 1 × 10⁷/mL in 2 × incorporation buffer (HKML buffer with 5 mM MnCl₂ and 1 mM DTT). The membranes (2 × 10⁷ cell equivalents) were transferred to another tube containing GDP-[3,4-³H]Man and 2 mM UDP-GlcNAc for 5 min at 30°C. The reaction was chased with 1 mM nonradioactive GDP-Man for 20 min at 30°C. Reactions were terminated by adding chloroform and methanol to reach a final CHCl₃:CH₃OH:H₂O ratio of 10:10:3 (v/v/v) and sonicated in a sonicating water bath for 10 min. After overnight extraction, the chloroform/methanol/water extract was centrifuged and the supernatant dried under a stream of nitrogen and partitioned between butan-1-ol and water, as described in Acosta-Serrano et al. (2004). The labeled glycolipid products were treated with 100 μL of 0.1 M TFA, 98°C, 15 min, cooled on ice and extracted twice with butan-1-ol to remove labeled GPIs. The aqueous phase, containing the acid-released glycans from the Dol-PP-oligosaccharides, were dried and redissolved in 5 μL of 0.25 M sodium phosphate buffer (pH 7.5). The samples were loaded into nanotips (Micromass-type F) and analyzed by ES-MS in positive-ion mode on an ABI Q-Star-XL spectrometer. Data were collected manually at 500 shots per spectrum with the laser intensity set at 2000 V.

**sVSG isolation**

Soluble form VSG was isolated from 2 × 10⁸ bloodstream form T. brucei cells as described in Cross (1975, 1984). Briefly, cells were chilled on ice for 10 min, centrifuged at 2500 × g for 10 min, washed in an isotonic buffer and resuspended in 300 μL of 10 mM sodium phosphate buffer, pH 8.0, containing 0.1 mM TLCK, 1 μg/mL leupeptin, and 1 μg/mL aprotinin. After 5 min at 37°C, the mixture was cooled on ice and centrifuged (14,000 × g, 5 min). The supernatant was applied to a small (0.2 mL) DE52 anion exchange column equilibrated in a 10 mM sodium phosphate buffer, pH 8.0 and eluted with 0.8 mL of the same buffer. The entire column eluate (containing about 100 μg sVSG) was concentrated and dialyzed with water on a Microcon YM-10 concentrator (Millipore, Watford, UK) and recovered in 100 μL of water.

**ES-MS analysis of intact VSG**

sVSG obtained from small scale VSG purification was diluted to ∼0.05 μg/μL in 50% methanol, 1% formic acid, loaded into nanotips (Micromass-type F). The positive-ion electrospray mass spectra were recorded on the ABI QSTAR-XL system with tip potential of 900 V and declustering potential of 60 V. Data were deconvolved by using the ABI Analyst software and using the Bayesian protein reconstruct program.

**ES-MS and ES-MS/MS analysis of Pronase glycopeptides**

Aliquots of sVSG (approximately 50 μg in 50 μL water) were mixed with 5 μL of 1 M ammonium bicarbonate and 10 μL of 1 mg/mL Pronase in 5 mM calcium acetate and incubated at 37°C for 36 h. The Pronase glycopeptides were purified on Envi-carb graphitized carbon microcolumns which were prepared as follows: the contents of an Envi-carb cartridge (Supelco, Poole, UK) were suspended in methanol and a bed of approximately 20 μL of graphitized carbon was packed into a 100 μL C4 OMI/X (Varian, Oxford, UK) pipette tip. The microcolumns were prepared by attaching them to a Gilson pipette, set at 100 s, chilled on ice for 10 min, centrifuged at 2500 × g, and pipetting up and down 10 times with 80% methanol, 1% formic acid; 60% methanol, 1% formic acid; and 1% methanol, 1% formic acid. The sample (10 μL of Pronase digest) was mixed with 90 μL of 1% methanol, 1% formic acid and applied to the microcolumn by pipetting up and down 20 times. The microcolumns were washed by pipetting up and down 20 times with 1% methanol, 1% formic acid. The pipette was reset to 50 μL and the glycopeptides eluted by pipetting up and down 20 times with 50 μL of 60% methanol, 1% formic acid. Aliquots of these samples were loaded into nanotips (Micromass-type F) and analyzed by ES-MS in positive-ion mode on an ABI Q-Star-XL instrument with tip and declustering potentials of 900 V and 60 V respectively. The product ion spectra of selected ions were collected using collision energies of 30–60 V. The ES-MS spectra were processed using the Bayesian peptide reconstruction program in the ABI Analyst software. After analysis, the
remaining material was dried, redissolved in 5 µL of 0.5 M sodium acetate buffer (pH 5.0) and digested with 5 µL of Aspergillus saitoi α-mannosidase (ASAM), 16 h, 37°C, re-purified on graphitized carbon (as described above) and re-analyzed by ES-MS and ES-MS/MS.

The ES-MS spectra of mostly proton and/or sodium-adducted doubly charged ions, with some triply charged ions for the larger glycopeptides and singly charged ions for the smaller glycopeptides (supplementary Table S1), were processed by Bayesian peptide reconstruction to produce the spectra shown in supplementary Figure S4, A–H and supplementary Table SII, A–H. The majority of these ions could be assigned to the internal Asn263 N-glycosylation site, the C-terminal Asn428 N-glycosylation site or the C-terminal Ser433 GPI-peptide according to their masses (supplementary Table SII, A–H), their characteristic MS/MS product ion spectra (see supplementary Figure S5, A–C for examples) and the known structural glycoforms of VSG221 (Zamze et al. 1991; Mehliert, Richardson, et al. 1998; Jones et al. 2005). For the N-linked glycopeptides, the relatively abundant [GlcNAc-peptide + H]+ MS/MS fragment ions were particularly useful to assign glycans to their correct peptidic sites, which appeared as NET and RNET peptides for Asn263 and as NT, NTT, and TNTT peptides for Asn428. The GPI-peptides appeared exclusively as Ser-ethanolamine-P-Man3-GlcN-inositol-1,2,5-cyclic-P fragments with between two and six Gal side-chain residues.

An estimate of the relative molar abundances of individual glycans at each site was assigned by adding together all of the intensities in the deconvolved spectra for a particular glycoform (e.g., intensities for (NET)-Hex3HexNAc2, (N E T-intensities in the deconvolved spectra for a particular glycan at each site was assigned by adding together all of the and six Gal side-chain residues.

**Funding**

Wellcome Trust Programme (grant 071463 to M.A.J.F).

**Acknowledgements**

We thank Craig Lapsley for processing and maintaining tsetse flies and Angela Mehliert for helpful suggestions and assistance in preparing bloodstream form trypanosomes. S.M. was supported by a BBRC PhD studentship, A.A.S by a Wellcome Trust Research Career Development Fellowship, and I.I. by a Marie Curie Fellowship.

**Conflict of interest statement**

None declared.

**Abbreviations**

ASAM, Aspergillus saitoi a1-2-specific α-mannosidase; ConA, concanavalan A; Dol-PP, dolichol pyrophosphate; ER, endoplasmic reticulum; ES-MS/MS, electrospray tandem mass spectrometry; FBS, fetal bovine serum; GPI, glycosylphosphatidylinositol; HPTLC, high-performance thin layer chromatography; LacNAc, N-acetyllactosamine; MI, α-mannosidase inhibitors, GPI-PLC, GPI-specific phospholipase C; OST, oligosaccharyltransferase; PI, phosphatidylinositol; PNGaseF, peptide N-glycosidase F; sVSG, soluble form VSG; TFA, trifluoroacetic acid; UGGT, unfolded glycoprotein glucosyltransferase; VSG, variant surface glycoprotein.

**References**


